Modelling. Virtual. Realities. A Practical Introduction to Virtual (and Augmented) Reality

Zoe Schubert
Jan G. Wieners
Welcome to DHSI 2019!

Thanks for joining the DHSI community!

In this booklet, you will find essential course materials prefaced by some useful information about getting settled initially at UVic, finding your way around, getting logged in to our network (after you’ve registered the day before our courses begin), and so on.

Given our community’s focus on things computational, it will be a surprise to no one that we might expect additional information online for some of the classes - your instructors will let you know - or that the most current version of all DHSI-related information may be found on our website at dhsi.org.

Do check in there first if you need anything that’s not in this coursepak.

To access the DHSI wifi network, simply go into your wireless settings and connect to the “DHSI” network and enter the password “dhsi2019”.

And please don’t hesitate to be in touch with us at institut@uvic.ca or via Twitter at @AlyssaA_DHSI or @DHInstitute if we can be of any help....
DHSI Wi-Fi

Network name: DHSI
Passkey: dhsi2019
The 2019 schedule is just taking shape nicely! A very few things to confirm, add, etc, still but this is the place to be to find out what is happening when / where ...

Psst: Some Suggested Outings

If you're here a day or two before we begin, or staying a day or two afterwards, here are a few ideas of things you might consider doing ....

Suggested Outing 1, Botanical Beach (self-organised; car needed)
A self-guided visit to the wet, wild west coast tidal shelf (and historically-significant former research site) at Botanical Beach; we recommend departing early (around 8.00 am) to catch low tide for a better view of the wonderful undersea life! Consider bringing a packed lunch to nibble-on while looking at the crashing waves when there, and then have an afternoon drink enjoying the view from the deck of the Port Renfrew Hotel.

Suggested Outing 2, Butchart Gardens (self-organised)
A shorter journey to the resplendently beautiful Butchart Gardens and, if you like, followed by (ahem) a few minutes at the nearby Church and State Winery, in the Saanich Peninsula. About an hour there by public bus from UVic, or 30 minutes by car.

Suggested Outing 3, Salt Spring Island (self-organised; a full day, car/bus + ferry combo)
Why not take a day to explore and celebrate the funky, laid back, Canadian gulf island lifestyle on Salt Spring Island. Ferry departs regularly from the Schwartz Bay ferry terminal, which is about one hour by bus / 30 minutes by car from UVic. You may decide to stay on forever ....

Suggested Outing 4, Paddling Victoria's Inner Harbour (self-organised)
A shorter time, seeing Victoria's beautiful city centre from the waterways that initially inspired its foundation. A great choice if the day is sunny and warm. Canoes, kayaks, and paddle boards are readily rented from Ocean River Adventures and conveniently launched from right behind the store. Very chill.

And more!

Self-organised High Tea at the Empress Hotel, scooter rentals, visit to the Royal BC Museum, darts at Christies Carriage House, a hangry breakfast at a local diner, whale watching, kayaking, brew pub sampling (at Spinnaker's, Swans, Moon Under Water, and beyond!), paddle-boarding, a tour of used bookstores, and more have also been suggested!

9:00 to 4:00

Early Class Meeting: 4. [Foundations] DH For Department Chairs and Deans (David Strong Building C124, Classroom)
Further details are available from instructors in mid May to those registered in the class. Registration materials will be available in the classroom.

3:00 to 5:00

DHSI Registration (MacLaurin Building, Room A100)
After registration, many will wander to Cadboro Bay and the pub at Smuggler's Cove OR the other direction to Shelbourne Plaza and Maude Hunter's Pub OR even into the city for a nice meal.

Monday, 3 June 2019

Your hosts for the week are Alyssa Arbuckle, Ray Siemens, and Jannaya Friggstad Jensen.

7:45 to 8:15

Last-minute Registration (MacLaurin Building, Room A100)
<table>
<thead>
<tr>
<th>Time</th>
<th>Event Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:30 to 10:00</td>
<td>Welcome, Orientation, and Instructor Overview (MacLaurin A144)</td>
</tr>
<tr>
<td></td>
<td>• Welcome to the Territory</td>
</tr>
<tr>
<td></td>
<td>• Welcome to DHSI: Ray Siemens, Alyssa Arbuckle</td>
</tr>
<tr>
<td></td>
<td>• Welcome from UVic: Jonathan Bengtson (University Librarian), Alexandra D'Arcy (Associate Dean Research, Humanities)</td>
</tr>
<tr>
<td></td>
<td>Classes in Session (click for details and locations)</td>
</tr>
<tr>
<td></td>
<td>• 1. [Foundations] Digitisation Fundamentals and their Application (Clearihue A103, Lab)</td>
</tr>
<tr>
<td></td>
<td>• 2. [Foundations] Introduction to Computation for Literary Criticism (Clearihue A102, Lab)</td>
</tr>
<tr>
<td></td>
<td>• 3. [Foundations] Making Choices About Your Data (Digital Scholarship Commons, McPherson Library A308, Classroom)</td>
</tr>
<tr>
<td></td>
<td>• 4. [Foundations] DH For Department Chairs and Deans (David Strong Building C124, Classroom)</td>
</tr>
<tr>
<td></td>
<td>• 5. [Foundations] Developing a Digital Project (With Omeka) (Clearihue A031, Lab)</td>
</tr>
<tr>
<td></td>
<td>• 6. [Foundations] Race, Social Justice, and DH: Applied Theories and Methods (Cornett A229, Classroom)</td>
</tr>
<tr>
<td></td>
<td>• 8. [Foundations] Fundamentals of Programming/Encoding for Human(s)ists (Clearihue A108, Lab)</td>
</tr>
<tr>
<td></td>
<td>• 9. Out-of-the-Box Text Analysis for the Digital Humanities (Human and Social Development A160, Lab)</td>
</tr>
<tr>
<td></td>
<td>• 10. Sound and Digital Humanities (Cornett A120, Classroom)</td>
</tr>
<tr>
<td></td>
<td>• 11. Critical Pedagogy and Digital Praxis in the Humanities (Clearihue D132, Classroom)</td>
</tr>
<tr>
<td></td>
<td>• 12. Digital Humanities for Japanese Culture: Resources and Methods (McPherson Library A003, Classroom)</td>
</tr>
<tr>
<td></td>
<td>• 13. Conceptualising and Creating a Digital Edition (McPherson Library 210, Classroom)</td>
</tr>
<tr>
<td></td>
<td>• 15. Retro Machines &amp; Media (McPherson Library 129, Classroom)</td>
</tr>
<tr>
<td></td>
<td>• 16. Geographical Information Systems in the Digital Humanities (Clearihue A105, Lab)</td>
</tr>
<tr>
<td></td>
<td>• 17. Introduction to IIIF: Sharing, Consuming, and Annotating the World’s Images (Cornett A121, Classroom)</td>
</tr>
<tr>
<td></td>
<td>• 18. Web APIs with Python (Human and Social Development A170, Lab)</td>
</tr>
<tr>
<td></td>
<td>• 19. Ethical Data Visualization: Taming Treacherous Data (Cornett A128, Classroom)</td>
</tr>
<tr>
<td></td>
<td>• 20. Linked Open Data and the Semantic Web (Cornett A132, Classroom)</td>
</tr>
<tr>
<td></td>
<td>• 21. Palpability and Wearable Computing (McPherson Library A025, Classroom)</td>
</tr>
<tr>
<td></td>
<td>• 22. The Frontend: Modern JavaScript &amp; CSS Development (Clearihue A030, Lab)</td>
</tr>
<tr>
<td></td>
<td>• 23. Modelling. Virtual. Realities. A Practical Introduction to Virtual (and Augmented) Reality (Human and Social Development A150, Lab)</td>
</tr>
<tr>
<td></td>
<td>• 25. Information Security for Digital Researchers (David Strong Building C114, Classroom)</td>
</tr>
<tr>
<td>10:15 to Noon</td>
<td>Lunch break / Unconference Coordination Session (MacLaurin A144)</td>
</tr>
<tr>
<td></td>
<td>(Grab a sandwich and come on down!)</td>
</tr>
<tr>
<td></td>
<td>Discussion topics, scheduling, and room assignments from among all DHSI rooms will be handled at this meeting.</td>
</tr>
<tr>
<td>12:15 to 1:15</td>
<td>Classes in Session</td>
</tr>
<tr>
<td>1:30 to 4:00</td>
<td>Institute Lecture: Jacqueline Wernimont (Dartmouth C): &quot;Sex and Numbers: Pleasure, Reproduction, and Digital Biopower&quot;</td>
</tr>
<tr>
<td></td>
<td>Chair: Anne Cong-Huyen (U Michigan)</td>
</tr>
<tr>
<td></td>
<td>(MacLaurin A144)</td>
</tr>
<tr>
<td>4:10 to 5:00</td>
<td>Abstract: Drawing from Numbered Lives (MIT 2018), this talk will consider a long history of sex-number entanglement in Anglo-American Cultures. Drawing on historical and contemporary objects and practices, Wernimont will ask &quot;in what ways do theories of biopower, critical gender and critical race studies, and media studies&quot; suggest that we can understand this set of entanglements and their impacts. NB: While relevant, this talk will not include discussions of sexual trauma or violence. It will include frank discussion of sex acts and various ways of translating sexual behavior into numbers.</td>
</tr>
<tr>
<td>5:00 to 6:00</td>
<td>Opening Reception (University Club)</td>
</tr>
<tr>
<td>9:00 to Noon</td>
<td>Classes in Session</td>
</tr>
<tr>
<td>12:15 to 1:15</td>
<td>Lunch break / Unconference</td>
</tr>
<tr>
<td></td>
<td>&quot;Mystery&quot; Lunches</td>
</tr>
<tr>
<td>1:30 to 4:00</td>
<td>Classes in Session</td>
</tr>
<tr>
<td></td>
<td>DHSI Conference and Colloquium Lightning Talk Session 1 (MacLaurin A144)</td>
</tr>
</tbody>
</table>
Wednesday, 5 June 2019

9:00 to Noon

Classes in Session

Lunch break / Unconference

"Mystery" Lunches

Presentation: An Introduction to Scholarly Publishing with Manifold (MacLaurin A144)
Lunch included for those who [register here]

This presentation introduces Manifold Scholarship, a Mellon-funded digital publishing platform developed by the CUNY Graduate Center, The University of Minnesota Press, and Cast Iron Coding. Manifold allows you to create beautiful, dynamic open access projects that can include text, images, video, embedded resources, and social annotation. We will provide an overview of Manifold and demonstrate how faculty, students and staff in the digital humanities can use Manifold to publish open access scholarly works, conduct and participate in peer review, and create custom edited versions of public domain course texts and OER.

1:30 to 4:00

Classes in Session

DHSI Conference and Colloquium Lightning Talk Session 2 (MacLaurin A144)
Chair: Kim O'Donnell (Simon Fraser U)

- Catherine Ryu (Michigan State U), "Tone Perfect: Developing a Multimodal Audio Database for Mandarin Chinese as an Open Source"
- Kenzie Burchell (U Toronto Scarborough), "Making Responsible Reporting Practices Visible: Comparing newswire coverage of humanitarian crises in Syria"
- Jessica Linzel (Brock U), "The Shopkeeper Aristocracy: Mapping Trade Networks in Colonial Niagara"
- Kirsten Painter (U Washington), "From Bogatyr to Bread: Digitization & Online Exhibition of Rare Russian Children's Books at the U Washington"
- John Barber (Washington State U), "A Mighty Span"

4:15 to 5:15

DHSI Conference and Colloquium Lightning Talk Session 3 (MacLaurin A144)
Chair: Kim O'Donnell (Simon Fraser U)

- Colleen Kolba (U South Florida), "What Comics can Teach our Students about Multimodal Literacy"
- Trish Baer (ETCL; U Victoria), "Preserving Digital Legacies: Archived Websites and Digital Discoverability"
- Suchismita Dutta (U Miami), "The Importance of Archival Transcription for Genre Building"
- Jeffrey Lawler (California State U, Long Beach), "Twining our way through the Past: Video Game Authoring as History Pedagogy"

6:00 to 7:00

"Half Way There!" [An Informal, Self-Organized Birds of a Feather Get-Together] (Felicitas, Student Union Building)
Bring your DHSI nametag and enjoy your first tipple on us! [A great opportunity for an interest group meet-up ....]

Thursday, 6 June 2019

9:00 to Noon

Classes in Session

12:15 to 1:15

"Mystery" Lunches

[Istructor lunch meeting]

1:30 to 4:00

Classes in Session

DHSI Conference and Colloquium Lightning Talk Session 3 (MacLaurin A144)
Chair: Kim O'Donnell (Simon Fraser U)

- Marion Grant (Ryerson U), "Visualizing Networks: Yellow Nineties Print and Performance"
- Megan Perram (U Alberta), "Configuring the Postdigital Body Through the Digital Illness Narratives of Women with Polycystic Ovarian Syndrome"
- Kristen Starkowski (Princeton U), "Mapping Minor Characters: Quantifying and Visualizing Character Space in Dickens’s Novels and in their Adaptations"
- Leah Henrickson (Loughborough U), "Who is the author of the computer-generated text?"
Friday, 7 June 2019 [DHSI; ADHO Pedagogy SIG Conference Opening]

9:00 to Noon  Classes in Session

12:15 to 1:15  Lunch Reception / Course E-Exhibits (MacLaurin A100)

1:30 to 1:50  Remarks, A Week in Review (MacLaurin A144)

2:00 to 3:00  Joint Institute Lecture (DHSI and ADHO Pedagogy SIG Conference):
Matt Gold (CUNY Graduate Center and Association for Computers and the Humanities): “Thinking Through DH: Proposals for Digital Humanities Pedagogy”
Chair: Diane Jakacki (Bucknell U)
(MacLaurin A144)

Abstract: How do we teach digital humanities, and how should DH be taught? What, indeed, should we teach when we teach DH? This talk will present a proposal for grounding digital humanities pedagogical practice in the research interests of our students and the epistemological foundations of our methods rather than through an approach grounded more central in data and methods.

3:30 to 5:00  Joint Reception: DHSI and ADHO Pedagogy SIG Conference (University Club)
E-Poetry Event (Chris Tanasescu)
Watch this space for details, including how to participate!

Saturday, 8 June 2019 [Conference, Colloquium, and Workshop Sessions]

8:00 to 9:00  Conference / Workshop Registration (MacLaurin A100)

Coffee, Tea, &c?  Looking for some morning coffee or tea, or a small nibble? Options and hours of operation for weekend campus catering are available here. Mystic Market usually opens around 10.00.

9:00 to 4:00  DHSI Conference and Colloquium Sessions
ADHO Pedagogy SIG Conference Sessions
Right2Left Workshop Sessions

9:00 to 4:00  All Day DHSI Workshop Session (click for workshop details and free registration for DHSI participants)

9:00 to 9:10  Informal Greetings, Room Set-up (Lobby, outside Hickman 105)

Session 1
DHSI Colloquium and Conference (Hickman 105)
Digital Humanities & Literature, Chair: Kim O’Donnell (Simon Fraser U)
- Youngmin Kim (Dongguk U), “Transdiscursivity in the Convergence of Digital Humanities and World Literature”
- Caroline Winter (U Victoria), “Digitizing Adam Smith’s Literary Library”
- Kaitlyn Fralick (U Victoria); Kailey Fukushima (U Victoria); Sarah Karlson (U Victoria), “Victorian Poetry
9:10 to 10:30
ADHO Pedagogy SIG Conference (Hickman 110)
Chair: Katherine Faull (Bucknell U)
- Aaron Tucker and Nada Savicewic (Ryerson U), "Write Here, Right Now: An Open Source eTextbook for the Flipped Classroom"
- Heather McAlpine (U Fraser Valley), "Digital Meters: Using Text Encoding to Teach Literature in the Undergraduate Classroom"
- Tiina H. Airaksinen (U Helsinki), "Digital Humanities in Cultural Studies: Creating a MOOC course for University Students and A-Level Students"

Right2Left Workshop (Hickman 116)

10:30 to 10:40
Break

10:40 to Noon
Session 2
DHSI Colloquium and Conference (Hickman 105)
Digital Humanities & Society, Chair: Eleanor Reed (Hastings C)
- Joel Zapata (Southern Methodist U), "Uncovering the Southern Plains’ Mexican American Civil Rights Movement"
- Ayo Opaisanwo (U Ibadan), "Online Newspaper Construction of Agitation for the Sovereign State of Biafra in Nigeria"
- Joseph Jones (U British Columbia), "Testbed for an Approach to Distant Reading: Fictions That Represent Vietnam War Resisters in Canada"
- Brendan Mackie (U California, Berkeley), "Visualizing Long-Term Cultural Change: An Example From The Birth of Civil Society"

10:40 to Noon
ADHO Pedagogy SIG Conference (Hickman 110)
Chair: Laura Estill (St Francis Xavier U)
- Jane Jackson (Chinese U of Hong Kong), "Interrogating digital spaces for intercultural meaning-making"
- Ryan Ikeda (UC Berkeley), "Disrupting Digital Literacy: Situating Electronic Literature Among Public Education Initiatives"
- Christopher Church, Katherine Hepworth (U Nevada, Reno), "We’re STEAMed! A call for balancing technical instruction and disciplinary content in the digital humanities"
- Chelsea Milbourne (Cal Poly, San Luis Obispo), "Finding the Right Fit between Technology and Class Content: Reflections on Including Web Development in a Digital Storytelling Course"

Right2Left Workshop (Hickman 116)
- Edward “Eddie” Surman (Claremont Graduate U), "Qualitative Digital Text Analysis and #Right2Left Languages: A Demonstration of Atlas.ti using the Hebrew Bible"

Noon to 1:10
Lunch (We recommend Mystic Market on weekends!)

1:10 to 2:30
Session 3
DHSI Colloquium and Conference (Hickman 105)
Digital Humanities & Community, Chair: Claire Carlin (U Victoria)
- Pia Russel (U Victoria); Emily Stremel (U Victoria), "Mentorship and disability: Supporting disabled employees in digital humanities"
- Amy Lueck (Santa Clara U), "Virtually Emplacing Indigenous Memory"
- Md. Shehabul Alam (National U Bangladesh), "Integrating Library Service with Union Information and Service Center: A Joint Initiative towards Digital Bangladesh"
- Veronica Gomez (Instituto de Humanidades y Ciencias Sociales (HuCSo) - UNL-CONICET), "Latin American E-literature and Location: The Nation Revisited in Electronic Literature Organization (ELO)"
- Francesca Giannetti (Rutgers U, New Brunswick), "Fr Ready Library Pedagogy and Digital Humanities Methodology"
- Najla Jarkas (American U Beirut) and David Joseph Wrisley (NYU Abu Dhabi), "RTL Software Localization and Digital Humanities: the Case Study of Translating Voyant Tools into Arabic"
2:30 to 2:40 Break

Session 4

DHSI Colloquium and Conference (Hickman 105)
Digital Humanities & Media, Chair: Caroline Winter (U Victoria)
- Ashleigh Casserme-Stanfield (U Chicago), “Sonifying Hamlet and Reading the Room”

ADHO Pedagogy SIG Conference (Hickman 110)
Chair: Aaron Tucker (Ryerson U)
- Youngmin Kim (Dongguk U), “Teaching Digital Humanities and World Literature in Class”
  Alice Fleerackers, Juan Pablo Alperin, Esteban Morales, Remi Kalir (Simon Fraser U, U Colorado Denver), “Online annotations in the classroom: How, why, and what do students learn from annotating course material?”
  Andie Silva (York C and Graduate Center, CUNY), “Keeping it Local: Undergraduate DH as Feminist Practice”

Right2Left Workshop (Hickman 116)
- Joanna Byszuk (Institute of Polish Language, Polish Academy of Sciences, Warsaw/Computational Stylistics Group) and Alexey Khismatulin (Institute of Oriental Manuscripts, Russian Academy of Sciences, Saint Petersburg), “Attribution of Authorship for Medieval Persian Quasidas with Stylometry”

2:40 to 4:00

Sunday, 9 June 2019 [Workshop Sessions]

8:00 to 5:00
The day’s events are included with your DHSI registration. If you’re not registered in DHSI, you’re very welcome to join us by registering here as a Conference / Colloquium / Workshop participant. We’ll have a nametag waiting for you!

Coffee, Tea, &c?
Looking for some morning coffee or tea, or a small nibble? Options and hours of operation for weekend campus catering are available here. Mystic Market usually opens around 10.00.

9:00 to 4:00
All Day Workshop Sessions (click for workshop details and free registration for DHSI participants)
- 55. Introduction to Machine Learning in the Digital Humanities [8-9 June; All day, each day] (David Strong Building C124, Classroom)
- 56. Pedagogy of the Digitally Oppressed: Anti-Colonial DH Methods and Praxis [9 June; All Day] (Hickman 116, Classroom)
- 57. Natural Language Processing and Network Coding Apps for Text & Textual Corpus Analysis in the Humanities [9 June; All Day] (David Strong Building C114, Classroom)

AM Workshop Sessions (click for workshop details and free registration for DHSI participants)
- 59. 3D Visualization for the Humanities [9 June; AM] (Cornett A229, Classroom)
- 60. It’s All Relational: AbTeC’s Indigenous Video Game Workshops as Storytelling Praxis [9 June; AM] (Cornett A121, Classroom)
- 61. Spatial DH: De-Colonizing Cultural Territories Online [9 June; AM] (Clearihue D130, Classroom)
- 63. Creating a CV for Digital Humanities Makers [9 June; AM] (David Strong Building C108, Classroom)

Noon to 1:00
Lunch (We recommend Mystic Market on weekends!)

PM Workshop Sessions (click for workshop details and free registration for DHSI participants)
- 65. Indigenous Futurities in the Classroom and Beyond [9 June; PM] (Cornett A121, Classroom)
- 66. DHSI Knits: History of Textiles and Technology [9 June; PM] (Fine Arts 109, Classroom)
- 68. Linked Open Datafication for Humanities Scholars [9 June; PM] (McPherson Library A003, Classroom)
- 69. Stylo - WYSIWYM Text Editor for Humanities Scholars [9 June; PM] (McPherson Library A025, Classroom)

After the day, many will wander to Cadboro Bay and the pub at Smuggler’s Cove OR the other direction to Shelbourne Plaza and Maude Hunter’s Pub OR even into the city for a bite to eat.

Monday, 10 June 2019
Your hosts for the week are Ray Siemens and Jannaya Friggstad Jensen.

7:45 to 8:15

DHSI Last-minute Registration (MacLaurin A100)

8:30 to 10:00

Welcome, Orientation, and Instructor Overview (MacLaurin A144)

Classes in Session (click for details and locations)

- 29. [Foundations] Understanding The Predigital Book: Technologies of Inscription (McPherson Library A003, Classroom)
- 30. [Foundations] Databases for Digital Humanists (McPherson Library 210, Classroom)
- 33. Digital Storytelling (Cornett A120, Classroom)
- 34. Text Mapping as Modelling (Clearihue D131, Classroom)
- 35. Stylometry with R: Computer-Assisted Analysis of Literary Texts (Clearihue A102, Lab)
- 36. Open Access and Open Social Scholarship (Clearihue D130, Classroom)
- 37. Digital Games as Tools for Scholarly Research, Communication and Pedagogy (Cornett A229, Classroom)
- 38. Queer Digital Humanities (David Strong Building C114, Classroom)
- 40. Introduction to Electronic Literature in DH: Research and Practice (Cornett A128, Classroom)
- 41. Surveillance and the Critical Digital Humanities (David Strong Building C108, Classroom)
- 42. Text Analysis with Python and the Natural Language ToolKit (Clearihue A103, Lab)
- 43. Creating LAMP Infrastructure for Digital Humanities Projects (Human and Social Development A170, Lab)
- 44. Processing Humanities Multimedia (Human and Social Development A150, Lab)
- 46. Digital Humanities Pedagogy: Integration in the Curriculum (Cornett A121, Classroom)
- 47. Accessibility & Digital Environments (Priestly Law Library 265, Classroom)
- 48. Agile Project Management (Cornett A132, Classroom/Lab)
- 49. XPath for Processing XML and Managing Projects (Clearihue A105, Lab)
- 50. Endings: How to End (and Archive) your Digital Project (Priestly Law Library 192, Classroom)
- 51. Introduction to Humanities Data Analysis & Visualization in R (Human and Social Development A160, Lab)
- 53. Introduction to Network Analysis in the Digital Humanities (Clearihue D132, Classroom)

10:15 to Noon

Lunch break / Unconference Coordination Session (MacLaurin A144)
(Grab a sandwich and come on down!)

"Mystery" Lunches

12:15 to 1:15

Classes in Session

1:30 to 4:00

Institute Lecture: Angel David Nieves (San Diego State U): "3D Mapping and Forensic Traces of Testimony: Documenting Apartheid-Era Crimes Through the Digital Humanities"
Chair: Constante Crompton (U Ottawa)
(MacLaurin A144)

Abstract: In 1989 the killing of a queer, 14-year-old youth in Winnie Mandela's house named Stompie Seipei (an event that few in South Africa are willing to recall, let alone discuss, in any detail) -- is perhaps one of the most glaring examples where the queer and activist community was suppressed or erased from anti-apartheid/liberation histories. Digital humanities may actually help both reconstruct and recover a history that is still very early in the telling, despite what is commonly believed about the liberation struggle and the contributions of queer activists in the dismantling of apartheid. Perhaps it could explain why a youth such as Seipei was killed -- or at the very least, provide a more complex and messy narrative that permits one to know more how the history of queer anti-apartheid activists was suppressed. This talk outlines a methodology for "messy thinking and writing" in the digital humanities that -- through a queer and feminist intersectional framework -- permits a more complex layering of oral histories and 3D historical reconstructions.

4:10 to 5:00

Lunch break / Unconference Coordination Session (MacLaurin A144)
(Grab a sandwich and come on down!)

"Mystery" Lunches

5:00 to 6:00

Reception (University Club)

Tuesday, 11 June 2019

9:00 to Noon

Classes in Session

12:15 to 1:15

Lunch break / Unconference

"Mystery" Lunches
1:30 to 4:00

Classes in Session

4:15 to 5:15

DHSI Conference and Colloquium Lightning Talk Session 4
Chair: Lindsey Seatter (U Victoria)

- Ashley Caranto Morford (U Toronto); Kush Patel (U Michigan); ArunJacob (McMaster U), "#OurDHIs anti-colonial: Questions and challenges in dismantling colonial influences in digital humanities pedagogy"
- Julia King (U Bergen), "Developing Network Visualizations of Syon Abbey's Books, 1415-1539"
- Luis Meneses (ETCL; U Victoria), "Identifying Changes in the Political Environment in Ecuador"
- Alicia Brown (Texas Christian U), "Digital Cartography of the Ancient World"
- Laura Horak (Carleton U), "Building the Transgender Media Portal"
- Andrew Boyles Peterson (Michigan State U), "Last Mile Tracking: Implications of Rental Scooter Surveillance"

6:00 to 8:00

DHSI Newcomer's Gathering
Grad House Restaurant, Graduate Student Centre
Come down, buy meal and a beverage, and make some new friends!

9:00 to Noon

Classes in Session

12:15 to 1:15

Lunch break / Unconference

"Mystery" Lunches

Presentation: An Introduction Jupyter Notebooks for Researchers
Chair: Lindsey Seatter (U Victoria)

This presentation introduces Jupyter Notebooks for researchers, via a partnership between Compute Canada and the Pacific Institute for the Mathematical Sciences (PIMS) including a large number of Canadian institutions. Read more here. Presenting is James Colliander, PIMS Director and team.

1:30 to 4:00

Classes in Session

4:15 to 5:15

DHSI Conference and Colloquium Lightning Talk Session 5
Chair: Lindsey Seatter (U Victoria)

- Calin Murgu (New College of Florida), "Putting local metadata to strategic use: A Dashboard for visualizing 60 years of theses metadata"
- Jason Lajoie (U Waterloo), "Queer Critical Making and the Logic of Control"
- John Barber (Washington State U), "Zambezi River Bridge"
- Kent Emerson (U Wisconsin-Madison), "Digital Mappa and the George Moses Horton Project"

6:00 to 7:00

"Half Way There (yet again)!" [An Informal, Self-Organized Birds of a Feather Get-Together]
Felicitas, Student Union Building
Bring your DHSI nametag and enjoy your first tipple on us! [A great opportunity for an interest group meet-up ....]

9:00 to Noon

Classes in Session

12:15 to 1:15

Lunch break / Unconference

"Mystery" Lunches

[Instructor lunch meeting]

1:30 to 4:00

Classes in Session

4:10 to 5:00

Institute Lecture: Karina van Dalen-Oskam (Huygens Institute and U Amsterdam; Alliance of Digital Humanities Organizations): "The Riddle of Literary Quality: Some Answers"
Chair: Aaron Mauro (Penn State, Behrend C)

Abstract: What is literature, and can you measure it? That is the key question of the project The Riddle of Literary Quality. "The Riddle" is a research project of the Huygens Institute for the History of the Netherlands (Amsterdam) in collaboration with the Fryeke Akademy (Leeuwarden) and the Institute for Logic, Language and Computation (University of Amsterdam). The Riddle combines computational analysis of writing style with the results of a large online survey of readers, completed by almost 14,000 participants. In my talk, I will go into
some of the main results of the project.

**Friday, 14 June 2019**

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00 to Noon</td>
<td>Classes in Session</td>
</tr>
<tr>
<td>12:15 to 1:15</td>
<td>Lunch Reception / Course E-Exhibits (MacLaurin A100)</td>
</tr>
<tr>
<td>1:30 to 2:00</td>
<td>Closing, DHSI in Review (MacLaurin A144)</td>
</tr>
</tbody>
</table>

**Contact info:**

[enstitut@uvic.ca](mailto:enstitut@uvic.ca)  P: 250-472-5401  F: 250-472-5681
Coursepak

Modelling. Virtual. Realities.
A Practical Introduction to Virtual (and Augmented) Reality

Zoe Schubert and Jan G. Wieners
Institute of Digital Humanities | University of Cologne | Germany

Course description on DHSI website:
http://www.dhsi.org/courses.php#ModellingVR

Further material on GitHub:
https://github.com/DH-Cologne/DHSI2019_VR

Table of Contents
1) General Information and Introduction
2) Schedule
3) Readings
1) General Information and Introduction

Welcome

You are all very welcome! This coursepak includes general information, scheduling details, and relevant readings for our course *Modelling. Virtual. Realities. A Practical Introduction to Virtual (and Augmented) Reality*. For participants it would be useful to take a closer look at the provided material in advance, as it will be essential to the theoretical and practical work.

First we would like to give some helpful general information:

- [ToDo: add information about hardware/ software requirements => Should the participants bring hardware/ preinstall software?]
- We will give a brief introduction and overview about state of the art WebXR frameworks and the technology behind it. For the hands-on sessions we will focus on BabylonJS - a 3D engine based on WebGL/Web Audio and JavaScript (https://www.babylonjs.com/). For software development purposes we will use an integrated development environment like IntelliJ “Webstorm” or free open source code editors like “Visual Studio Code”, “Atom” or “Sublime Text”. GitHub (https://github.com) and OpenProject (https://www.openproject.org) will support the agile software development.
- The course will be divided into theoretical and reflective sessions and practical work. Both parts are closely linked and influence each other. We would like to give as many input and feedback as needed to provide all participants to work on their own VR project influenced by their disciplines and scientific backgrounds.
- Even if the used frameworks, tools and software are offering a quick start for beginners, there are time consuming and frustrating moments during programming. We will help you to plan your own project which can be built during one week, but we might not be able to protect you from frustrating situations which are usually unavoidable.
- The preliminary timetable takes up basic topics and is based on the experience of previous practical courses. The topics and the duration of the theoretical and practical parts can be adapted at any time if it seems appropriate for the group. Working independently and gaining personal experience is one of our most important concerns.

Any questions in advance of the gathering can be directed to us (Jan and / or Zoe) at jan.wiener@uni-koeln.de and zoe.schubert@uni-koeln.de

Summary: Modelling. Virtual. Realities. A Practical Introduction to Virtual (and Augmented) Reality

*As it says on the DHSI website:*
Current technical developments offer many people the opportunity to have first Virtual Reality experiences. A wide selection of devices, specifically made for this purpose, have become more affordable. Also mobile devices of the last generations make VR available when used in conjunction with Google Cardboards and recent developments like Google Daydream View or Samsung Gear VR. State of the art JavaScript libraries like Three.js and frameworks like BabylonJS, Mozilla’s A-Frame, or Facebook’s React VR allow highly accessible web- and browser-based VR implementations for mobile phones, brought to life through affordable physical environments like Google Cardboard.

This course provides a practical introduction into the implementation of virtual reality environments by using classic web standards and state of the art WebVR frameworks. After a practical introduction to the basic functionalities, the possible contents for presentation in VR will be examined theoretically. Not only virtual worlds but also existing 3D models can be explored and experienced anew by and in VR. In addition, the transfer of other media content into this environment is suitable, which currently implies few specifications for presentation. The partly (still) experimental use allows finding new ways of interaction and presentation. In this course, we want to try out new ways of representation and give the participants the opportunity to use WebVR for their work.

Having completed this course, participants will have a better understanding of the possibilities (and restrictions) web-based virtual reality provides and be able to implement their own web-based VR environments by using current standards in web development (HTML, CSS, JavaScript) and state of the art WebVR frameworks. This offering is co-sponsored by the Department for Digital Humanities, University of Cologne.

This is a hands-on course. Consider this offering to build on: Fundamentals of Coding / Programming for Human(s) lists; Web Development / Project Prototyping for Beginners with Ruby on Rails. Consider this offering in complement with and / or to be built on by: 3D Modelling for the Digital Humanities and Social Sciences, Text Mapping as Modelling, Physical Computing and Desktop Fabrication; Understanding Topic Modelling; Data, Math, Visualization, and Interpretation of Networks; and more.

Instructors

**Zoe Schubert** (U Cologne, zoe.schubert@uni-koeln.de) is working as a research associate and lecturer for Digital Humanities at the Institute for Digital Humanities / Computer Science for the Humanities (Historisch-Kulturwissenschaftliche Informationsverarbeitung) at the University of Cologne, Germany. Zoe Schubert holds a Master’s degree in Computer Science for the Humanities and Media Science. She is writing her dissertation about „Virtual Reality as a transformative technology in the Humanities - Theater in VR“. Supervisors are Prof. Dr. Øyvind Eide and Prof. Dr. Manfred Thaller from the University of Cologne, Germany. Her research interests include Virtual Reality and Augmented Reality, Media Transformations, visualisations and web technologies.

**Jan G. Wieners** (U Cologne, jan.wieners@uni-koeln.de) is a research associate, lecturer for Digital Humanities and software developer at the Institute for Digital Humanities at the University of Cologne, Germany. Jan G. Wieners holds a Master’s degree (Magister Artium) in Computer Science for the Humanities, German Philology and Philosophy and a PhD in Digital Humanities at the University of Cologne. His research interests include Virtual Reality and Augmented Reality, Media Transformations, Artificial Intelligence, Computer Vision, Cognitive Mapping and Modelling.
Further Information

Additional material for this course will be on GitHub. You can access it quickly through the following domain: https://github.com/DH-Cologne/DHSi2019_VR.

There might be new information added over the coming weeks / during the sessions.
2) Preliminary Schedule

Day 1: Getting started

Topics:
Realities | Objects and Stories | Technologies, Tools and Frameworks | First steps | Projects

Part 1 (10:15 - 12:00)
During this first part we would like to start with an introduction into our topic. Therefore we talk about a theoretical framework and the available technologies in the context of mixed -, augmented - and virtual reality. We will have a closer look on Objects and narrations in VR as well as give an technological overview on WebVR / WebXR and projects using this technology (also in the humanities).

Part 2 (13:30 - 16:00)
For the practical start we will give an overview about tools and (web) frameworks to build VR-applications. In the following we guide through the first steps in developing a VR application: Setting up a project, drawing basic 3D-objects (primitives). We will also discuss individual project ideas for the course.

Day 2: Modelling realities

Topics:
Modelling | Transformation | Agile Software Development | Practical work | Interactivity in VR

Part 1 (9:00 - 12:00)
A requirement analysis for the individual projects is part of this hands-on: agile software development part. We will define necessary resources, functionalities of the VR world and visual objects. Wireframes and User Stories are helpful and part of the project management. We will introduce OpenProject as Software for supporting these things and introducing a ticket system.

Part 2 (13:30 - 16:00)
The afternoon is divided into two parts and deals theoretically with interactivity in VR and practically with the implementation of the VR projects. Modelling and transformation will be of particular importance.

Day 3: Time, navigation and interactivity in VR

Topics:
Modelling | Hands-On | Practical work | Interaction & Movement in VR

Part 1 (9:00 - 12:00)
All objects and models should be integrated into the virtual worlds. Focussing on detailed modelling is also possible for selected objects. Interacting with the world and its objects is the second huge step in the context of the practical projects.

Part 2 (13:30 - 16:00)
Theoretical and practical work on time, navigation and interactivity in VR.
Day 4: Final sprint
Topics:
Programming | Critical time review

Part 1 (9:00 - 12:00)
Independent practical work on projects with individual support.

Part 2 (13:30 - 16:00)
Putting all together: final review.

Day 5: Summing up
Topics:
Review | Analyses | Presentation | Outlook

(9:00 - 12:00)
The last session will be used to review analyse and present the practical projects and also the ideas of the course. We will have a critical discussion about use cases and benefits of using WebXR in the context of digital humanities.
3) Readings

The following additional readings intend to give a broader perspective on virtual and augmented reality, software developing and modelling.

Table of Contents

I. Arianna Ciula, Øyvind Eide; Modelling in digital humanities: Signs in context, Digital Scholarship in the Humanities, Volume 32, Issue suppl_1, 1 April 2017, Pages i33–i46. https://doi.org/10.1093/llc/fqw045


Modelling in digital humanities: Signs in context

Arianna Ciula
Department of Humanities, University of Roehampton, London, UK

Øyvind Eide
Historisch-Kulturwissenschaftliche Informationsverarbeitung, Universität zu Köln, Germany

Abstract

In this paper we focus on modelling as a creative process to gain new knowledge about material and immaterial objects by generating and manipulating external representations of them. We aim at enriching the current theoretical understanding by contextualising digital humanities practices within a semiotic conceptualisation of modelling. A semiotic approach enables us to contextualise modelling in a scholarly framework well suited to humanistic enquiries, forcing us to investigate how models function as signs within specific contexts of production and use. Kralemann and Lattmann’s semiotic model of modelling complemented by Elleström’s theories on iconicity are some of the tools we use to inform this semiotic perspective on modelling. We contextualise Kralemann and Lattmann’s theory within modelling practices in digital humanities by using three examples of models representing components and structure of historical artefacts. We show how their model of models can be used to understand and contextualise the models we study and how their classification of model types clarify important aspects of digital humanities modelling practice.

1 Introduction

In this article we focus on modelling as a creative process to gain new knowledge (meaning) about material and immaterial objects by generating and manipulating external representations of them. Modelling is widely understood and used as a heuristic strategy in the sciences (Mahr, 2009; Frigg and Hartmann, 2012) as well as in digital humanities (hereafter DH) research where it is considered a core practice (McCarty, 2005, pp. 20–72). In the past two decades there has been a significant development of theory that complements the practice-based tradition of the field (e.g. ibid, Buzzetti, 2002; Beynon et al., 2006; Jannidis and Flanders, 2012; Flanders and Jannidis, 2015).

We aim at enriching the current theoretical understanding by contextualizing DH practices within a semiotic conceptualization of modelling. A semiotic approach enables us to contextualize DH modelling in a scholarly framework well suited to humanistic enquiries, forcing us to investigate how models function as signs within specific contexts of production and use. Kralemann and Lattmann’s (2013) semiotic model of modelling complemented by Elleström’s (2013) theories on iconicity are some of the tools we use to inform this semiotic perspective on modelling.
We then go on to contextualize Kralemann and Lattmann’s theory within modelling practices in DH by using three examples of DH models representing components and structure of historical artefacts. We show how their model of models can be used to understand and contextualize the models we study and how their classification of model types clarifies important aspects of DH modelling practice.

2 What is Modelling?

In this article we take a pragmatic definition of modelling as a starting point. Indeed, interdisciplinary theories around modelling are used mainly to inform our analysis of modelling practices. By modelling we intend a creative process of thinking and reasoning where meaning is made and negotiated through the creation and manipulation of external representations. We narrow this definition further by applying it to modelling as a research strategy: modelling is a process by which researchers make and manipulate external representations—what Godfrey-Smith (2009) calls ‘imaginary concreta’—to make sense of the conceptual objects and phenomena they study.

Modelling in DH is often understood as ‘any act of formal structuring’ of data intended as ‘formal information’ (Flanders and Jannidis, 2015, p. 4). Our point of departure (see also Ciula and Eide, 2014; Ciula and Marras, 2016) is however wider to allow us to explore whether a more encompassing definition can overcome some limitations of a narrower take on modelling. Rather than prioritizing a conceptualization of modelling directed first and foremost at communicating with the computer, we rather attempt at seeing modelling as a means to create ‘tools for thinking’ (Bradley, 2015).

3 Semiotics and DH Modelling

Rather than framing our reflection on modelling around human–machine communication or on implementative purposes in a strict sense, we propose to consider modelling as a process of signification and reasoning in action. Contextualizing modelling within a semiotic framework means indeed to consider it as a strategy to make sense (signification) via practical thinking (creating and manipulating models). We use an interdisciplinary perspective on modelling to guide us both in understanding how models as signs are made (the construction of models) as well as in understanding how something new is discovered in the process of making and using models (the epistemic and heuristic value of models).

3.1 Dynamic relation models/objects/interpretations

Kralemann and Lattmann (2013) claim that models should be understood as signs in the Peircean sense. In Peirce’s seminal theory of signs, the sign is a triadic relation between a ‘representamen’ (the sign from which the relation begins, sometimes also called in the literature the sign-vehicle), its object, and the interpreting thought. Often represented as a tripod where the three ‘composing elements’ (Olteanu, 2015, p. 127)—object, ‘representamen’, and ‘interpretant’—intersect, the sign for Peirce is hence, first and foremost, relational. The experience of interpreting signs or signification (semiosis) is therefore intrinsically dynamic. As a consequence, a semiotic approach which considers models as signs gives high prominence to a dynamic view on models reinstating in renewed terms the value of modelling as an open process—in particular a process of signification.

Figure 1 shows in abstract and simplified terms (cf. Kralemann and Lattmann, 2013, Fig. 5) how the composing elements of the sign-relation (here translated into a model-relation) interact.

3.2 Models as icons

The semiotic theory of signs proposed by Peirce identifies three types of signs based on the relation between the object and the sign: symbols (e.g. conventional names used to denote objects), icons (e.g. onomatopoetic words such as ‘splash’), and indexes (signs used to point directly to their meaning, such as ‘there’). In this respect, Kralemann and Lattmann (2013, pp. 3399–400) claim that models are icons, because the dominant relation with the objects they represent is one of similarity, as shown in Fig. 2. In Peircean theory, such iconic relation of similarity is
what makes icons signify; icons act as signs based on how the relation of similarity is enacted: via simple qualities of their own in case of images, via analogous relations between parts and whole and among parts in the case of diagrams, and via parallelism of qualities with something else in the case of metaphors (Olteanu, 2015, pp. 77, 193).

Different shades of iconic similarity between sign and object as theorized by Peirce correspond to three kinds of models in Kralemann and Lattmann:

- **image-like** models, for example, real-life sketches where single qualities such as forms and shapes enable them to act as signs of the original objects they represent in given circumstances;
- **relational or structural** models, for example, diagrams such as the relation exhibited in the graph of a mathematical equation, where the ‘interdependence between the structure of the sign and the structure of the object’ (Kralemann and Lattmann, 2013, p. 3408) enables the modeller to make inferences about the original by manipulating its model;
- **metaphor-like** models which represent attributes of the original by a non-standard kind of parallelism with something else which generates further models (metaphors are metamodels; Kralemann and Lattmann, 2013, p. 3409).

In Kralemann and Lattmann’s theory as well as in Peirce’s original theory, models do not act as signs in virtue of themselves. What establishes the model as a sign is the interpretative act of a subject, whether as creator or reader. The practical act of modelling connects the model to its interpretation, that is, to its specific semantic content in a given social and institutional context (Kralemann and Lattmann, 2013, pp. 3402–3). The modeller’s judgement depends on his or her presuppositions connected to ‘theory, language or cultural practice’ (Kralemann and Lattmann, 2013, p. 3417). Models are contingent. Kralemann and Lattmann also reiterate the concept of models as middle ground between theory and objects. The relationship of iconicity between the model and the object being modelled is partly externally determined (it relies on the similarity between the model and the object) and partly internally determined (it depends on theory, languages, conventions, scholarly tradition, etc.). Based on this duality, they stress, on the one hand, the subjectively determined dependency of models on prior knowledge and theory and, on the other, their independence from these in light of the specific conditions of the objects being modelled.

### 3.3 Similarity, iconicity, and reasoning

One consequence of seeing models as icons is that through an understanding of the process by which icons are made and used we can gain new insights on how models are built and used. This understanding highlights similarity as a key to link models to the modelled:

Representation based on resemblance generally falls under the heading of ‘iconicity’. When something is understood to be a sign of something else because of shared, similar qualities, it is referred to as an iconic sign (Elleström, 2013, p. 95).

The notion of iconicity is however not only about how models (as signs) appear with respect to similarity to their objects. It also encompasses the possibility of manipulating models and reasoning with them. This is another point of connection...
between models and icons, a point that goes to the core of DH practice.

Modelling in DH has a hybrid nature which combines implementation-oriented work with methodological inquiries bearing implications beyond the specific implementation. This distinction has recently been verbalized as one between altruistic and egoistic modellers in Jannidis and Flanders (2013, p. 138) and as one between modelling for production and modelling for understanding in Eide (2015a). An altruistic modeller will create a model for others’ use, often as part of a production project, whereas an egoistic modeller will create a model to be used at the individual level or by a group to inquire into a specific area of interest. In the latter case, using models to reason with is considered to be a main goal of modelling, whereas in the former it rather forms part of a process with a mainly practical goal, for example, the publication of a collection of documents.

This distinction can be useful in analytical terms, but is problematic in that it ignores that all models are used as external representations to facilitate reasoning. Any model used in DH will to some extent be used for reasoning, and especially shared reasoning or negotiation of meaning. A model gives us a common language to talk about the world. To take one example: The Text Encoding Initiative (TEI) does not only give us a method for marking up texts, but also a language and formalism in which to think about textual phenomena such as manuscripts or poems. As stressed in Stachowiak (1973, p. 60), stringent and exact systems for making deductions are useful also when no generally agreed upon objective reality exists; they can even be more necessary when reality is elusive and negotiable.

The use of models as external representations to reason with has important points of connection with Peirce’s thinking about icons and reasoning:

Similarity, which is the root of iconicity, is not simply an absolute trait that is ready to be picked up in the external world; instead it is a perceived quality processed by subjective attention and selection, and a potent force in cognition. (Ellestrohm, 2013, p. 97)

According to Peirce, ‘it is by icons only that we really reason’ (Peirce, 1933, CP 4.127 [1893]). In more recent literature, cognitive sciences and the philosophy of scientific modelling have been brought together (Nersessian, 2008). In particular, within theorizations of distributed cognition (Hoffmann, 2011, p. 199), thinking processes are seen as being distributed in the world and shared among different people through external mediations. Historical accounts of scientific practices

---

Fig. 2 The Peircean trichotomy of signs into icons, indexes, and symbols based on the relation with its object (of similarity in the case of the icons) and the subsequent classification of icons (or rather pure icons or hypoicons) into images, diagrams, and metaphors based on how the respective similarity relations signify. Highlighted in grey are the sign types associated with models by Kralemann and Lattmann (2013, Fig. 2)
establish model-based reasoning as a social problem-solving strategy comparable to practices in everyday life (Nersessian, 2008). When we share our scholarly ideas using models in reasoning and discussion, this is a type of process which is fundamentally icon based in Peirce’s sense. The role of graphical representations in ‘external cognition’ is described by Hoffmann (2011, p. 192) as ‘diagrammatic reasoning to solve problems, to cope with complexity, to learn something new, or to resolve conflicts’. Seen as icons such diagrams fall into a wide variety of model types,10 from toy cars used as scale models to mathematical formulae and semantic networks. Why do we make such external representations? To take one example: Wood (1993, p. 51) distinguishes between the process of mapping and the one of mapmaking, which consists of the difference between a gesture leaving no physical trace and making a permanent inscription. The choice is based on the needs in concrete communication situations: if the communication need is complex, a map is better than just an allusive gesture. This distinction is not sharp and it is connected to the ‘continuum’ between communication and reasoning, as pointed out by Hoffman:

When I draw a map to explain a friend how to drive to a certain location, I would communicate by means of a diagram but I would not reason with it. Diagrammatic reasoning is about problem solving, decision making, knowledge development, and belief change by means of diagrams. However, I do not presuppose a clear cut distinction between diagrammatic communication and diagrammatic reasoning. There might be a continuity between both these possibilities. (Hoffmann, 2011, p. 193–4)

Especially in project-based DH practice, where interdisciplinary groups work together to solve problems at practical as well as theoretical levels, reasoning and communication act as two sides of the same coin.

3.4 Grades of iconicity

As shown above, Peirce distinguished between three types of hypoicons: images, diagrams, and metaphors. Let us take the example of an apple. The image of an apple put up in the window of a grocery shop has a signification immediately perceived by a hungry tourist passing by. She will assume that the sign on paper, through its image-like resemblance with a real apple, indicates that apples are sold in the shop. While this immediacy is not there to be seen for everyone and in every circumstance (it would not work for a person who does not know what an apple is or what it looks like, and it would not necessarily be experienced by somebody not interested in buying apples there and then), it is still general enough to be defined as an immediate image for apples within a given context.

A botanical visualization of the reproduction system of the apple plant can be used to exemplify a diagrammatic icon of apples. The diagram exhibits the structural similarity between the form of the organs as represented in the diagram and the organs we find in actual apples.

Finally, a metaphorical icon can be exemplified by a representation of an apple as a sign of sin. This can be expressed in various forms, such as an apple in a biblical painting or expressions such as ‘she gave me the apple’. The whole expression—reduced to ‘sin is an apple’—is the metaphor implying a relation between the apple and sin (the object of the model). This sign relation makes it possible for the object of the sign ‘apple’ to become an icon for the object of the sign ‘sin’ (cf. the example provided by Kralemann and Lattmann, 2013, pp. 3408–9), establishing a chain of signs. Hence the words of the poet Pablo Neruda ‘innocence is round like an unbitten apple’ (Ode to the Apple).

The representamen of an image is perceptually close to its object, which means that the object may be sensuously perceived in much the same way as the representamen (this is a conception that is close to Peirce’s own few remarks on the image). The representamen of a metaphor is at a greater distance from its object, which means that the interpretation
of a metaphor includes one or several cognitive leaps that make the similarity between representamen and object apparent. (Ellestrom, 2013, p. 104)

What we see clearly in the semiotic understanding of modelling is how the analytical dichotomy objects vs. models is useful, but also misleading. For analytical purposes the object is the apple and the models (icons) are the three different examples. But the object changes when the model changes; the meaning of the apple in the metaphorical example above is different from the apple in the diagrammatic example. The context of the interpretation changes the sign but the sign also changes the context of interpretation.

3.5 Space and time

In his model of media modalities, Elleström distinguishes between four, namely, the material, sensorial, spatiotemporal, and semiotic modalities (Ellestrom, 2010). This is not a claim for any linear development through the modalities; it is rather an analytical distinction to clarify various aspects of a media expression. Different configurations of the four modalities can be used to specify the characteristics of specific media.

While the focus in this article is on the semiotic modalities of models as media expressions, our analysis, as we will see later with the examples, also considers the other three modalities. For our purpose it is especially important to understand how the spatiotemporal modality structures the experience of the material interface through which we encounter a media expression into conceptions of space and time. When we read a text and when we study a map we act in time. But the time operates differently. In most types of text the space of the printed or written page is turned into one or several sequences of characters and words, read in a predefined order. In studying a map we can let our eyes wander in any pattern while still getting to the meaning of the map.

4 Examples in DH

In this article we take previous research (Ciula and Eide, 2014; Ciula and Marras, 2016) one step further by mapping Kralemann and Lattmann’s trichotomy of models as icons to examples of digital modelling in DH research dealing with historical artefacts. These prototypical cases were chosen to investigate how model types relate to the cultural objects they represent and how modellers reason with them.

If we accept Kralemann and Lattmann’s argument, it follows that by modelling we link models to qualities and relationships already existing in the objects being modelled. Such linking is based on choices which are made for a certain end informing and motivating the act of modelling. Models are contingent, created in actual scholarly situations of production and use. A model is partially arbitrary in that the same inferences drawn by manipulating one model could have been reached in other ways, for instance using a different model.

In this framework, models operate as sign-functions initiating a sign-relation (model-relation). To understand their epistemic role, we need to look at
both how they come to be and how the similarity relation with the object is realized. By analysing the association of syntactic attributes of the source object with the attributes of the model we focus on the latter; that is, the representational correspondence. To explain the semantics of the model, the analysis of the similarity relation needs to be complemented with an analysis of the overall sign-relation in which production and use of models are enacted, as indicated in Fig. 1. Three examples will be used to analyse the three types of sign-functions and relations in a DH context.

In general one could say that every DH model is a diagram in that it is a formalism of logical and mostly mathematical nature; in this respect, Flanders and Jannidis talk about ‘data structure’ as different from ‘data modelling’ (Flanders and Jannidis, 2015, p. 8). However, we believe we can in fact identify different grades of iconicity corresponding to the three model types mentioned above, namely image, diagram, and metaphor. The classic example that comes to mind to represent an image-like model is a 3D graphic model such as, for instance, the model of an historical monument. The digital model acts as a surrogate of or a substitute for the reconstruction of the real object. A diagrammatic version of the same model could be the mathematical equations used to create the graphical 3D model. Below we dwell on three examples in detail.

4.1 Example 1: Image-like model
We will use an example from digital palaeography research (Ciula, 2005, 2009), where the abstract model letter acts as an image-like model of the samples it was algorithmically generated from. What we can learn about the objects of analysis (the medieval handwritten letterforms) depends on the features being selected in the modelling process. What is relevant for the scope of this article is that the inferential power of the model is mainly based on a strong immediate similarity (what above was called resemblance) between model and object. We can unpack this further by stating that the similarity is first and foremost of spatial nature: the handwritten letter is a two-dimensional spatial object as its spatial model is. However, their temporalities are different. We encounter single instances of letters in the manuscript pages, while the morphing models shown in Fig. 4 incorporate variants that can be visualized in sequence.

This specific palaeographical model is based on immediate similarity relevant for this context. The ‘a’ of the model looks very much alike the ‘a’ of the handwriting in the manuscript, they have the same spatiality. Its hermeneutical power relies, however, also on a different temporality between object and model. Anchoring the reasoning on the spatial similarity and going beyond it enable us to learn new things about the object. Indeed, new inferences are fostered by the availability of an ‘actual’ temporal element in the morphing of the model. While we have to look at all single instances in the manuscript, we get a model which incorporates all variants, and by sliding from left to right, we can ‘see’ those variants in real time. The object itself, however, is not temporal in this sense. So while the model is an abstraction—a fuzzy image which loses the precision of the instances out of which it was generated (representation is indeed asymmetrical) while keeping a basic (symmetrical) similarity to it—it gains an actual temporal mode that the single instance objects do not hold. If the modeller can make any inferences this is also due to her awareness of scribal variants and of what morphological traits are more revealing of different dating and location than others. So context and prior knowledge are important not only for the creation of models but also—not surprisingly—for their interpretation.13

4.2 Example 2: Relational model
As an example we will use models of landscapes described in historical sources, where textual information is modelled in the form of maps (Eide, 2015b). The inferential power of the model relies on the analogous relational structure between object and model. When the text says ‘A is north of B’ it makes a claim about a geometrical relationship between places denoted in the text. A map showing A north of B makes a claim expressing a similar geometrical structure. What new we can know about the object of analysis depends very much on the correspondence between the structuring of the textual expressions in the modelling process and the structure of the map model.
The model–object relationship here is not between an expression and a landscape but between two expressions in different media, as shown in Fig. 5. These media express structural relationships in fundamentally different ways. To see the structural similarity one needs to understand the written language being used in the text, the schemata used in topographical maps to convey meaning, and have experience of real landscapes. These elements define the context of the model.

In this example ‘similarity’ is not immediate resemblance. The digital model—the map—looks completely different from the source object—the text, but there is a structural similarity between the two. This structural similarity possesses a strong hermeneutical potential. It can be used to reveal gaps; there are things expressed in the text that cannot be put on the map. Examples of things that cannot be expressed include open, borderless expressions such as ‘the area north of the river’ and ambiguous expressions such as ‘Either A or B is on the border’. The analogy breaks at some point; the examples show how the signification of rich expressions in the text cannot be communicated via the structure of the map. Realizing this can lead to new knowledge, or rather to renegotiating what a text can mean, the meanings of a text. Based on the structural correspondence and non-correspondence between the virtual geographical space of the text and the geographical space of the map, the map makes the virtual space ‘visible’ and in so doing reveals a dissimilarity. It pinpoints the degree to which the text is underspecified spatially, how open the virtual space of the text is. This forces our understanding of the text to change.

4.3 Example 3: Metaphor-like model

Finally, we will use the example of network models used to capture information about references to persons in historical sources. These can be used to tie specific textual passages to real-world historical
entities, but also to form parts of networks of co-references (Eide, 2009). The association of things shaped as woven networks (e.g. leaf venation, a spider, or a fishing net) or of technical networks (e.g. in telecommunication) to describe relationships between people is metaphorical. The inferential power of the model leverages on a deep conceptual similarity between the model (the topography of a network) and the object (e.g. kinship of historical characters). It can generate unexpected connections between the objects it represents, which exist ‘only’ metaphorically in a network.

In the example in Fig. 6 we see a historical picture of a man and a woman laying her hand on his. The literature over the reading of this 15th century painting by Jan van Eyck is vast. For example, one interpretation of this image sees it as a claim that the two depicted persons are married; another suggests more subtly that the joining of arms is rather an act of presentation by the man in the picture of the child to be borne in the woman’s womb to the destinatary in the mirror, hence exhibiting the fatherhood of the painter (Lancioni, 2012). Whatever the symbolic link between the figures, the physical link establishes a bond between them. This bond can be associated to, and hence expressed as, a link between two nodes in a network.

There are also other types of links deduced from historical documents that can be expressed using a network model. One is co-reference, for instance in the case where two person references expressed by two different statements, such as names in texts or pictures of identifiable persons, refer to the same person. A source can for instance claim that B and C, the person on the image and a name in a text, refer to the same person. Such claims can also be expressed as links between nodes in a network.

Both these types of links are metaphorical. There are no strings attaching occurrences of names referring to the same historical characters to each other, and there are no connections between historical persons that bear any structural similarity to the topography of a net. The social network in the model is a projection of a conceptual framework. Concepts from our understanding of social relations are combined with a sequential object, the text, and a two-dimensional painting, to form a spatial network model.

But the development and use of such models change our view on history, we start seeing relationships as networks. The network gains hermeneutical
power and makes visible as well as quantifiable aspects of a past family network or societal relations. However, different types of relationships (family vs. co-reference) easily lose their particularity and become ‘just’ links. The chain of signs become greedy and takes over another cognitive space or plane which in fact deals with relations with a different semantics, in our example moving from the plane of assertion of social relations to the plane of assertion of co-reference.

One meaning can trigger others; for example, the links between entities not only connote a relation (e.g. kinship), but their length or thickness might also be interpreted as more or less distance between those entities (i.e. more or less related); in this sense the sign (model) takes a life of its own. A link in the net is just a link, and a documented co-reference relationship becomes like a supposed marriage. This feeds back to our view of the modelled objects; in other words: the context/prior knowledge influences the construction and interpretation of the model, but is also in turn influenced by it.

Common for all three types of models is the inferential power operating at the interplay between their ‘intrinsic structure’ and their ‘extrinsic mapping’ (Kralemann and Lattmann, 2013, p. 3409). Indeed, the features being selected in the modelling process are influenced by contextual elements of different kinds, including hypothesis, scholarly methods and conventions, sample selection, and the technologies being used. However, the inferential and epistemic power of the model relies both on extrinsic and intrinsic aspects of the model relation. In the former case, examples show us how—sometimes with vivid immediacy—similarity of existing verifiable qualities between object and model enable DH modellers to manipulate models to make new sense of those objects. In the latter case, examples show us again how models are conductive to new meaning and further modelling through our exercising of a certain imaginative freedom in selecting salient qualities and associating concepts.

5 Conclusions

In the article, we focused on some aspects highlighted in Kralemann and Lattmann’s semiotic theory of models with respect to the role of ‘context’
in modelling acts and the nature of the ‘representational relation’ between objects and models through practical examples. We believe that these two foci are where modelling practices in DH meet with this semiotic framework in productive ways to explain both formal and open aspects of modelling practices.

We contextualized this framework with specific examples of image-like, relational, and metaphor-like modelling in DH research. Prior knowledge is a sine qua non to create models in the first place and to use them as interpretative tools with respect to the objects they are signs of (Ciula and Eide, 2014). The relationships between modelling processes and interpretative outcomes are neither mechanical nor directly causal (Ciula and Marras, 2016); however, the type of similarity on which modelling relies shapes the interpretative affordances of those ‘anchor’ models. Modelling processes bring about investments and burdens with respect to our knowledge of the objects we model. In particular, models as signs relate to the interpretation of those objects in different ways, from the immediate similarity on the image end of the iconic ‘continuum’ to the imaginative ramifications of conceptual similarity on the metaphorical end. To understand the inferential, epistemic, and heuristic role of models as sign-relations, we need to look at both how they come to be (context; i.e. how we make our prior knowledge explicit and in most cases formalized) and how the similarity relation with the object is used to create meaning (new knowledge).

In summary, studying the ‘single respects’ (Kraelmann and Lattmann, 2013, p. 3401; in Peircian terms ‘the ground of the representantem’) by which a model becomes a sign for an object is useful to explain both the logic and syntax of DH models within specific contexts. It demonstrates how these models are built as well as how the relation with the object is realized, for example, in terms of spatio-temporal modalities. The selection of salient qualities or features to exhibit in the models plays a crucial role both in the creation and interpretation of these models. Such selection is however not necessarily human-driven only. We increasingly use computing algorithms to facilitate or even propose that selection, especially in complex environments where variables are many and interconnected (e.g. pattern recognition in image processing or textual similarity in stylometry).

Our examples showed how the relationship of iconicity between the model and the object being modelled is partly extrinsically determined (it relies on the similarity between the model and the object) and partly guided by intrinsic choices (it depends on theory, conventions, imaginative associations, and prior knowledge). Indeed we showed how the inferential power operates at the interplay between their ‘intrinsic structure’ and their ‘extrinsic mapping’ (Kraelmann and Lattmann, 2013, p. 3409). A future challenge would be to explore how the interplay between intrinsic structures of models (selection of salient qualities) and extrinsic mapping (their iconic ground) develops in the creation of scholarly arguments in the humanities.

From this exploration of the semiotics of models we gained a different way to look at and analyse models: models as a type of signs mediating between the impressions of experience and freedom of association. In future research we aim to combine further studies of modelling practice in DH with interdisciplinary studies of modelling in the sciences and the long tradition of abstraction, representation, and modelling in the humanities to expand the model of models presented here. The main challenge remains to grasp the iterative and generative translation of informal models into formal ones and vice versa.

References


Extensive literature in philosophy of science especially focuses on making computer models with which we can deliberately dwell upon our personal understanding of something of interest for its own sake, and without any functional use yet in mind’ (2006, 146).

1 Following Nersessian, we subscribe to an expanded understanding of reasoning as ‘creative reasoning’ beyond logic and spanning the ‘continuum’ between ordinary and scientific problem-solving. Model-based reasoning is not a simple recipe always leading to correct solutions, and reasoning cannot be equated with logic. Most scientific practice do not fit the traditional philosophical ‘gold standard’ of deductive reasoning. ‘The “hypothetico-deductive” method, which comprises hypothesis generation and the testing of deductive consequences of these, is a variation that focuses the fallibility of science with respect to the premises. This leaves out of the account the prior inferential work that generates the hypotheses. […] In model-based reasoning, inferences are made by means of creating models and manipulating, adapting, and evaluating them. […] Analogical, visual, and simulative modeling are used widely in ordinary and in scientific problem solving, ranging from mundane to highly creative usage. On a cognitive-historical account, these uses are not different in kind, but lie on a continuum.’ (2008, pp. 11–12). We wish to thank Gabor Toth for pointing out the relevance of Nersessian’s work to our research.

2 This echoes of course McCarty’s approach to modelling as ‘orientation to questioning rather than to answers, and opening up rather than glossing over the inevitable discrepancies between representation and reality on which that questioning focuses’ (McCarty, 2005, p. 38).

3 This work of contextualizing modelling within a semiotic approach builds on Kralemann and Lattmann (2013) as well as its recent applications to modelling in DH (Ciula and Marras, 2016; Ciula and Eide, 2014).

4 The distinction between the three types of hypoicons is not meant to be clear-cut. We follow Elleström (2013) amongst others in seeing these types as grades of a ‘continuum’ or even of a development rather than separate categories.

5 Beynon et al. defend such pragmatic or empirical approach to modelling (based on William James’ pluralist philosophy of ‘radical empiricism’) which emphasizes the role of informal semantics over the ‘formal semantics of computuation’ (2006, p. 154). ‘[…] all kinds of conception of model are possible through assuming different kinds of context, observation, and agency’. (2006, p. 155) On the historical contingency of models especially within the context of economics see Morgan (2012, pp. 1–37).

6 Extensive literature in philosophy of science especially focusing on the use of models in the empirical sciences recognizes models (including computational models) as mediators between theory and objects of analysis (e.g. Winsberg, 2003; Morrison, 2009). Within a semiotic context, this finds a parallel in the concept of sign-vehicles functioning as mediators between denotational and connotational qualities, between thing and meaning (MacEachren, 2004, p. 246).

Notes

1 Our pragmatic understanding of modelling is comparable to what Beynon et al. call Empirical Modelling: ‘Model-building in EM [Empirical Modelling] evolves through an extended process of observation and experiment in which exploration and negotiation of meaning play a fundamental role’ (2006, p. 152). In our work we make specific reference to Peirce’s semiotic pragmatism rather than Jamesian pragmatism, since the latter implies a different understanding of experience and hence of the use of the term ‘pragmatism’. See Olteanu 2015 (81–104) for an informed and detailed explanation of this. What is particularly insightful in Peirce’s philosophy for us is his ‘understanding of life in term of phenomena of signification’ (idem: 83), which goes beyond and even against the epistemological account of (relativist) experience in James.

2 This echoes of course McCarty’s approach to modelling as ‘orientation to questioning rather than to answers, and opening up rather than glossing over the inevitable discrepancies between representation and reality on which that questioning focuses’ (McCarty, 2005, p. 38).

3 This work of contextualizing modelling within a semiotic approach builds on Kralemann and Lattmann (2013) as well as its recent applications to modelling in DH (Ciula and Marras, 2016; Ciula and Eide, 2014).

4 The distinction between the three types of hypoicons is not meant to be clear-cut. We follow Elleström (2013) amongst others in seeing these types as grades of a ‘continuum’ or even of a development rather than separate categories.

5 Beynon et al. defend such pragmatic or empirical approach to modelling (based on William James’ pluralist philosophy of ‘radical empiricism’) which emphasizes the role of informal semantics over the ‘formal semantics of computuation’ (2006, p. 154). ‘[…] all kinds of conception of model are possible through assuming different kinds of context, observation, and agency’. (2006, p. 155) On the historical contingency of models especially within the context of economics see Morgan (2012, pp. 1–37).

6 Extensive literature in philosophy of science especially focusing on the use of models in the empirical sciences recognizes models (including computational models) as mediators between theory and objects of analysis (e.g. Winsberg, 2003; Morrison, 2009). Within a semiotic context, this finds a parallel in the concept of sign-vehicles functioning as mediators between denotational and connotational qualities, between thing and meaning (MacEachren, 2004, p. 246).

7 Following Nersessian, we subscribe to an expanded understanding of reasoning as ‘creative reasoning’ beyond logic and spanning the ‘continuum’ between ordinary and scientific problem-solving. Model-based reasoning is not a simple recipe always leading to correct solutions, and reasoning cannot be equated with logic. Most scientific practice do not fit the traditional philosophical ‘gold standard’ of deductive reasoning. ‘The “hypothetico-deductive” method, which comprises hypothesis generation and the testing of deductive consequences of these, is a variation that focuses the fallibility of science with respect to the premises. This leaves out of the account the prior inferential work that generates the hypotheses. […] In model-based reasoning, inferences are made by means of creating models and manipulating, adapting, and evaluating them. […] Analogical, visual, and simulative modeling are used widely in ordinary and in scientific problem solving, ranging from mundane to highly creative usage. On a cognitive-historical account, these uses are not different in kind, but lie on a continuum.’ (2008, pp. 11–12). We wish to thank Gabor Toth for pointing out the relevance of Nersessian’s work to our research.

8 As pointed out by Beynon et al., ‘it is now possible to make computer models with which we can deliberately dwell upon our personal understanding of something of interest for its own sake, and without any functional use yet in mind’ (2006, 146).


11 Note that the term ‘sensuously’ (rather than e.g. ‘sensorial’) occurs here for specific reasons. While one of our current senses of ‘sensuous’ has hedonistic and even erotic connotations, this was not the case for philosophers in the 19th century. For continental philosophy in particular (e.g. Kant and Hegel), the term ‘sensuousness’ is used in connection to the immediacy of nature and in relation or opposition to conceptual understanding. Sensuous encounter is hence considered to be devoid of analytical consciousness and intention. Peirce uses the term to refer to the impression of experience in its (conscious) immediacy as well as individuality situated in space and time with no ontological or moral bearing.

12 On Elleström’s system for media modalities applied to modelling of spatial information in DH, see Eide (2015b).
Note that interpretation involves multiple and intertwined processes of signification; iconic signs are indeed ‘mixed with indexical and symbolic ways of interpreting’ (cf. Elleström, 2013, p. 113).

Various attempts have been made to put such things on maps. See Eide (2015b) for an extensive discussion.

This is exactly what happened in the modelling experiments described in Eide (2015b), where differences between the structures expressed in the text and structures expressible as maps were found. The model could not express what the source object expressed.

In Kralemann and Lattmann (2013) these models are claimed to be based on semiotic similarity, but this appears categorically misleading to us so we privilege the concept of metaphor taken from Peirce.

For a recent discussion on the benefits and pitfalls of the use of network as metaphor in social sciences, see Erickson (2012).

The National Gallery, London, image number NG186.

While it is outside the scope of this article to account for the nuanced and precise terminology adopted by Peirce, it should be noted that he defines a subclass of icons called ‘hypoicons’ which are in their turn divided into images, diagrams, and metaphors; for a recent detailed and comprehensive overview of Peirce’s categories and taxonomy of signs, see Olteanu (2015, 61–79).
Ancient Vase 3D Reconstruction and 3D Visualization

Sorin Hermon
STARC - The Cyprus Institute, Cyprus. s.hermon@cyi.ac.cy

Despina Pilides
Department of Antiquities of Cyprus, Cyprus.

Giancarlo Iannone, Ropertos Georgiou, Nicola Amico, Paola Ronzino
STARC - The Cyprus Institute, Cyprus.

Abstract:
The paper describes the process of 3D virtual reconstruction of an ancient fragmented vessel. The work followed several steps: identification of significant potsherds for the 3D reconstruction, the 3D acquisition of the fragments with laser scanner, the analysis of the 3D model (diameter, vertical projection, orientation and thickness), 3D reconstruction and modelling of the vessel, photographic acquisition and ortho-stereoscopic rendering for 3D visualization.

Key Words: Archaeology, 3D Modelling, Laser Scanning, Open-Source Software, Ortho-Stereoscopy.

Introduction

Many archaeological finds uncovered during excavations are pottery fragments. Archaeologists select identifiable ones in order to assign their type, to understand cultural, economic, chronological and social aspects of the site under investigation. The main steps of potsherds study are: orientation of fragments, diameter estimation, profile estimation and drawing (diameter, vertical projection, profile). Our case study concerns the study and 3D reconstruction of a set of fragments belonging to one vessel, (roughly and partially) restored by the conservation laboratory of the Archaeological Museum of Larnaca, Cyprus. Since conservators were unable to fully understand the original shape of the vessel, we tried to virtually reconstruct it. The main difficulties related to this type of work are: a large amount of small fragments, inability to place them properly along the 3D surface and the inaccurate physical restoration that forced us to define a virtual error correction (Goel and Priyank 2005). All potsherds larger than 10cm were digitally acquired with a laser scanner (multi-stripe laser triangulation) and open-source software were used for post-processing (MeshLab, Blender). During the post-processing the correct orientation of fragments was calculated through geometric analysis; the potsherds physically restored were virtually separated and repositioned in the right way. Once the 3D virtual shape was obtained, the vessel was digitally rebuilt and textured using photographs with colour calibration. Blender software was used for modelling and for the stereoscopic virtual set up of the vessel in order to obtain the “ortho-stereoscopic” rendering. The final result is the 3D model of the vessel, which was used for 3D stereoscopic vision simulation. Illusionary depth perception and immersive view experience allows to better understand the shape and the volume of an object which is unreadable in the fragmented conditions or in traditional restoration.
Case Study

The vase was found in 2010 at the 13th – 12th centuries BC site of Pyla - Kokkinokremos, near the south-east coast of Cyprus during the excavation carried out under the direction of Prof. Vassos Karageorghis (Leventis Foundation), Dr. Athanasia Kanta (Ephoros of antiquities, Heraklion, Crete) and Dr. Maria Hadjicosti (Director of the Cyprus Department of Antiquities) (Karageorghis 2011). When discovered, the vessel was extremely fragmented (approximately 60 pieces) (Fig. 1). It had typical handles, of the “elbow reversed handle” type (Sebis 1996). Based on analyses mostly done in Sardinia and adjacent islands (e.g. Lipari), this vessel typology starts from Late Bronze Age (1300-1950 BC) and continues throughout the first Iron Age (900-750 BC) (Campus and Leonelli 2000). After the recovery, all pieces were sent to the Restoration Department of the Larnaca Archaeological Museum, where they reconstructed only three sets of fragments: base, neck and a part of body with the handle. The position of the majority of the potsherds were uncertain and thus the original shape remained unknown.

Digital Acquisition and Virtual Reconstruction Pipeline

The curator of the Department of Antiquities Dr. D. Pilides decided to involve The Cyprus Institute in the virtual reconstruction process. After a preliminary visual investigation, all sets of restored fragments and some of the bigger potsherds were acquired with a 3D laser scanner. The device used is NextEngine 3D Desktop laser scanner (Abernathy 2007) based on multi-stripe laser triangulation (MLT). It has an internal calibrated camera that allows the acquisition with the RGB colour. Seven pieces were scanned in Macro Mode, low resolution with 127 micron accuracy. The time spent for acquisition was 15min for each piece at 360°, using a rotary servo-positioner. The operating distance was set at 16.5cm (Fig. 2). After the digital acquisition, the obtained range maps have been processed with the open-source software MeshLab, developed by CNR-ISTI of Pisa, Italy, used for the range maps optimization.

Post-processing

Each piece was meshed with Meshlab software, in order to obtain solid 3D models. These meshes were exported in the software JRC 3D Reconstructor (Sgrenzaroli and Vassena 2007) (complete processing software for 3D data), where we were able to virtually cut the 3D models according to horizontal and vertical plans (Fig. 3). The profile thus obtained was exported in .dxf format for AutoCAD, in order to complete the virtual reconstruction process (Andrews and Laidlaw 2002) (Fig. 4).
From each section of each fragment a cutting plane was obtained, which served to draw an arc from which we were able to define the respective diameter and to build a hypothesized circumference of the potsherds. These arcs were aligned on a perpendicular plane to the central axis of the vessel. Subsequently, all circumferences with an equal diameter were overlapped so as to define an alignment and trace the shape of the object (Fig. 5). Finally using the tool bottom loft, the circumferences were merged together using as trajectory the vertical axis generated from the circumferences.

Regarding the height of the vessel, it was decided to test four different heights with a range of 5cm. (35-50). The height of 35cm indicates the minimum height obtained by the curvature of the larger fragment. For each possible height, the handles have been re-positioned, in order to find the correct typological position. The final result demonstrated that the vessel had originally an estimated height of approximately 38cm with handles angulations of 92° to 94° to the body.

Reconstruction Process

After the post-processing and the creation of the possible profile (Kampel and Sablatnig 2008) through the CAD system, the profile was exported into the AutoCAD .dxf format, a drawing interchange format for enabling data interoperability between AutoCAD and other modelling applications.

Blender (Blender 2011) is open source software suitable for our needs supporting the necessary formats for the efficient communication between the two modelling applications. It supports Python scripts for the revolution of the profile around its vertical axis and the 3D stereoscopic simulation of the final result.
Inside the 3D application environment, “spline primitives” are drawn on top of the profile cross-sections by following the path of the line of each digital drawing imported from CAD. This step simulates the half silhouette profile of the vessel represented as NURBS curves. A genetic algorithm “Lathe” was then applied by compiling a Blender experimental Python Script, which creates the vessel’s volume, leading to a robust reconstruction of its whole. Each scanned piece was then aligned along the surface of the reconstructed vessel in order to check matching and find its probable original location (Fig. 6).

**Texturing Process**

A High Resolution camera (Nikon D3x) was used for the photographic acquisition of the fragments using a standard colour reference, determining an optimal exposure index and dynamic range.

The images were used to create the texture map of the model and to create a ‘template map’ from which the colours were extracted. The template map has been applied to the 3D model using the UVW mapping technique, preventing graphic artefacts which accompany more simplistic texture mappings such as planar projection (Fig. 7).

**Visualization**

The visualization process follows an “orthostereoscopic” (Schneider 2011) approach to the simulation of the vessel by animating the still left/right imagery captured from the Blender (Blender 2011) environment through real environment calculations, such as the horizontal size of the projection screen and the average distance of the viewer to the projection screen (Fig. 8). This is done by running and compiling a Python script for creating two off-axis virtual cameras with the analogous adjustable attributes for the correct visualization of the artefact, without distortion of geometry, having a 1:1 ratio of real and virtual spaces, thus conveying the correct visual perception to the spectator.

**Conclusions**

In this article we have described a method of virtual reconstruction by the digital acquisition of the fragments using a very cheap laser scanning: NextEngine 3D laser scanner (Guidi et al. 2007). The post-processing work has been done with MeshLab, open-source software with a relatively easy-to-use interface but with a wide community of users and excellent technical support. Once we have obtained the 3D models of each fragment, the virtual restoration and visualization were performed within CAD and...
Blender systems (for profile extraction of each fragment and its location in the Euclidian space). Future works will consist of creating a methodology that solves the problems related on understanding the correct location of the pottery fragments on the original vessels and on finding a semi-automated way to obtain the original shape of very fragmented vases based on pottery fragments.

The study reconstruction of objects such as ceramic vessels today can make a productive use of 3D technologies and advanced display systems, achieving results that were unthinkable only a few years ago. The study of an object in every detail, seeing colours and shapes in three dimensions, has become an indispensable support for researchers in this field. Thanks to technology, the case study proposed here allows you to open a window to our past and investigate typologically a more or less absolute chronology (in this case involving the elbow reversed handle). In any case, the real goal is that thanks to virtual reconstruction historical and archaeological research can be narrated and represented as never before.

**Acknowledgments**

The authors want to thank prof. Vassos Karageorgis (director of the excavation), for allowing us the digital acquisition of the potsherds and for his support in the interpretation process. This study is the result of collaboration between: STARC (Science and Technology in Archaeology Research Center), Cyprus Department of Antiquities, The A.G. Leventis Foundation. This research benefited from the support of the EU-funded project 3DCOFORM (www.3d-coform.eu) (FP7/2007-2013) under grant agreement n° 231809.

**Bibliography**


An anonymous Latin treatise of the 4th century AD, known as *De Rebus Bellicis*, describes, among a number of new mechanical contrivances, which in the opinion of its author ought to form part of the equipment of the Roman army, an ox-powered paddlewheel *liburnian*. There is no other data to accredit the use of this system, no archaeological or iconographical evidence so researchers are dubious about this text. One of the first modern editors of *De Rebus Bellicis*, R. Schneider (1908), wrote: “Die Liburna ist geradezu verrückt” (“The liburnian is just crazy”), and recently, in 1989 A. Giardina considered that, if the chariot *à faux*, described by Anonymous was at first a reality on the battlefield, and then a legend of military engineering, the paddlewheel boat is exactly the contrary: born as a fantastic machine powered by oxen, it is steam which finally provided the invention with a real application (Giardina 1989). The aim of this paper is to show that this self-propelled *liburnian*, described in *De Rebus Bellicis*, can be reconstructed with the text transmitted (and its illustrations) and with indications provided by technical knowledge in antiquity. The virtual reconstruction will show that the *liburnian* can work and that it could have been built and used in antiquity. It is an important challenge for the history of technology because, if this hypothesis is correct, the first propulsion of a boat with a paddlewheel should no longer be ascribed to Denis Papin at the end of the 17th century (with a steam engine), or to Robert Fulton at the beginning of the 19th century for the operational model: this technology would thus go back thirteen centuries at least, with an animal-driven engine. First, we will look to the text itself, its reliability, and the illustrations in the manuscripts. Then, we will study a history of the techniques in order to find indications lending credence to the system. To conclude, the virtual reconstruction of the *liburnian* will be shown.

**Abstract:**

This paper will focus on the virtual reconstruction of a war machine, an ox-powered Roman paddle wheel boat used for its impressive speed during naval manoeuvres. The ship is described and drawn in an anonymous 4th century work but is often considered as imaginary by researchers. This is the first time, to our knowledge, that the reconstruction of this ship has been attempted and we shall show, through the experiment with a 3D model, that this warship of antiquity can work and, even more so, that it is probably the ancestor of the paddle steamer. Through this reconstruction, we would like to insist on the value of experimental history produced using scientifically justified virtual reconstructions, a medium of the future to develop innovative educational applications and write new pages in the history of science and technology.

**Key Words:** Virtual Reality, History of Technology, Roman Machinery
Ancient Source Materials

_De Rebus Bellicis_ is a short anonymous treatise in which the author sends to the emperor some proposals for reforms in fiscal, monetary and military matters (Condorelli 1971; Giardina 1989; Jouffroy 2004; Reinach 1922; Thompson 1952). The author was not a “mechanical engineer” as was Vitruvius, the author of _De Architectura_ in the first century AD. If we had to qualify him by modern notions, we would speak about a specialist in economic and financial affairs. The title given to his treatise (“Military Affairs”), probably in a late period, is inappropriate since war machines concern no more than half of the work. At the time of the writing, the author was not exercising an official function, but the vocabulary which he uses to deal with administrative and financial questions indicates that he was probably a state employee of the imperial civil administration, in close contact with the environment and culture of _apparitores_, a class of state employees in the service of Roman magistrates which also included Vitruvius. As a result he is reliable and his proposals are probably real innovations. The machines he describes are not only intended to save manpower which was becoming scarce. What he suggests is an increase in the power of the army thanks to new machines. His treatise, in a way, is not a technical treatise: no one, either today or in antiquity, would be able to build the machines proposed by the anonymous author of _De Rebus Bellicis_ with only the instructions given in the text. It is a theoretical book. Its dating has been much debated but today it is generally agreed that it was written between the reigns of Constance II (337-361 AD) and Theodosius (379-395 AD). In a recent edition, R.I. Ireland suggests specifically that the book would have been dedicated to Valens at the end of 368 or at the beginning of 369 AD (Ireland 1984).

The text was transmitted via four main manuscripts of the 15th and 16th centuries. Three passages concern the self-propelled liburnian.

1 Anon., _De rebus bellicis_, Praef. 12.
2 Anon., _De rebus bellicis_. 17.

It is important to note that the facts are identical in the four main manuscripts.

I shall, in fact, demonstrate how a particularly fast type of warship is able through a brilliant invention to outmatch ten other ships, sending them to the bottom without the aid of a large crew.

A Liburnian ship suitable for naval warfare, so large that human weakness more or less precluded its being operated by men’s hands, is propelled in any required direction by animal power harnessed by the aid of human ingenuity to provide easy locomotion. Inside its hull or hold pairs of oxen are yoked to the machines and turn wheels attached to the ship’s sides; the spokes project beyond the circumference or rim of the wheels, and, striking the water forcibly like oars as the wheels rotate, work with a wondrous and ingenious effect, their impetus producing locomotion. Moreover this warship, thanks to its massiveness and the machines working inside it, joins battle with such furious strength that it easily crushes and destroys all opposing warships that come to close quarters with it.

But if the enemy flees the land and besets the seas with naval warfare, victory will be restored to you without delay by a warship swiftly traversing the...
waves by means of a new speed device, for it is propelled by wheels and oxen as though it were on land. For who will resist its strength when it has the firmness of a land vehicle and has the advantage of a ship’s easy movement?³

De Rebus Bellicis was accompanied by illustrations, announced by the author himself and preserved in three of the four manuscripts⁴. These illustrations were probably produced by the author himself or at least checked by him. They thus provide an essential basis for the reconstruction. Concerning the liburnian, variants between the three versions are of little significance.

**Indications Confirming Ancient Source Materials**

Therefore, we have a reliable text and an illustration which has accompanied the text from the time of its original writing in the late 4th century AD, but there is no reliable evidence which would allow us to prove that the system was built and used by the Romans. However, a series of indications show that this Roman paddlewheel boat powered by oxen could have been used before the invention of the steam engine and even before the 4th century: on the one hand paddlewheels, gear transmissions and animal mills are widely known in antiquity; on the other hand, while the use of animal mills on water seems to have been a rare practice, it is attested to, for certain, in the 19th century.

The paddlewheel has been known since at least the 3rd century BC: Philo of Byzantium describes in his *Pneumatics* a small paddlewheel which turns using the force of water. In the 1st century BC, Lucretius refers to it and Vitruvius describes water-driven wheels for raising water⁵ (Fleury 1993; Lewis 1997; Wikander 1981).

Following this text, Vitruvius describes a water mill which is based, he says, on the same principle (Fig. 2): its mechanical elements are exactly the same as the ones of the liburnian except that water imparts motion to the paddles instead of *vice versa*. During the siege of Rome

---

³ Anon., *De rebus bell.* 18, 9-10.
⁴ Anon., *De rebus bell.* 6, 5.
⁵ Vitr. 10, 5, 1.
in 537-538 AD, Belisarius installed water mills on boats anchored in the Tiber, but this is not the only evidence of paddlewheels on ships: Vitruvius and Heron of Alexandria (1st century AD) had already described an instrument that indicates the distance travelled by a ship with paddlewheels (Figs 3 and 4).6

Water mill and odometer imply a gear transmission. These transmission systems were also used in engines for raising water, in machines for sawing, and maybe also in hoisting machines. Besides textual sources we have for gears, archaeological and iconographic documentation which is certainly not very plentiful but which is sufficient to give evidence of their use in antiquity. These systems were still used in wind mills and water mills in the 20th century (Fig. 5).

The paddlewheels mentioned above drive wheels which activate a system. Except in De Rebus Bellicis, we have no examples of paddlewheels activated by a motive force and used for propulsion. However the type of engine considered here, animal mills, was also well known in antiquity, in particular for the mills of the Pompeian model (Fig. 6) and in systems for raising water. It is the principle of the “Persian wheel” or saqiya, a system used in Egypt from the 2nd century BC Is it possible to transpose the principle of the animal mill to a ship? We do not have any serious evidence either from Antiquity or from the Renaissance but, in the 19th century the system is attested in England and described by an engineer who saw it working:

“At Yarmouth the Horse Packet is about 60 feet in length and 18 feet abeam. It is worked by four horses in a file which walk in a circle 18 feet in diameter in which they are much too confined and so do only half the work. The drive-shaft has two bevelled wheels, one at each end, by which the motion is communicated from the horses to the axle of the paddle-wheels 7 feet in diameter. The boat goes at the rate of about six miles an hour” (Stevenson 1946, 59).

On the other hand the usage of animal mills with horses on rafts is attested in Finland in the 20th century for floating wood (Figs 7 and 8).

**The Reconstruction of the Liburnian**

From all these elements, we have reconstructed the self-propelled liburnian described in De Rebus Bellicis (www.unicaen.fr/ersam).
The ship, the system’s support, is not part of the project of research presented here in the first version; thus its reconstruction is of a summary and little detailed nature. Vegetius, a contemporary of the anonymous source, uses the word *liburna* in the sense of “warship”. He considers it is the type of ship adopted by the Roman navy since the battle of Actium in 31 BC. For him, liburnians can be of several sizes:

*With respect to size: the smallest ships of war have a single tier of oars; those a little larger, two; those of good size allot three, four or sometimes five positions for the rowers*.  

Thus, we shall distrust some modern definitions determining the liburnian as a “low and fast” ship or some definitions of authors further from the anonymous source than Vegetius, such as Lucan or Appian who assimilate the liburnian to a ship with two rows of oars on each side. In reality the word “liburnian” used by the anonymous source does not inform us about the size of the ship: it can be a light or heavy unit. Prudentius, another contemporary, speaks of a liburnian equipped with a tower, thus of

---

7 Veg., Mil. 4, 37: *Quod ad magnitudinem pertinet, minimae liburnae remorum habent singulos ordines, paulo maiores binos, idoneae mensurae ternos uel quaternos interdum quinos sortiuntur remigio gradus.*

8 Lucan. 3, 533-534: Appianus, Ill. 3.

9 Prud. C. Symm. 2, 530-531.
a ship of respectable size (Casson 1971; Starr 1960). For the virtual experiment, we modelled a ship with a length of 35.75m and a beam of 5m. We kept on each side the outrigger for the oars: this is the structure for the paddlewheel axes. The breadth, including outriggers and paddlewheels, is 7.5m. These dimensions allow the placing of three animal mills with two oxen (Fig. 9), as in the manuscripts’ illustrations. The anonymous source does not specify the number of animal mills; he speaks simply of “pairs of oxen yoked to the machines”\(^{10}\). In this first version of the virtual model, as in the manuscripts’ illustrations, we placed oxen on the deck of the ship to allow a better visualization of the system, but it is more likely they were placed below. Placing animal mills inside the hull offers a triple advantage: 1) animals are protected from missiles thrown by the enemy, 2) the centre of gravity of the ship is lowered, and 3) the deck is clear for soldiers. It is also what the anonymous source means by the expression *In cuius alueo uel capacitate*, “In its hull or hold”. But whatever the position of the animal mills, above or below the deck, the principle of mechanical transmission does not change: it is a question of transforming rotation on a horizontal plane into rotation on a vertical plane.

The virtual reconstruction allows carrying out different experiments. Thus, we suggest fixing to the animal mill’s axis of rotation a wooden cogwheel which pulls a lantern, united itself to paddlewheels which are fixed on each side of the ship (Fig. 10). In our reconstruction, the paddlewheels have a 3.5m diameter, and thus a circumference of 11m. An ox, depending on its race and weight, can walk at an average speed of between 2 and 3km/h and supply a one-time effort of up to 4.5km/h. We multiplied the speed of the oxen by three with the system of gears. In order for the liburnian to advance at 5 knots, the oxen have to walk at 2.6km/h; in order for it to advance at 8 knots, they have to walk at 4.2km/h. Other calculations were made concerning the power developed by the oxen. They show that the total power is relatively low (3hp or 2.1kW) but that it allows the liburnian to advance.

**Conclusions**

The virtual reconstruction is a way to show that the liburnian can work (however maybe not as fast and with as much power as the anonymous source claims) and that the system is not “crazy” or “purely imaginary”. Therefore it is possible that this boat could have really been built. Why did not the system become more widespread? Why was it necessary to wait for the steam engine to see the development of ships with paddlewheels? At first, the invention of the anonymous source would have a limited use in the navy since the paddlewheel system is cumbersome and vulnerable during a fight.

\(^{10}\) Anon., *De rebus bell.* 17,3: *bini boues machinis adiuneti.*
Figure 10. Mechanism of the liburnian (4th century AD).

Even at the time of the introduction of steam propulsion in the navy in the 19th century, ships with paddlewheels made only a brief appearance and they were very quickly replaced by ships with propellers (Mollat 1970, 9). Secondly this ship is not intended for the open sea: oxen (and other animals) can only work on a calm sea while rowers can contend with pitching and rolling even if the naval actions of antiquity took place only in relatively calm weather. This limits the practical use of this liburnian to rivers, to their mouths and to sheltered bays. But, after all, at this time this was the Roman navy’s main field of intervention.

Acknowledgements

The English text has been proofread and edited by Alice and Edward Mills. We address them all our gratitude.

Bibliography


All Reality: Virtual, Augmented, Mixed (X), Mediated (X,Y), and Multimediated Reality

Steve Mann, Tom Furness, Yu Yuan, Jay Iorio, and Zixin Wang

ABSTRACT
The contributions of this paper are: (1) a taxonomy, framework, conceptualization, etc., of the “Realities” (Virtual, Augmented, Mixed, Mediated, etc.), and (2) some new kinds of “reality” that come from nature itself, i.e. that expand our notion beyond synthetic realities.

VR (Virtual Reality) replaces the real world with a simulated experience (virtual world). AR (Augmented Reality) allows a virtual world to be experienced while also experiencing the real world at the same time. Mixed Reality provides blends that interpolate between real and virtual worlds in various proportions, along a “Virtuality” axis, and extrapolate to an “X-axis” defined by Wyckoff’s “XR” (eXtended reality), and Sony’s X-Reality™.

Mediated Reality goes a step further by mixing/blending and also modifying reality. This modifying of reality introduces a second axis called “Mediality”. Mediated Reality is useful as a seeing aid (e.g., modifying reality to make it easier to understand), and for psychology experiments like Stratton’s 1896 upside-down eyeglasses experiment.

We propose Multimediated Reality (“All Reality” or “All R”) as a multidimensional multisensory mediated reality that includes not just interactive multimedia-based “reality” for our five senses, but also includes additional senses (like sensory sonar, sensory radar, etc.), as well as our human actions/actuators. These extra senses are mapped to our human senses using synthetic synesthesia. This allows us to directly experience real (but otherwise invisible) phenomena, such as wave propagation and wave interference patterns, so that we can see radio waves and sound waves and how they interact with objects and each other.

Multimediated reality is multidimensional, multimodal, multisensory, and multiscale, including not just “wearables” but also smart buildings, smart cities, etc.. It is also multidisciplinary, in that we must consider not just the user, but also how the technology affects others, e.g. how its physical appearance affects social situations. Finally, it is multiveilant (surveillance, data-veillance, and other “veillances”). For example, cameras in multimediated reality devices affect the privacy of non-users of the technology as well.

KEYWORDS
Virtual reality, VR, augmented reality, AR, mixed reality, mediated reality, wave propagation, education, physics, lock-in amplifier, standing waves, sitting waves

Figure 1: Above: Mixed Reality Continuum, adapted from Milgram and Kishino, 1994 [26]. The blue arrow is suggestive of a one-dimensional “slider” or “fader” that “mixes” between the Real world and the Virtual world, as illustrated below:

Real Environment → Mixed Reality → Augmented Reality (AR) → Augmented Virtuality (AV) → Virtual Reality (VR)

Figure 2: Disk Jockey (DJ) Mixer Metaphor: Imagine two record players (turntables), feeding into an audio/video mixer. Real-world and virtual world mixtures are selected by sliding a one-dimensional “fader” left or right. This allows us to choose various points along an “X” axis between the extremes of Reality, “R”, and Virtuality, “V”.

1 HISTORICAL BACKGROUND & CONTEXT
1.1 Virtual, Augmented, Mixed, and X-Reality
VR (Virtual Reality) is a computer-generated simulation of a realistic experience. Typically VR blocks out the real world (“Reality”) and replaces it with a “Virtual” world. The virtual world may be generated by a computer, or by interactively playing back recorded media. An example of the latter is the Aspen Movie Map of 1978 that used computers to play back analog laser disk recordings to render an interactive virtual world as hypermedia [28], or, more recently, Google Street View with Earth VR.

AR (Augmented Reality) is a similar concept, but instead of blocking out reality, the computer-generated content is added onto, or embedded into, the real world experience, so that both can be experienced together [3].

It has been suggested, by Milgram and Kishino [26], that Augmented Reality exists along a continuum between the real and virtual worlds, giving rise to “mixed reality” (see Fig. 2). In this context we can think of AR as a setting on a “mixer” or “fader” or “slider” that is somewhere between reality and virtuality.

This “slider” is analogous to the “X-axis” of an X-Y plot or graph, treating “X” as a mathematical variable that can assume any quantity on the real number line. Thus mixed reality is sometimes referred to as “X-reality” or “XR” [4, 15, 29]. Specifically, a 2009 special
issue of IEEE "Pervasives computing" on "Cross-Reality Environments" defines X-reality as a proper subset of Mixed Reality. Paradiso and Landay define "cross-reality" as:

"the union between ubiquitous sensor/actuator networks and shared online virtual worlds ... We call the ubiquitous mixed reality environment that comes from the fusion of these two technologies cross-reality." [29]

In that same issue of IEEE "Pervasives computing", Coleman defines "cross-reality" and "X-Reality" as being identical:

"Cross-reality (also known as x-reality) is an informational or media exchange between real-and virtual-world systems." [4]

XR as extrapolation ("extended reality" or "extended response") dates back as early as 1961 when Charles Wyckoff filed a patent for his "XR" film which allowed people to see nuclear explosions and other phenomena beyond the range of normal human vision [8, 40, 41]. In 1991, Mann and Wyckoff worked together to build "XR vision" devices into wearable computers (AR/VR headsets, etc.) for human augmentation and sensory extension by way of High Dynamic Range (HDR) imaging blended with virtual/augmented reality [15].

The terms "XR", "X-Reality", "X-REALITY", and "XREALITY" appear as trademarks registered to Sony Corporation, filed in 2010, and used extensively in the context of mobile augmented reality across Sony’s “Xperia” X-Reality™ for mobile products, as shown in Fig. 3 below:

![Figure 3: Sony’s trademarked X-Reality and XR](image)

Sony’s use of XR and X-Reality is consistent with the Wyckoff-Mann conceptualization of extended human sensory perception through high dynamic range.

There is some confusion, though, since XR (X-Reality) now has at least three definitions, one in which it is a proper superset of mixed reality, another in which it is mixed reality, and another in which it is a proper subset of mixed reality. We shall enumerate and classify these, chronologically, as follows:

- **Type 1 XR/X-Reality** in which "X" as a mathematical variable, i.e. any number on the real number line, that defines an axis for either:
  - **Type 1a XR/X-Reality: extrapolation**, i.e. XR/X-Reality in Wyckoff-Mann sense (1991), as technologies that extend/augment/expand human sensory capabilities through wearable computing. In this sense "X" defines an axis that reaches past "reality".
  - **Type 1b XR/X-Reality: interpolation**, i.e. XR/X-Reality in the Milgram sense (1994), as technologies that augment human senses by creating a blend (mixture) between the extremes of reality and virtuality. In this sense "X" defines an axis that miXes (interpolations) between reality and virtuality.

- **Type 2 XR/X-Reality** in which "X" means "Cross" in the Paradiso-Landay/Coleman sense (2009), i.e. as a form of mixed reality (a proper subset of mixed reality) in which the reality portion comes from sensor/actuator networks,

and the virtuality portion comes from shared online virtual worlds.

The taxonomy of these three definitions of XR/X-Reality is summarized as a Venn diagram in Fig. 4, showing also XY-Reality (XYR) which will be defined in the next section.

What these definitions of XR/X-Reality all have in common is that XR/X-Reality defines an "X-axis" defining a number line that passes through both "reality" and "virtuality".

1.2 Mediated Reality (X-Y Reality)

Many technologies function as an intermediary between us and the environment around us. Technology can modify or change (mediate) our "reality", either as a result of deliberate design of the technology to mediate reality, or sometimes as an accidental or unintended side-effect. These two variants of mediated reality are further discussed below.

1.3 Deliberately mediated reality

Examples of deliberate modification of reality include the upside-down eyeglass invented 122 years ago by George Stratton to study the effects of optically mediated vision on the brain [36]. More recently others have done similar experiments with deliberate mediation of reality, such as left-right image reversing eyeglasses [7].

Multimedia devices such as hand-held camera viewfinders have also been used to study long-term adaptation to a photographic negative view of the world in which light areas of the image are made dark, and dark areas of the image are made light [1]. Computer-mediated reality has also been explored [13, 16].

Mediated Reality is not just for psychology experiments, though. It has many practical everyday applications such as eyeglasses that filter out advertisements, and, more generally, helping people see
better by getting rid of clutter. HDR (High Dynamic Range) welding helmets use computer vision to diminish the otherwise overwhelming brightness of an electric arc, while augmenting dark shadow detail. In addition to this Mediated Reality the HDR welding helmet also adds in some virtual content as well [6]. Mediated Reality has also been examined in the context of wearable computing, prosthetic, and veillance (surveillance, sousveillance, metaveillance, and dataveillance) [9, 37].

1.4 Unintentionally Mediated Reality

In Augmented Reality there is often an attempt made to not alter reality at all. But when we experience augmented reality through a smartphone or tablet, or by using video see-through eyeglasses, the simple fact that we have a technology between us and our outside world means that the virtual objects overlaid onto reality are actually being overlaid onto an unintentionally modified reality (ie. both the virtual and real objects are presented by a video display device). Thus the Mediated Reality framework is directly applicable to the research and understanding of video-see-through implementations of AR.

It has also been observed that head-mounted displays can cause semi-permanent or permanent and lasting harm (e.g. brain damage) or good, such as in the treatment of PTSD (Post Traumatic Stress Disorder), phobias, and in treating (reversing the effects of) brain damage [30, 31, 33, 35, 39]. The fact that HMDs can both damage the brain as well as treat brain damage suggests that we need to be extra careful when using technology as an intermediary, and that there is a special importance to mediality in general.

1.5 The Mediated Reality (X,Y) Continuum

Summarizing the previous two subsections: consider either of the following:

- devices and systems designed to intentionally modify reality;
- the unintended modification of reality that occurs whenever we place any technology between us and our surroundings (e.g. video-see-through Augmented Reality).

Both of these situations call for at least one other axis beyond the mix between reality and virtuality.

Moreover the above are just two of the many more examples of “reality” technologies that do not fit into the one-dimensional “mixer” of Fig 2, and thus we need at least one additional axis when describing technology that specifically modifies reality. For this reason, Mediated Reality [10, 13, 37, 38] has been proposed. See Fig 5. In this Mediated Reality taxonomy (continuum), there are two axes: the virtuality axis (“X”) exactly as proposed by Milgram, and a second axis, the Mediality axis (“Y”). This allows us to consider other possibilities like mediated-augmented-reality (“augmented reality”) [6] (e.g. HDR welding helmets), as well as mediated virtuality (e.g. taking an existing VR system and then flipping the image upside-down, to allow us to repeat George Stratton’s 1896 upside-down eyeglass experiment but in a virtual world).

2 MULTIMEDIATED REALITY

2.1 Technologies for sensory attenuation

The Milgram Continuum (Fig 2) (Milgram 1994 [26]) and the Mann Continuum (Fig 5) (Mann 1994 [13]) both place reality at the left or the lower left, i.e. the “origin” in Cartesian coordinates.

Neither Milgram’s nor Mann’s Continuum directly addresses visual sensory attenuation technologies like sunglasses and sleep masks, or attenuation of other senses by such technologies as ear plugs or sensory attenuation tanks (also known as “sensory deprivation tanks” or “flotation tanks”).

Other visual useful sensory attenuation devices include the sun visor of a car, the brim of a baseball cap, or the “blinders” attached...
to a horse’s bridle so that the horse is not distracted by peripheral motion cues.

Sensory attenuation technologies form an underexplored yet richly interesting space for technology. Consider for example, some of the following possibilities:

- Interactive sleep masks for shared lucid dreaming.
- Interactive multimedia bathing environments like computer-mediated sensory attenuation tanks, and interactive relaxation tanks that use water for sensory attenuation (Mann 2004), as well as “immersive multimedia” and “Fluid User Interfaces” [18] that use water to alter the senses in conjunction with interactive multimedia. See Fig. 7.
- Interactive darkroom experiences such as interactive light-painting with reality-based media such as persistence-of-exposure and “Phenomenological Augmented Reality” (e.g. being able to see radio waves and see sound waves by way of a darkroom environment with eyes adjusted to the dark).

2.2 Multimedia in Photographic Darkrooms

A light bulb waved around in a dark room will tend to create the visual appearance or impression of shapes, curves, patterns, etc., by way of a “persistence-of-exposure” effect in human vision as well as in photographic or videographic media. There is a long history of photographic “lightpainting” (http://lpwa.pro/event/15). There is also a well established “flow arts” community doing artistic dance in a dark environment with light sources, e.g. LED (Light Emitting Diodes), as well as “fire spinning” and juggling light-emitting objects as a medium of creative expression. Flow art is similar to lightpainting but for direct viewing rather than through photography. Some tools (specialized light sources) and devices are used for both lightpainting and flow arts.

The tradition of darkness (sensory attenuation) combined with sensory media (e.g. controlled lighting) dates back to the early days of theatre. Theatrical productions typically take place in a space in which all or most of the walls are painted black, and there is usually a black floor, and black curtains, such that lighting can be controlled carefully. In fact the world’s first use of the term “Virtual Reality” came from theatre in 1938 [2].

The tradition of sensory attenuation and controlled sensory stimulus connects well to multimedia: Morton Heilig produced the “Sensorama” (U.S. Pat. #3050870), a multi-sensory experience which was also the world’s first “3D film”.

2.3 Multimediated Reality Darkroom

In the 1970s, the idea of an interactive darkroom was taken a step further, by conducting a series of experiments to make otherwise invisible phenomena visible. These experiments involved light bulbs connected to the output of powerful amplifiers that were driven by transducers or antennae that sensed a physical quantity of interest.

In one example, a light bulb was used to “sweep” for video “bugs” and the light bulb glowed more brightly when in the field of view of a surveillance camera, than it did when not in the camera’s field of view (Mann 2014). This works very simply by using video feedback: a receive antenna is connected to the input of a very sensitive lock-in amplifier with extremely high gain. The lock-in amplifier was

![Figure 6: Submersive Reality (SR): Meditation in a VR/AR/MR Flotation tank. (a) Waterproof 24-electrode underwater EEG (ElectroEncephaloGram) system is part of the many multidimensional inputs to a real-time computer system (b) Interactive multimedia VR/AR/MR environment with water pumps, waves, vibration, lighting, and data projectors controlled by brainwave mediation/meditation[17, 21].](image)

![Figure 7: We also developed submersive reality for (a) musical performances, underwater concerts, etc., and (b) physical fitness, e.g. swimming in virtual environments and interaction with water jets [17, 21].](image)
Figure 8: A very sensitive lock-in amplifier picks up weak radio signals and drives a powerful light bulb causing video feedback. The amplifier "locks in" on the signal that is due to itself, and thus glows brilliantly whenever it is within the field of view of the surveillance camera. In this way, waving the light bulb back and forth "paints" out an image that reveals the otherwise hidden "sightfield" of the camera.

specially designed to drive high capacity loads such as powerful electric light bulbs (1500 to 2500 watts).

When the light shines on a camera it causes the camera to exhibit small but measurable changes, causing video feedback. Waving the light bulb back and forth in front of the camera makes the "sightfield" of the camera visible. See Fig. 8. It has been suggested that this is a form of augmented reality (Mann 2014) but it is a special kind of reality in the sense that it comes directly from nature itself. Unlike many other forms of reality, it does not come from a computer simulation. Here the light bulb filament has a dual role: it is both the mechanism by which a physical quantity is sensed, and it is also the display mechanism. Therefore, due to the fundamental physics of the situation, the alignment between the "real" physical world, and the "augmented" world is exact (there is no need for any tracking mechanism since the process itself is self-tracking).

Figure 9: The electrical signal from a microwave motion sensor (radar) is received by an antenna that moves together with a light bulb on an XY plotter. The bulb traces out the waveform of the received signal in exact alignment with physical reality. For example, we can measure the distance between cycles of the waveform as being exactly $(300,000,000 \text{ m/s}) / (10.525 \text{ GHz}) = 2.85\text{cm}$. Note the hand grasping the bulb, making it possible for a human user to feel the radio wave in addition to seeing it. Thus we have a visual and haptic multimediated reality experience that extends our senses beyond the five human senses, and maps this "sixth sense" (radio) onto two or three (we can also hear it) existing senses.

We proffer to call this Phenomenological Reality, because it makes visible true physical quantities by way of directly physical means, i.e. a direct connection between a physical quantity and the sensed quantity.

When the quantity we wish to sense is not the same thing as light itself, we can use a separate sensor but mount it directly to the light bulb. For example, Fig. 9 shows a light bulb being moved by an XY plotter driven by the output of a lock-in amplifier. It is important to emphasize here that the path the light bulb takes is in perfect alignment with the physical quantity, and thus provides us with a direct glimpse into the physical nature of the signal that is being shown. Here the quantity that we’re measuring is electromagnetic energy, phase-coherently, with an antenna affixed directly to the light bulb. A special technique is used to transform the space in such coordinates that the speed of light is exactly zero, and thus we can see the radio wave sitting still. The first such photograph of radio waves was made using a linear array of light bulbs, controlled by a wearable computer, in 1974 [20]. See Fig10 (See [19] for an explanation of the methodology of photography of radio waves and sound waves.)

Fig 11 shows an example of Phenomenological Reality directed to seeing sound. Here a linear array of 600 LEDs forms a simple dot-graph indicator of voltage (zero volts at the bottom and 5 volts at the top), and is connected to the output of a lock-in amplifier driven by a moving sound transducer referenced against a stationary sound transducer. The relative movement between the two transducers causes corresponding relative movement along the phase front of the sound wave, thus making it visible in coordinates in which the speed of sound is exactly zero (Mann 2016). Because the waveform
Figure 10: Early photograph of electromagnetic radio waves using the Sequential Wave Imprinting Machine (S. Mann, July 6, 1974, from Wikimedia Commons). Center: Sequential Wave Imprinting Machine exhibited at Smithsonian Institute and National Gallery, showing the linear array of 35 electric lights and the early wearable computer with lock-in amplifier and antenna sensor. Right: Early (1980) version of the apparatus with early wearable multimedia computer prototype.

There are 21 cycles of this sound wave over its 1.5 metre distance of travel. Each cycle is \(150\text{cm}/21 = 7\text{cm}\) long. This moving speaker emits a 5000 CPS (Cycles Per Second) tone, which this microphone hears. Row of green lights moves with speaker and displays output of lock-in amplifier from microphone. Measured speed of sound = \(0.07\text{ metres/cycle} \times 5000\text{ cycles/second} = 350\text{ m/s}\). Since the temperature was 27 degrees C, we know from theory that the speed of sound is 347 m/s. Thus we have an experimental error of approximately 0.86 percent.

This multimedia display technologies such as video, as well as special eyeglasses, can be used to sample and hold the data captured by a moving sensor. A multimedia darkroom setup of this type is shown in Fig. 12 and Fig. 13. In this way a volumetric dataset is retained as the 3D waveform interference pattern is scanned through space. While prior SWIM (Sequential Wave Imprinting Machine) techniques (e.g. long exposure photography) are able to capture 2D and sometimes 3D waveforms, they are recorded from a single two-dimensional perspective. By reading and storing radio wave or sound wave information in point-cloud form, it can be reconstructed, manipulated, and analyzed a-posteriori in three dimensions, allowing for infinite angles and slicing. This gives a new level of richness to the observation of the wave, and this richness can be explored using the multimedia reality eyeglasses.
Figure 12: Multimediated Reality darkroom with 3D mechanical position control device that scans a space with transducers connected to a special lock-in amplifier (visible in lower left portion of upper right picture). The signal generator drives the transmit transducer here at 40kHz (signal generator waveform appears on the oscilloscope that’s sitting on top of the amplifier). Top left: with the lights on (HDR photograph); Top right: with the lights off, the sensory attenuation experience helps us concentrate on the multimediated reality that shows interference patterns between two sound sources. Bottom: Experimental apparatus for multimediated reality. An XY(Z) plotter carries a listening device (transducer) together with an RGB (Red Green Blue) LED (Light Emitting Diode) through all possible positions in space. At each position the sound is sensed phase-coherently by way of a L.I.A. (Lock In Amplifier), while sound is produced by a transmit array comprised of two transmitters, receiving the same signal to which the L.I.A. is referenced. The outputs Re (Real) and Im (Imaginary) of the L.I.A. are converted to RGB values for display on the LED. A picture is taken of this movement and presented to a video display, to provide a persistence-of-exposure. Alternatively the video display may be driven directly by data stored in the Control & Data Logging system. In this case, it can be animated by multiplication by a complex number of unit modulus, so that the waves on the screen slowly “crawl” at any desired speed-of-sound (e.g. the speed of sound can be set to zero or to some small value so as to be able to see it clearly).

2.4 Comparison with existing measuring instruments

Many measuring instruments allow us to use one of our senses to measure a quantity that pertains to another sense. For example, we can have a photographic darkroom light meter that emits sound, so that we can use it in total darkness. An oscilloscope allows us to see sound waves. A multimeter is a device that allows us to see or hear electrical signals which would otherwise be difficult or impractical to sense directly. What is unique about multimediated reality as compared with these traditional measurement instruments is that multimediated reality provides a direct physical alignment between the measured quantities and the real world from which they come. Multimediated reality allows sound waves or radio waves to be seen in perfect alignment with the world in which they exist, i.e. at 1:1 scale, and “frozen” in time (i.e. visible in a set of coordinates
Figure 13: With our multimediated reality headset we see the multidimensional interference pattern embedded in 3D space. Here the interference pattern between three transducers (a triangular array) is shown, and we can also interact with the waveforms by way of patterns traced by the moving print head.
in which the speed of sound or speed of light is zero, or a small number that allows the wave propagation to be studied easily).

3 MULTIMEDIATED REALITY IS MULTISCALE, MULTIMODAL, MULTISENSORY, MULTIVEILLANT, AND MULTIDIMENSIONAL

Multimediated reality is more than just a taxonomy of real and synthetic experience. It also considers how we interact with the world around us and each other, through the use of technology as a true extension of our own minds and bodies. Specifically we consider the concept of AI (Artificial Intelligence) as well as human-in-the-loop-AI, also known as HI (Humanistic Intelligence) [27]. HI posits that technology should function as an intermediary between us and our environment in such a way that the intelligence it affords us arises through a computational feedback loop of which we are a part. See Fig. 14.

Multimediated reality involves multiple physical scales, including both wearable technology as well as technology in the environment around us, like smart rooms (e.g. smart darkrooms). This multiscale and multiveillant nature of multimediated reality is illustrated in Fig. 15.

3.1 Multisensory Synthetic Synesthesia

Synesthesia is a neurological condition in which there is crosstalk between human senses, e.g. chromesthesia which is hearing colors of light, or “The Man Who Tasted Shapes” [5].

Multimediated reality often involves a multimedia-induced (synthetic) synesthesia among and across our existing senses (e.g. seeing sound), or, extrasensory, i.e. beyond our existing senses (e.g. seeing or feeling radio waves). In this way, multimediated reality is multisensory and multimodal.

3.2 Multidimensional Multimediated Reality

Synthetic synesthesia of extra senses like electromagnetic radio waves and radar (or sonar) provides us with a richly multidimensional perceptual space, in the context of multimediated reality. Whereas existing virtual reality systems might use radio waves or sound waves for tracking and position sensing, they do not directly engage the waveforms phase-coherently as part of the resulting output space. When we engage with the radio waveforms or sound waveforms directly, we have many new sensory possibilities for direct experience of multidimensional signal properties like multipolarization and multipath propagation.

Even simple radio waves become complex-valued when they are brought down to baseband (e.g. when experienced in coordinates where the speed of sound or light is zero). In this case, the chosen synesthesia is to be able to see complex numbers on a color wheel where phase is color and magnitude is the quantity (photoquantity [23]) of light presented [12].

We see this in Fig. 16, where the phase of the sound waves is displayed as color. Thus we can see clearly the interference patterns of two sound waves where they interfere constructively and destructively, as variations in the quantity of light, and the phase fronts, as variations in color.

This system may also be used to study underwater sound waves. Fig. 17 shows a sonar array being explored. We have also constructed a Multimediated Reality aquatics facility for use with underwater MR eyewear. In water, the weightless state allows us to feel as if we are floating among the underwater sound waves.

4 MULTIMEDIATED REALITY CONTINUUM

Many of the systems presented in this paper do not fit nicely into existing taxonomies of VR and AR, or any of the more general taxonomies of synthetic experience [32]. We proffer a more general “reality” continuum in which the space is multidimensional, and in which the origin is the absence of sensory stimulation, allowing us to consider technologies such as sleep masks, interactive sleep masks, sensory deprivation tanks, interactive sensory deprivation tanks [17, 21], aquatics facilities, theatres, darkrooms, therapy systems, and the like, as a basis upon which to create new realities directly connected to physical or intellectual phenomena. See Fig. 18. Note the many dimensions and the many ways they can be combined. For example we can have a mix of Reality and Virtuality that gives AR (Augmented Reality), and then further add some phenomenonality to get PAR (Phenomenological Augmented Reality [20]). We can add to AR some Fluency to get SR (Submersive Reality [17, 21]). And if we do PAR while swimming fully...
submerged in water, we’re spanning the four dimensions of Reality, Virtuality, Phenomenality, and Fluidity/Fluentity.

Note also that many of the dimensions are inherently combined and thus certainly do not form an orthonormal basis. For example, Metaveillance includes Surveillance (oversight) and AI (Artificial Intelligence), as well as Sousveillance (undersight) and HI (Humanistic Intelligence). Sousveillance and HI are very closely related to Wearability. Thus there is strong overlap between Metaveillance and Wearability. Wearability is closely correlated to Existentiality [14], giving us the “Wearable Reality” (Canadian Pat. 2388766) proposed by Mann in 1974 as embodied by the SWIM (Sequential Wave Imprinting Machine), itself an example of Phenomenality [20]. Thus these three axes, Metaveillance, Wearability, and Phenomenality, are closely related.

Not all dimensions are desirable. For example, sensing can be done in a discrete (quantized) or continuous manner. We prefer systems in which computers are used to sense “undigitally”, i.e. above the Nyquist rate spatiotemporally. Today’s digital cameras do this, for the most part, but still exhibit tonal/level quantization, thus the need for being undigital [25] with digital cameras, e.g. HDR sensing and metasensing [20, 34]. Ideally we would like to use computers for undigital senses / sensors as well as for undigital effectors / actuators, i.e. for all six signal flow paths of Fig. 14. This allows us to achieve undigital experiences, i.e. what we proffer to call an undigital reality, which is a multimediated reality that is free from the artefacts of digital computation. This is especially important for SR (Submersive Reality) in which the free flowing nature of swimming makes it possible to forget that one is wearing a VR/AR/MR headset and imagine, in the totally weightless world, another reality that is submerged in water, where things like clothes, and other technological prostheses.

Figure 15: Multimediated reality is multiveillant (surveillance AND sousveillance) as well as multiscale (wearables AND smart environments). We can identify at least three axes: Firstly, a physical scale axis (of physical reality) defines the environment (that which surrounds us) and the invironment (us ourselves). At the border between the environment and invironment are things like clothes, and other technological prostheses. A Virtuality axis defines also a scale from “Bits” all the way out to “Big Data”. A sociopolitical or “Veillance” axis defines sousveillance (individual/internal) out to surveillance (external). At the origin are “Bits”, “Atoms”, and “Genes”. Genes, for example, are the smallest unit of “humanness” (human expression). The axes are labeled $\alpha$, $\beta$, and $\gamma$. The first of the three axes ($\alpha$) is denoted pictorially, at physical scales starting from naked, to underwear, to outerwear, to a vehicle (car), to the “smart city”.

4.1 Multimediated Reality is “★R” (All R)

In most search engines and computing environments, the asterisk symbol, “★”, is a “wildcard” that can mean anything. It can be replaced with any other characters, symbols, words or phrases, and dates back as far as the TOPS-10 operating system in 1967. Thus for the space defined by the Multimediated Reality continuum, we proffer to call it “★R”, pronounced “All R” (all realities). In situations where there is only room for two characters, and where an asterisk cannot be used conveniently (e.g. file names in the Unix operating
Figure 16: Interference pattern of two acoustic wavefronts made visible using the multimediated reality darkroom of Fig. 12. As the two sound sources (speakers) are moved closer together, their acoustic interference fringes spread further apart, confirming a simple law of physics. Left photograph: wide spacing between speakers gives closely spaced interference pattern. Right photograph: narrow spacing between speakers gives broad spaced interference pattern. Sound waves from two movable speakers emitting the same pure tone are photographed with a lock-in amplifier driving an RGB (Red, Green, Blue) light source moving in space (X, Y, Z). This is not computer graphics! These are photographs of an actual physical process generated by nature itself, and is merely facilitated (made visible) by computation. Those wearing a multimediated reality headset see patterns in a high-dimensional space, whereas those without headsets can still see the sound waves, but at a reduced-dimensionality.

Figure 17: Photograph of apparatus showing underwater sound at 40 kHz from two hydrophones, toward the left (the spacing is being controlled by the user’s right hand), and a test/reference hydrophone toward the right (near the user’s left hand). The wearer of the multimediated reality eyeglass sees in a high dimensional space (3D spatially, plus time animation, plus complex-valued quantities, etc.). Others see only a flat 2D slice through the interference pattern.
system), we proffer “ZR”, to emphasize the possibility of complex-valued multidimensional data of the form $Z = X + iY$ where $i = \sqrt{-1}$ is the imaginary unit. We must realize, though, that there are many dimensions, and some of these many dimensions may be complex-valued, or beyond, i.e. not limited to the two-dimensions implied by the Argand plane (https://en.wikipedia.org/wiki/Complex_plane).

Figure 18: The Multimediated Reality Continuum. Reality is the main axis going from left-to-right, starting at “Total sensory deprivation” (the origin, indicated by a sensory deprivation tank), then to “Sensory attenuation”, then Reality, and then beyond Reality to give also Extended reality. Virtuality is the secondary axis pointing upwards. Augmentation exists in the 2-dimensional space spanned by Reality and Virtuality. A third axis, phenomenality, indicates any kind of phenomenological reality, such as phase-coherent photography of radio waves or sound waves, such as by Sequential Wave Imprinting Machine (SWIM). In this sense, PAR (Phenomenological Augmented Reality) [20] is a combination of AR and P (Phenomenality). A point in this space is indicated by the red dot as “Augmented Phenomenality”. As another example, consider a point (indicated in blue) that comes out from AR along the Fluentity axis. An example of this kind of reality is the Internet-based underwater virtual/augmented reality performance space [17, 21]. When we submerge ourselves in a large swimming pool, with an underwater SWIM to see sound waves (e.g. to test out an underwater sound system and see the interference patterns between two underwater speakers), we’re in a reality described by adding the two vectors (red dot and blue dot) taking us into an additional higher dimension. This works like a giant version of our apparatus of Fig. 17 in which we submerge ourselves fully in the pool while wearing the All Reality headset. The All Reality Continuum thus allows us to understand sensory attenuation (interpolation between sensory deprivation and reality) as well as eXtended reality (extrapolation beyond reality), in addition to the many other dimensions shown, such as Metaveillance [19] (Sur/Sous-Veillance, smart cities, smart buildings, etc.), Wearability (Mann’s “Wearable Reality” of 1974 and Canadian Patent 2388766), Kineveillance (drone and swarm-based realities [22]), Imaginality (e.g. lucid dreaming in the sensory deprivation tank), and Therapality (the axis of lasting effects that persist even after shutting off and removing a technology). Not all axes are desirable, e.g. Digitality is the axis of quantization noise that embodies the undesirable artifacts of being digital, pixelation, etc. Ideally we wish to use computers for being undigital [24].
Figure 19: Venn diagram showing various realities. VR is a proper subset of AR. AR is a proper subset of mixed reality. Mixed reality and X-Reality™ are proper subsets of mediated reality (mixed, X-, and mediated/Y- reality are denoted by a dotted line in the diagram in order to indicate that either could be represented by this circle). These realities are a proper subset of multimediated reality, which we proffer to call "XR", pronounced "All R", where "X" denotes any of the other forms of interactive multimedia reality. Alternatively, e.g. in systems like the UNIX operating system where "*" is a special reserved character, we can abbreviate "All R" as "ZR" to emphasize its multidimensional and complex-valued nature, i.e. $Z = X + iY$, where $i$ is the imaginary unit number, $i = \sqrt{-1}$.

5 SUMMARY AND CONCLUSIONS

We have presented Multimediated Reality as a proper superset of mediated (XY), mixed (X), augmented (A), and virtual (V) reality (R).

Multimediated reality uses interactive multimedia in a way that is:

- Multidimensional, in a true interactive sense, i.e. direct interaction with concepts like cross-polarization, and complex-valued multidimensional mathematical quantities, e.g. Fieldary User Interfaces [Gerson et al 2015];
- Multisensory, cross-sensory (e.g. synthestic synesthesia);
- Multimodal, in many senses of the word, e.g. multimodal interaction (human–computer interaction using multiple modalities of input/output), and multimodal artifacts (artifacts that use multiple media modes, including social media, gaming, storytelling, etc) [Gunther Kress, 2010. Multimodality: A Social Semiotic Approach to Contemporary Communication. New York: Routledge];
- Multidisciplinary (i.e. as a field of inquiry, involving the arts, the sciences, engineering, medicine, health, wellness (e.g. meditation in multimediated sensory attenuation tanks), etc.

Combining Fig. 4 and Fig. 19, we have mostly a nested set except for XR2 which is a proper subset of mixed reality, i.e. limited to a specific kind of virtual world. See Fig. 20.

Additionally, we have presented some new forms of multisensory multidimensional interactive reality, such as multidimensional complex-valued fields and other quantities that are captured and displayed to the human senses or recorded on photographic media, video display media, or wearable media. Whereas virtual reality and other realities (e.g. augmented reality) usually show artificial graphical content (computer graphics), multimmediated reality also has the capacity to show a "real reality", i.e. to make visible what is really present (but otherwise invisible) in the world around us.

ACKNOWLEDGMENTS

Authors would like to acknowledge the many students who were involved with the project: Max, Sen, Jackson, and former students Arkin and Alex, as well as Kyle. The work is supported by Mannlab Shenzhen.

REFERENCES


The Scrum Guide™

The Definitive Guide to Scrum:
The Rules of the Game

November 2017

Developed and sustained by Scrum creators: Ken Schwaber and Jeff Sutherland
# Table of Contents

Purpose of the Scrum Guide ................................................................. 3  
Definition of Scrum ........................................................................... 3  
Uses of Scrum .................................................................................... 4  
Scrum Theory ..................................................................................... 4  
Scrum Values ..................................................................................... 5  
The Scrum Team .................................................................................. 6  
  The Product Owner ........................................................................... 6  
  The Development Team ................................................................... 7  
  The Scrum Master ........................................................................... 7  
Scrum Events ...................................................................................... 9  
  The Sprint ......................................................................................... 9  
  Sprint Planning ............................................................................... 10  
  Daily Scrum ..................................................................................... 12  
  Sprint Review .................................................................................. 13  
  Sprint Retrospective ....................................................................... 14  
Scrum Artifacts .................................................................................. 14  
  Product Backlog ............................................................................. 15  
  Sprint Backlog ............................................................................... 16  
  Increment ........................................................................................ 17  
Artifact Transparency .......................................................................... 17  
  Definition of “Done” ...................................................................... 18  
End Note ............................................................................................. 19  
Acknowledgements ............................................................................ 19  
People ................................................................................................ 19  
History ............................................................................................... 19
Purpose of the Scrum Guide

Scrum is a framework for developing, delivering, and sustaining complex products. This Guide contains the definition of Scrum. This definition consists of Scrum’s roles, events, artifacts, and the rules that bind them together. Ken Schwaber and Jeff Sutherland developed Scrum; the Scrum Guide is written and provided by them. Together, they stand behind the Scrum Guide.

Definition of Scrum

Scrum (n): A framework within which people can address complex adaptive problems, while productively and creatively delivering products of the highest possible value.

Scrum is:

- Lightweight
- Simple to understand
- Difficult to master

Scrum is a process framework that has been used to manage work on complex products since the early 1990s. Scrum is not a process, technique, or definitive method. Rather, it is a framework within which you can employ various processes and techniques. Scrum makes clear the relative efficacy of your product management and work techniques so that you can continuously improve the product, the team, and the working environment.

The Scrum framework consists of Scrum Teams and their associated roles, events, artifacts, and rules. Each component within the framework serves a specific purpose and is essential to Scrum’s success and usage.

The rules of Scrum bind together the roles, events, and artifacts, governing the relationships and interaction between them. The rules of Scrum are described throughout the body of this document.

Specific tactics for using the Scrum framework vary and are described elsewhere.
**Uses of Scrum**

Scrum was initially developed for managing and developing products. Starting in the early 1990s, Scrum has been used extensively, worldwide, to:

1. Research and identify viable markets, technologies, and product capabilities;
2. Develop products and enhancements;
3. Release products and enhancements, as frequently as many times per day;
4. Develop and sustain Cloud (online, secure, on-demand) and other operational environments for product use; and,
5. Sustain and renew products.

Scrum has been used to develop software, hardware, embedded software, networks of interacting function, autonomous vehicles, schools, government, marketing, managing the operation of organizations and almost everything we use in our daily lives, as individuals and societies.

As technology, market, and environmental complexities and their interactions have rapidly increased, Scrum’s utility in dealing with complexity is proven daily.

Scrum proved especially effective in iterative and incremental knowledge transfer. Scrum is now widely used for products, services, and the management of the parent organization.

The essence of Scrum is a small team of people. The individual team is highly flexible and adaptive. These strengths continue operating in single, several, many, and networks of teams that develop, release, operate and sustain the work and work products of thousands of people. They collaborate and interoperate through sophisticated development architectures and target release environments.

When the words “develop” and “development” are used in the Scrum Guide, they refer to complex work, such as those types identified above.

**Scrum Theory**

Scrum is founded on empirical process control theory, or empiricism. Empiricism asserts that knowledge comes from experience and making decisions based on what is known. Scrum employs an iterative, incremental approach to optimize predictability and control risk.

Three pillars uphold every implementation of empirical process control: transparency, inspection, and adaptation.
Transparency

Significant aspects of the process must be visible to those responsible for the outcome. Transparency requires those aspects be defined by a common standard so observers share a common understanding of what is being seen.

For example

- A common language referring to the process must be shared by all participants; and,
- Those performing the work and those inspecting the resulting increment must share a common definition of "Done".

Inspection

Scrum users must frequently inspect Scrum artifacts and progress toward a Sprint Goal to detect undesirable variances. Their inspection should not be so frequent that inspection gets in the way of the work. Inspections are most beneficial when diligently performed by skilled inspectors at the point of work.

Adaptation

If an inspector determines that one or more aspects of a process deviate outside acceptable limits, and that the resulting product will be unacceptable, the process or the material being processed must be adjusted. An adjustment must be made as soon as possible to minimize further deviation.

Scrum prescribes four formal events for inspection and adaptation, as described in the Scrum Events section of this document:

- Sprint Planning
- Daily Scrum
- Sprint Review
- Sprint Retrospective

Scrum Values

When the values of commitment, courage, focus, openness and respect are embodied and lived by the Scrum Team, the Scrum pillars of transparency, inspection, and adaptation come to life and build trust for everyone. The Scrum Team members learn and explore those values as they work with the Scrum roles, events, and artifacts.

Successful use of Scrum depends on people becoming more proficient in living these five values. People personally commit to achieving the goals of the Scrum Team. The Scrum Team members have courage to do the right thing and work on tough problems. Everyone focuses on the work of the Sprint and the goals of the Scrum Team. The Scrum Team and its stakeholders agree to be open about all the work and the challenges with performing the work. Scrum Team members respect each other to be capable, independent people.
The Scrum Team

The Scrum Team consists of a Product Owner, the Development Team, and a Scrum Master. Scrum Teams are self-organizing and cross-functional. Self-organizing teams choose how best to accomplish their work, rather than being directed by others outside the team. Cross-functional teams have all competencies needed to accomplish the work without depending on others not part of the team. The team model in Scrum is designed to optimize flexibility, creativity, and productivity. The Scrum Team has proven itself to be increasingly effective for all the earlier stated uses, and any complex work.

Scrum Teams deliver products iteratively and incrementally, maximizing opportunities for feedback. Incremental deliveries of “Done” product ensure a potentially useful version of working product is always available.

The Product Owner

The Product Owner is responsible for maximizing the value of the product resulting from work of the Development Team. How this is done may vary widely across organizations, Scrum Teams, and individuals.

The Product Owner is the sole person responsible for managing the Product Backlog. Product Backlog management includes:

- Clearly expressing Product Backlog items;
- Ordering the items in the Product Backlog to best achieve goals and missions;
- Optimizing the value of the work the Development Team performs;
- Ensuring that the Product Backlog is visible, transparent, and clear to all, and shows what the Scrum Team will work on next; and,
- Ensuring the Development Team understands items in the Product Backlog to the level needed.

The Product Owner may do the above work, or have the Development Team do it. However, the Product Owner remains accountable.

The Product Owner is one person, not a committee. The Product Owner may represent the desires of a committee in the Product Backlog, but those wanting to change a Product Backlog item’s priority must address the Product Owner.

For the Product Owner to succeed, the entire organization must respect his or her decisions. The Product Owner’s decisions are visible in the content and ordering of the Product Backlog. No one can force the Development Team to work from a different set of requirements.
The Development Team

The Development Team consists of professionals who do the work of delivering a potentially releasable Increment of “Done” product at the end of each Sprint. A “Done” increment is required at the Sprint Review. Only members of the Development Team create the Increment.

Development Teams are structured and empowered by the organization to organize and manage their own work. The resulting synergy optimizes the Development Team’s overall efficiency and effectiveness.

Development Teams have the following characteristics:

- They are self-organizing. No one (not even the Scrum Master) tells the Development Team how to turn Product Backlog into Increments of potentially releasable functionality;
- Development Teams are cross-functional, with all the skills as a team necessary to create a product Increment;
- Scrum recognizes no titles for Development Team members, regardless of the work being performed by the person;
- Scrum recognizes no sub-teams in the Development Team, regardless of domains that need to be addressed like testing, architecture, operations, or business analysis; and,
- Individual Development Team members may have specialized skills and areas of focus, but accountability belongs to the Development Team as a whole.

Development Team Size

Optimal Development Team size is small enough to remain nimble and large enough to complete significant work within a Sprint. Fewer than three Development Team members decrease interaction and results in smaller productivity gains. Smaller Development Teams may encounter skill constraints during the Sprint, causing the Development Team to be unable to deliver a potentially releasable Increment. Having more than nine members requires too much coordination. Large Development Teams generate too much complexity for an empirical process to be useful. The Product Owner and Scrum Master roles are not included in this count unless they are also executing the work of the Sprint Backlog.

The Scrum Master

The Scrum Master is responsible for promoting and supporting Scrum as defined in the Scrum Guide. Scrum Masters do this by helping everyone understand Scrum theory, practices, rules, and values.

The Scrum Master is a servant-leader for the Scrum Team. The Scrum Master helps those outside the Scrum Team understand which of their interactions with the Scrum Team are helpful and which aren’t. The Scrum Master helps everyone change these interactions to maximize the value created by the Scrum Team.
Scrum Master Service to the Product Owner
The Scrum Master serves the Product Owner in several ways, including:

- Ensuring that goals, scope, and product domain are understood by everyone on the Scrum Team as well as possible;
- Finding techniques for effective Product Backlog management;
- Helping the Scrum Team understand the need for clear and concise Product Backlog items;
- Understanding product planning in an empirical environment;
- Ensuring the Product Owner knows how to arrange the Product Backlog to maximize value;
- Understanding and practicing agility; and,
- Facilitating Scrum events as requested or needed.

Scrum Master Service to the Development Team
The Scrum Master serves the Development Team in several ways, including:

- Coaching the Development Team in self-organization and cross-functionality;
- Helping the Development Team to create high-value products;
- Removing impediments to the Development Team’s progress;
- Facilitating Scrum events as requested or needed; and,
- Coaching the Development Team in organizational environments in which Scrum is not yet fully adopted and understood.

Scrum Master Service to the Organization
The Scrum Master serves the organization in several ways, including:

- Leading and coaching the organization in its Scrum adoption;
- Planning Scrum implementations within the organization;
- Helping employees and stakeholders understand and enact Scrum and empirical product development;
- Causing change that increases the productivity of the Scrum Team; and,
- Working with other Scrum Masters to increase the effectiveness of the application of Scrum in the organization.
Scrum Events

Prescribed events are used in Scrum to create regularity and to minimize the need for meetings not defined in Scrum. All events are time-boxed events, such that every event has a maximum duration. Once a Sprint begins, its duration is fixed and cannot be shortened or lengthened. The remaining events may end whenever the purpose of the event is achieved, ensuring an appropriate amount of time is spent without allowing waste in the process.

Other than the Sprint itself, which is a container for all other events, each event in Scrum is a formal opportunity to inspect and adapt something. These events are specifically designed to enable critical transparency and inspection. Failure to include any of these events results in reduced transparency and is a lost opportunity to inspect and adapt.

The Sprint

The heart of Scrum is a Sprint, a time-box of one month or less during which a “Done”, useable, and potentially releasable product Increment is created. Sprints have consistent durations throughout a development effort. A new Sprint starts immediately after the conclusion of the previous Sprint.

Sprints contain and consist of the Sprint Planning, Daily Scrums, the development work, the Sprint Review, and the Sprint Retrospective.

During the Sprint:

- No changes are made that would endanger the Sprint Goal;
- Quality goals do not decrease; and,
- Scope may be clarified and re-negotiated between the Product Owner and Development Team as more is learned.

Each Sprint may be considered a project with no more than a one-month horizon. Like projects, Sprints are used to accomplish something. Each Sprint has a goal of what is to be built, a design and flexible plan that will guide building it, the work, and the resultant product increment.

Sprints are limited to one calendar month. When a Sprint’s horizon is too long the definition of what is being built may change, complexity may rise, and risk may increase. Sprints enable predictability by ensuring inspection and adaptation of progress toward a Sprint Goal at least every calendar month. Sprints also limit risk to one calendar month of cost.
Cancelling a Sprint
A Sprint can be cancelled before the Sprint time-box is over. Only the Product Owner has the authority to cancel the Sprint, although he or she may do so under influence from the stakeholders, the Development Team, or the Scrum Master.

A Sprint would be cancelled if the Sprint Goal becomes obsolete. This might occur if the company changes direction or if market or technology conditions change. In general, a Sprint should be cancelled if it no longer makes sense given the circumstances. But, due to the short duration of Sprints, cancellation rarely makes sense.

When a Sprint is cancelled, any completed and “Done” Product Backlog items are reviewed. If part of the work is potentially releasable, the Product Owner typically accepts it. All incomplete Product Backlog Items are re-estimated and put back on the Product Backlog. The work done on them depreciates quickly and must be frequently re-estimated.

Sprint cancellations consume resources, since everyone regroups in another Sprint Planning to start another Sprint. Sprint cancellations are often traumatic to the Scrum Team, and are very uncommon.

Sprint Planning
The work to be performed in the Sprint is planned at the Sprint Planning. This plan is created by the collaborative work of the entire Scrum Team.

Sprint Planning is time-boxed to a maximum of eight hours for a one-month Sprint. For shorter Sprints, the event is usually shorter. The Scrum Master ensures that the event takes place and that attendants understand its purpose. The Scrum Master teaches the Scrum Team to keep it within the time-box.

Sprint Planning answers the following:

- What can be delivered in the Increment resulting from the upcoming Sprint?
- How will the work needed to deliver the Increment be achieved?

Topic One: What can be done this Sprint?
The Development Team works to forecast the functionality that will be developed during the Sprint. The Product Owner discusses the objective that the Sprint should achieve and the Product Backlog items that, if completed in the Sprint, would achieve the Sprint Goal. The entire Scrum Team collaborates on understanding the work of the Sprint.

The input to this meeting is the Product Backlog, the latest product Increment, projected capacity of the Development Team during the Sprint, and past performance of the Development Team. The number of items selected from the Product Backlog for the Sprint is solely up to the Development Team. Only the Development Team can assess what it can accomplish over the upcoming Sprint.
During Sprint Planning the Scrum Team also crafts a Sprint Goal. The Sprint Goal is an objective that will be met within the Sprint through the implementation of the Product Backlog, and it provides guidance to the Development Team on why it is building the Increment.

**Topic Two: How will the chosen work get done?**

Having set the Sprint Goal and selected the Product Backlog items for the Sprint, the Development Team decides how it will build this functionality into a “Done” product Increment during the Sprint. The Product Backlog items selected for this Sprint plus the plan for delivering them is called the Sprint Backlog.

The Development Team usually starts by designing the system and the work needed to convert the Product Backlog into a working product Increment. Work may be of varying size, or estimated effort. However, enough work is planned during Sprint Planning for the Development Team to forecast what it believes it can do in the upcoming Sprint. Work planned for the first days of the Sprint by the Development Team is decomposed by the end of this meeting, often to units of one day or less. The Development Team self-organizes to undertake the work in the Sprint Backlog, both during Sprint Planning and as needed throughout the Sprint.

The Product Owner can help to clarify the selected Product Backlog items and make trade-offs. If the Development Team determines it has too much or too little work, it may renegotiate the selected Product Backlog items with the Product Owner. The Development Team may also invite other people to attend to provide technical or domain advice.

By the end of the Sprint Planning, the Development Team should be able to explain to the Product Owner and Scrum Master how it intends to work as a self-organizing team to accomplish the Sprint Goal and create the anticipated Increment.

**Sprint Goal**

The Sprint Goal is an objective set for the Sprint that can be met through the implementation of Product Backlog. It provides guidance to the Development Team on why it is building the Increment. It is created during the Sprint Planning meeting. The Sprint Goal gives the Development Team some flexibility regarding the functionality implemented within the Sprint. The selected Product Backlog items deliver one coherent function, which can be the Sprint Goal. The Sprint Goal can be any other coherence that causes the Development Team to work together rather than on separate initiatives.

As the Development Team works, it keeps the Sprint Goal in mind. In order to satisfy the Sprint Goal, it implements functionality and technology. If the work turns out to be different than the Development Team expected, they collaborate with the Product Owner to negotiate the scope of Sprint Backlog within the Sprint.
**Daily Scrum**

The Daily Scrum is a 15-minute time-boxed event for the Development Team. The Daily Scrum is held every day of the Sprint. At it, the Development Team plans work for the next 24 hours. This optimizes team collaboration and performance by inspecting the work since the last Daily Scrum and forecasting upcoming Sprint work. The Daily Scrum is held at the same time and place each day to reduce complexity.

The Development Team uses the Daily Scrum to inspect progress toward the Sprint Goal and to inspect how progress is trending toward completing the work in the Sprint Backlog. The Daily Scrum optimizes the probability that the Development Team will meet the Sprint Goal. Every day, the Development Team should understand how it intends to work together as a self-organizing team to accomplish the Sprint Goal and create the anticipated Increment by the end of the Sprint.

The structure of the meeting is set by the Development Team and can be conducted in different ways if it focuses on progress toward the Sprint Goal. Some Development Teams will use questions, some will be more discussion based. Here is an example of what might be used:

- What did I do yesterday that helped the Development Team meet the Sprint Goal?
- What will I do today to help the Development Team meet the Sprint Goal?
- Do I see any impediment that prevents me or the Development Team from meeting the Sprint Goal?

The Development Team or team members often meet immediately after the Daily Scrum for detailed discussions, or to adapt, or replan, the rest of the Sprint’s work.

The Scrum Master ensures that the Development Team has the meeting, but the Development Team is responsible for conducting the Daily Scrum. The Scrum Master teaches the Development Team to keep the Daily Scrum within the 15-minute time-box.

The Daily Scrum is an internal meeting for the Development Team. If others are present, the Scrum Master ensures that they do not disrupt the meeting.

Daily Scrums improve communications, eliminate other meetings, identify impediments to development for removal, highlight and promote quick decision-making, and improve the Development Team’s level of knowledge. This is a key inspect and adapt meeting.
Sprint Review

A Sprint Review is held at the end of the Sprint to inspect the Increment and adapt the Product Backlog if needed. During the Sprint Review, the Scrum Team and stakeholders collaborate about what was done in the Sprint. Based on that and any changes to the Product Backlog during the Sprint, attendees collaborate on the next things that could be done to optimize value. This is an informal meeting, not a status meeting, and the presentation of the Increment is intended to elicit feedback and foster collaboration.

This is at most a four-hour meeting for one-month Sprints. For shorter Sprints, the event is usually shorter. The Scrum Master ensures that the event takes place and that attendees understand its purpose. The Scrum Master teaches everyone involved to keep it within the time-box.

The Sprint Review includes the following elements:

- Attendees include the Scrum Team and key stakeholders invited by the Product Owner;
- The Product Owner explains what Product Backlog items have been “Done” and what has not been “Done”;
- The Development Team discusses what went well during the Sprint, what problems it ran into, and how those problems were solved;
- The Development Team demonstrates the work that it has “Done” and answers questions about the Increment;
- The Product Owner discusses the Product Backlog as it stands. He or she projects likely target and delivery dates based on progress to date (if needed);
- The entire group collaborates on what to do next, so that the Sprint Review provides valuable input to subsequent Sprint Planning;
- Review of how the marketplace or potential use of the product might have changed what is the most valuable thing to do next; and,
- Review of the timeline, budget, potential capabilities, and marketplace for the next anticipated releases of functionality or capability of the product.

The result of the Sprint Review is a revised Product Backlog that defines the probable Product Backlog items for the next Sprint. The Product Backlog may also be adjusted overall to meet new opportunities.
**Sprint Retrospective**

The Sprint Retrospective is an opportunity for the Scrum Team to inspect itself and create a plan for improvements to be enacted during the next Sprint.

The Sprint Retrospective occurs after the Sprint Review and prior to the next Sprint Planning. This is at most a three-hour meeting for one-month Sprints. For shorter Sprints, the event is usually shorter. The Scrum Master ensures that the event takes place and that attendants understand its purpose.

The Scrum Master ensures that the meeting is positive and productive. The Scrum Master teaches all to keep it within the time-box. The Scrum Master participates as a peer team member in the meeting from the accountability over the Scrum process.

The purpose of the Sprint Retrospective is to:

- Inspect how the last Sprint went with regards to people, relationships, process, and tools;
- Identify and order the major items that went well and potential improvements; and,
- Create a plan for implementing improvements to the way the Scrum Team does its work.

The Scrum Master encourages the Scrum Team to improve, within the Scrum process framework, its development process and practices to make it more effective and enjoyable for the next Sprint. During each Sprint Retrospective, the Scrum Team plans ways to increase product quality by improving work processes or adapting the definition of “Done”, if appropriate and not in conflict with product or organizational standards.

By the end of the Sprint Retrospective, the Scrum Team should have identified improvements that it will implement in the next Sprint. Implementing these improvements in the next Sprint is the adaptation to the inspection of the Scrum Team itself. Although improvements may be implemented at any time, the Sprint Retrospective provides a formal opportunity to focus on inspection and adaptation.

**Scrum Artifacts**

Scrum’s artifacts represent work or value to provide transparency and opportunities for inspection and adaptation. Artifacts defined by Scrum are specifically designed to maximize transparency of key information so that everybody has the same understanding of the artifact.
**Product Backlog**

The Product Backlog is an ordered list of everything that is known to be needed in the product. It is the single source of requirements for any changes to be made to the product. The Product Owner is responsible for the Product Backlog, including its content, availability, and ordering.

A Product Backlog is never complete. The earliest development of it lays out the initially known and best-understood requirements. The Product Backlog evolves as the product and the environment in which it will be used evolves. The Product Backlog is dynamic; it constantly changes to identify what the product needs to be appropriate, competitive, and useful. If a product exists, its Product Backlog also exists.

The Product Backlog lists all features, functions, requirements, enhancements, and fixes that constitute the changes to be made to the product in future releases. Product Backlog items have the attributes of a description, order, estimate, and value. Product Backlog items often include test descriptions that will prove its completeness when “Done.”

As a product is used and gains value, and the marketplace provides feedback, the Product Backlog becomes a larger and more exhaustive list. Requirements never stop changing, so a Product Backlog is a living artifact. Changes in business requirements, market conditions, or technology may cause changes in the Product Backlog.

Multiple Scrum Teams often work together on the same product. One Product Backlog is used to describe the upcoming work on the product. A Product Backlog attribute that groups items may then be employed.

Product Backlog refinement is the act of adding detail, estimates, and order to items in the Product Backlog. This is an ongoing process in which the Product Owner and the Development Team collaborate on the details of Product Backlog items. During Product Backlog refinement, items are reviewed and revised. The Scrum Team decides how and when refinement is done. Refinement usually consumes no more than 10% of the capacity of the Development Team. However, Product Backlog items can be updated at any time by the Product Owner or at the Product Owner’s discretion.

Higher ordered Product Backlog items are usually clearer and more detailed than lower ordered ones. More precise estimates are made based on the greater clarity and increased detail; the lower the order, the less detail. Product Backlog items that will occupy the Development Team for the upcoming Sprint are refined so that any one item can reasonably be “Done” within the Sprint time-box. Product Backlog items that can be “Done” by the Development Team within one Sprint are deemed “Ready” for selection in a Sprint Planning. Product Backlog items usually acquire this degree of transparency through the above described refining activities.

The Development Team is responsible for all estimates. The Product Owner may influence the Development Team by helping it understand and select trade-offs, but the people who will perform the work make the final estimate.
Monitoring Progress Toward Goals
At any point in time, the total work remaining to reach a goal can be summed. The Product Owner tracks this total work remaining at least every Sprint Review. The Product Owner compares this amount with work remaining at previous Sprint Reviews to assess progress toward completing projected work by the desired time for the goal. This information is made transparent to all stakeholders.

Various projective practices upon trending have been used to forecast progress, like burn-downs, burn-ups, or cumulative flows. These have proven useful. However, these do not replace the importance of empiricism. In complex environments, what will happen is unknown. Only what has already happened may be used for forward-looking decision-making.

Sprint Backlog
The Sprint Backlog is the set of Product Backlog items selected for the Sprint, plus a plan for delivering the product Increment and realizing the Sprint Goal. The Sprint Backlog is a forecast by the Development Team about what functionality will be in the next Increment and the work needed to deliver that functionality into a “Done” Increment.

The Sprint Backlog makes visible all the work that the Development Team identifies as necessary to meet the Sprint Goal. To ensure continuous improvement, it includes at least one high priority process improvement identified in the previous Retrospective meeting.

The Sprint Backlog is a plan with enough detail that changes in progress can be understood in the Daily Scrum. The Development Team modifies the Sprint Backlog throughout the Sprint, and the Sprint Backlog emerges during the Sprint. This emergence occurs as the Development Team works through the plan and learns more about the work needed to achieve the Sprint Goal.

As new work is required, the Development Team adds it to the Sprint Backlog. As work is performed or completed, the estimated remaining work is updated. When elements of the plan are deemed unnecessary, they are removed. Only the Development Team can change its Sprint Backlog during a Sprint. The Sprint Backlog is a highly visible, real-time picture of the work that the Development Team plans to accomplish during the Sprint, and it belongs solely to the Development Team.
**Monitoring Sprint Progress**

At any point in time in a Sprint, the total work remaining in the Sprint Backlog can be summed. The Development Team tracks this total work remaining at least for every Daily Scrum to project the likelihood of achieving the Sprint Goal. By tracking the remaining work throughout the Sprint, the Development Team can manage its progress.

**Increment**

The Increment is the sum of all the Product Backlog items completed during a Sprint and the value of the increments of all previous Sprints. At the end of a Sprint, the new Increment must be “Done,” which means it must be in useable condition and meet the Scrum Team’s definition of “Done.” An increment is a body of inspectable, done work that supports empiricism at the end of the Sprint. The increment is a step toward a vision or goal. The increment must be in useable condition regardless of whether the Product Owner decides to release it.

**Artifact Transparency**

Scrum relies on transparency. Decisions to optimize value and control risk are made based on the perceived state of the artifacts. To the extent that transparency is complete, these decisions have a sound basis. To the extent that the artifacts are incompletely transparent, these decisions can be flawed, value may diminish and risk may increase.

The Scrum Master must work with the Product Owner, Development Team, and other involved parties to understand if the artifacts are completely transparent. There are practices for coping with incomplete transparency; the Scrum Master must help everyone apply the most appropriate practices in the absence of complete transparency. A Scrum Master can detect incomplete transparency by inspecting the artifacts, sensing patterns, listening closely to what is being said, and detecting differences between expected and real results.

The Scrum Master’s job is to work with the Scrum Team and the organization to increase the transparency of the artifacts. This work usually involves learning, convincing, and change. Transparency doesn’t occur overnight, but is a path.
Definition of “Done”

When a Product Backlog item or an Increment is described as “Done”, everyone must understand what “Done” means. Although this may vary significantly per Scrum Team, members must have a shared understanding of what it means for work to be complete, to ensure transparency. This is the definition of “Done” for the Scrum Team and is used to assess when work is complete on the product Increment.

The same definition guides the Development Team in knowing how many Product Backlog items it can select during a Sprint Planning. The purpose of each Sprint is to deliver Increments of potentially releasable functionality that adhere to the Scrum Team’s current definition of “Done.”

Development Teams deliver an Increment of product functionality every Sprint. This Increment is useable, so a Product Owner may choose to immediately release it. If the definition of "Done" for an increment is not a convention of the development organization, all Scrum Teams must follow it as a minimum.

If "Done" for an increment is not a convention of the development organization, the Development Team of the Scrum Team must define a definition of “Done” appropriate for the product. If there are multiple Scrum Teams working on the system or product release, the Development Teams on all the Scrum Teams must mutually define the definition of “Done.”

Each Increment is additive to all prior Increments and thoroughly tested, ensuring that all Increments work together.

As Scrum Teams mature, it is expected that their definitions of “Done” will expand to include more stringent criteria for higher quality. New definitions, as used, may uncover work to be done in previously “Done” increments. Any one product or system should have a definition of “Done” that is a standard for any work done on it.
End Note
Scrum is free and offered in this Guide. Scrum’s roles, events, artifacts, and rules are immutable and although implementing only parts of Scrum is possible, the result is not Scrum. Scrum exists only in its entirety and functions well as a container for other techniques, methodologies, and practices.

Acknowledgements

People
Of the thousands of people who have contributed to Scrum, we should single out those who were instrumental at the start: Jeff Sutherland worked with Jeff McKenna and John Scumniotailes, and Ken Schwaber worked with Mike Smith and Chris Martin, and all of them worked together. Many others contributed in the ensuing years and without their help Scrum would not be refined as it is today.

History
Ken Schwaber and Jeff Sutherland worked on Scrum until 1995, when they co-presented Scrum at the OOPSLA Conference in 1995. This presentation essentially documented the learning that Ken and Jeff gained over the previous few years, and made public the first formal definition of Scrum.

The history of Scrum is described elsewhere. To honor the first places where it was tried and refined, we recognize Individual, Inc., Newspage, Fidelity Investments, and IDX (now GE Medical).

The Scrum Guide documents Scrum as developed, evolved, and sustained for 20-plus years by Jeff Sutherland and Ken Schwaber. Other sources provide you with patterns, processes, and insights that complement the Scrum framework. These may increase productivity, value, creativity, and satisfaction with the results.
Git is the open source distributed version control system that facilitates GitHub activities on your laptop or desktop. This cheat sheet summarizes commonly used Git command line instructions for quick reference.

**INSTALL GIT**

GitHub provides desktop clients that include a graphical user interface for the most common repository actions and an automatically updating command line edition of Git for advanced scenarios.

- **GitHub for Windows**
  https://windows.github.com

- **GitHub for Mac**
  https://mac.github.com

Git distributions for Linux and POSIX systems are available on the official Git SCM web site.

- **Git for All Platforms**
  http://git-scm.com

**MAKE CHANGES**

Review edits and craft a commit transaction

- **$ git status**
  Lists all new or modified files to be committed

- **$ git diff**
  Shows file differences not yet staged

- **$ git add [file]**
  Snapshots the file in preparation for versioning

- **$ git diff --staged**
  Shows file differences between staging and the last file version

- **$ git reset [file]**
  Unstages the file, but preserve its contents

- **$ git commit -m "[descriptive message]"**
  Records file snapshots permanently in version history

**CONFIGURE TOOLING**

Configure user information for all local repositories

- **$ git config --global user.name "[name]"**
  Sets the name you want attached to your commit transactions

- **$ git config --global user.email "[email address]"**
  Sets the email you want attached to your commit transactions

- **$ git config --global color.ui auto**
  Enables helpful colorization of command line output

**CREATE REPOSITORIES**

Start a new repository or obtain one from an existing URL

- **$ git init [project-name]**
  Creates a new local repository with the specified name

- **$ git clone [url]**
  Downloads a project and its entire version history

**GROUP CHANGES**

Name a series of commits and combine completed efforts

- **$ git branch**
  Lists all local branches in the current repository

- **$ git branch [branch-name]**
  Creates a new branch

- **$ git checkout [branch-name]**
  Switches to the specified branch and updates the working directory

- **$ git merge [branch]**
  Combines the specified branch’s history into the current branch

- **$ git branch -d [branch-name]**
  Deletes the specified branch
## REFACCTOR FILENAMES
Relocate and remove versioned files

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>$ git rm [file]</code></td>
<td>Deletes the file from the working directory and stages the deletion</td>
</tr>
<tr>
<td><code>$ git rm --cached [file]</code></td>
<td>Removes the file from version control but preserves the file locally</td>
</tr>
<tr>
<td><code>$ git mv [file-original] [file-renamed]</code></td>
<td>Changes the file name and prepares it for commit</td>
</tr>
</tbody>
</table>

## SUPPRESS TRACKING
Exclude temporary files and paths

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>$ git ls-files --other --ignored --exclude-standard</code></td>
<td>Lists all ignored files in this project</td>
</tr>
</tbody>
</table>

## REVIEW HISTORY
Browse and inspect the evolution of project files

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>$ git log</code></td>
<td>Lists version history for the current branch</td>
</tr>
<tr>
<td><code>$ git log --follow [file]</code></td>
<td>Lists version history for a file, including renames</td>
</tr>
<tr>
<td><code>$ git diff [first-branch]...[second-branch]</code></td>
<td>Shows content differences between two branches</td>
</tr>
<tr>
<td><code>$ git show [commit]</code></td>
<td>Outputs metadata and content changes of the specified commit</td>
</tr>
</tbody>
</table>

## REDO COMMITS
Erase mistakes and craft replacement history

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>$ git reset [commit]</code></td>
<td>Undoes all commits after [commit], preserving changes locally</td>
</tr>
<tr>
<td><code>$ git reset --hard [commit]</code></td>
<td>Discards all history and changes back to the specified commit</td>
</tr>
</tbody>
</table>

## SAVE FRAGMENTS
Shelve and restore incomplete changes

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>$ git stash</code></td>
<td>Temporarily stores all modified tracked files</td>
</tr>
<tr>
<td><code>$ git stash pop</code></td>
<td>Restores the most recently stashed files</td>
</tr>
<tr>
<td><code>$ git stash list</code></td>
<td>Lists all stashed changesets</td>
</tr>
<tr>
<td><code>$ git stash drop</code></td>
<td>Discards the most recently stashed changeset</td>
</tr>
</tbody>
</table>

## SYNCHRONIZE CHANGES
Register a repository bookmark and exchange version history

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>$ git fetch [bookmark]</code></td>
<td>Downloads all history from the repository bookmark</td>
</tr>
<tr>
<td><code>$ git merge [bookmark]/[branch]</code></td>
<td>Combines bookmark's branch into current local branch</td>
</tr>
<tr>
<td><code>$ git push [alias] [branch]</code></td>
<td>Uploads all local branch commits to GitHub</td>
</tr>
<tr>
<td><code>$ git pull</code></td>
<td>Downloads bookmark history and incorporates changes</td>
</tr>
</tbody>
</table>
THE LONDON CHARTER

FOR THE COMPUTER-BASED VISUALISATION OF CULTURAL HERITAGE

Preamble

Objectives

Principles

Principle 1: Implementation
Principle 2: Aims and Methods
Principle 3: Research Sources
Principle 4: Documentation
Principle 5: Sustainability
Principle 6: Access

Glossary
PREAMBLE

While computer-based visualisation methods are now employed in a wide range of contexts to assist in the research, communication and preservation of cultural heritage, a set of principles is needed that will ensure that digital heritage visualisation is, and is seen to be, at least as intellectually and technically rigorous as longer established cultural heritage research and communication methods. At the same time, such principles must reflect the distinctive properties of computer-based visualisation technologies and methods.

Numerous articles, documents, including the AHDS Guides to Good Practice for CAD (2002) and Virtual Reality (2002) and initiatives, including the Virtual Archaeology Special Interest Group (VASIG) and the Cultural Virtual Reality Organisation (CVRO) and others have underlined the importance of ensuring both that computer-based visualisation methods are applied with scholarly rigour, and that the outcomes of research that include computer-based visualisation should accurately convey to users the status of the knowledge that they represent, such as distinctions between evidence and hypothesis, and between different levels of probability.

The London Charter seeks to capture, and to build, a consensus on these and related issues in a way that demands wide recognition and an expectation of compliance within relevant subject communities. In doing so, the Charter aims to enhance the rigour with which computer-based visualisation methods and outcomes are used and evaluated in heritage contexts, thereby promoting understanding and recognition of such methods and outcomes.

The Charter defines principles for the use of computer-based visualisation methods in relation to intellectual integrity, reliability, documentation, sustainability and access.
The Charter recognises that the range of available computer-based visualisation methods is constantly increasing, and that these methods can be applied to address an equally expanding range of research aims. The Charter therefore does not seek to prescribe specific aims or methods, but rather establishes those broad principles for the use, in research and communication of cultural heritage, of computer-based visualisation upon which the intellectual integrity of such methods and outcomes depend.

The Charter is concerned with the research and dissemination of cultural heritage across academic, educational, curatorial and commercial domains. It has relevance, therefore, for those aspects of the entertainment industry involving the reconstruction or evocation of cultural heritage, but not for the use of computer-based visualisation in, for example, contemporary art, fashion, or design. As the aims that motivate the use of visualisation methods vary widely from domain to domain, Principle 1: “Implementation”, signals the importance of devising detailed guidelines appropriate to each community of practice.
OBJECTIVES

The London Charter seeks to establish principles for the use of computer-based visualisation methods and outcomes in the research and communication of cultural heritage in order to:

Provide a benchmark having widespread recognition among stakeholders.

Promote intellectual and technical rigour in digital heritage visualisation.

Ensure that computer-based visualisation processes and outcomes can be properly understood and evaluated by users.

Enable computer-based visualisation authoritatively to contribute to the study, interpretation and management of cultural heritage assets.

Ensure access and sustainability strategies are determined and applied.

Offer a robust foundation upon which communities of practice can build detailed London Charter Implementation Guidelines.
PRINCIPLES

Principle 1: Implementation

The principles of the London Charter are valid wherever computer-based visualisation is applied to the research or dissemination of cultural heritage.

1.1 Each community of practice, whether academic, educational, curatorial or commercial, should develop London Charter Implementation Guidelines that cohere with its own aims, objectives and methods.

1.2 Every computer-based visualisation heritage activity should develop, and monitor the application of, a London Charter Implementation Strategy.

1.3 In collaborative activities, all participants whose role involves either directly or indirectly contributing to the visualisation process should be made aware of the principles of the London Charter, together with relevant Charter Implementation Guidelines, and to assess their implications for the planning, documentation and dissemination of the project as a whole.

1.4 The costs of implementing such a strategy should be considered in relation to the added intellectual, explanatory and/or economic value of producing outputs that demonstrate a high level of intellectual integrity.
Principle 2: Aims and Methods

A computer-based visualisation method should normally be used only when it is the most appropriate available method for that purpose.

2.1 It should not be assumed that computer-based visualisation is the most appropriate means of addressing all cultural heritage research or communication aims.

2.2 A systematic, documented evaluation of the suitability of each method to each aim should be carried out, in order to ascertain what, if any, type of computer-based visualisation is likely to prove most appropriate.

2.3 While it is recognised that, particularly in innovative or complex activities, it may not always be possible to determine, a priori, the most appropriate method, the choice of computer-based visualisation method (e.g. more or less photo-realistic, impressionistic or schematic; representation of hypotheses or of the available evidence; dynamic or static) or the decision to develop a new method, should be based on an evaluation of the likely success of each approach in addressing each aim.
Principle 3: Research Sources

In order to ensure the intellectual integrity of computer-based visualisation methods and outcomes, relevant research sources should be identified and evaluated in a structured and documented way.

3.1. In the context of the Charter, research sources are defined as all information, digital and non-digital, considered during, or directly influencing, the creation of computer-based visualisation outcomes.

3.2 Research sources should be selected, analysed and evaluated with reference to current understandings and best practice within communities of practice.

3.3 Particular attention should be given to the way in which visual sources may be affected by ideological, historical, social, religious and aesthetic and other such factors.
Principle 4: Documentation

Sufficient information should be documented and disseminated to allow computer-based visualisation methods and outcomes to be understood and evaluated in relation to the contexts and purposes for which they are deployed.

Enhancing Practice
4.1 Documentation strategies should be designed and resourced in such a way that they actively enhance the visualisation activity by encouraging, and helping to structure, thoughtful practice.

4.2 Documentation strategies should be designed to enable rigorous, comparative analysis and evaluation of computer-based visualisations, and to facilitate the recognition and addressing of issues that visualisation activities reveal.

4.3 Documentation strategies may assist in the management of Intellectual Property Rights or privileged information.

Documentation of Knowledge Claims
4.4 It should be made clear to users what a computer-based visualisation seeks to represent, for example the existing state, an evidence-based restoration or an hypothetical reconstruction of a cultural heritage object or site, and the extent and nature of any factual uncertainty.

Documentation of Research Sources
4.5 A complete list of research sources used and their provenance should be disseminated.

Documentation of Process (Paradata)
4.6 Documentation of the evaluative, analytical, deductive, interpretative and creative decisions made in the course of computer-based visualisation should be disseminated in such a way that the relationship between
research sources, implicit knowledge, explicit reasoning, and visualisation-based outcomes can be understood.

Documentation of Methods

4.7 The rationale for choosing a computer-based visualisation method, and for rejecting other methods, should be documented and disseminated to allow the activity's methodology to be evaluated and to inform subsequent activities.

4.8 A description of the visualisation methods should be disseminated if these are not likely to be widely understood within relevant communities of practice.

4.9 Where computer-based visualisation methods are used in interdisciplinary contexts that lack a common set of understandings about the nature of research questions, methods and outcomes, project documentation should be undertaken in such a way that it assists in articulating such implicit knowledge and in identifying the different lexica of participating members from diverse subject communities.

Documentation of Dependency Relationships

4.10 Computer-based visualisation outcomes should be disseminated in such a way that the nature and importance of significant, hypothetical dependency relationships between elements can be clearly identified by users and the reasoning underlying such hypotheses understood.

Documentation Formats and Standards

4.11 Documentation should be disseminated using the most effective available media, including graphical, textual, video, audio, numerical or combinations of the above.

4.12 Documentation should be disseminated sustainably with reference to relevant standards and ontologies according to best practice in relevant communities of practice and in such a way that facilitates its inclusion in relevant citation indexes.
Principle 5: Sustainability

Strategies should be planned and implemented to ensure the long-term sustainability of cultural heritage-related computer-based visualisation outcomes and documentation, in order to avoid loss of this growing part of human intellectual, social, economic and cultural heritage.

5.1 The most reliable and sustainable available form of archiving computer-based visualisation outcomes, whether analogue or digital, should be identified and implemented.

5.2 Digital preservation strategies should aim to preserve the computer-based visualisation data, rather than the medium on which they were originally stored, and also information sufficient to enable their use in the future, for example through migration to different formats or software emulation.

5.3 Where digital archiving is not the most reliable means of ensuring the long-term survival of a computer-based visualisation outcome, a partial, two-dimensional record of a computer-based visualisation output, evoking as far as possible the scope and properties of the original output, should be preferred to the absence of a record.

5.4 Documentation strategies should be designed to be sustainable in relation to available resources and prevailing working practices.
Principle 6: Access

The creation and dissemination of computer-based visualisation should be planned in such a way as to ensure that maximum possible benefits are achieved for the study, understanding, interpretation, preservation and management of cultural heritage.

6.1 The aims, methods and dissemination plans of computer-based visualisation should reflect consideration of how such work can enhance access to cultural heritage that is otherwise inaccessible due to health and safety, disability, economic, political, or environmental reasons, or because the object of the visualisation is lost, endangered, dispersed, or has been destroyed, restored or reconstructed.

6.2 Projects should take cognizance of the types and degrees of access that computer-based visualisation can uniquely provide to cultural heritage stakeholders, including the study of change over time, magnification, modification, manipulation of virtual objects, embedding of datasets, instantaneous global distribution.
APPENDIX – Glossary

The following definitions explain how terms are used within this document. They are not intended to be prescriptive beyond that function.

*Computer-based visualisation*

The process of representing information visually with the aid of computer technologies.

*Computer-based visualisation method*

The systematic application, usually in a research context, of computer-based visualisation in order to address identified aims.

*Computer-based visualisation outcome*

An outcome of computer-based visualisation, including but not limited to digital models, still images, animations and physical models.

*Cultural heritage*

The Charter adopts a wide definition of this term, encompassing all domains of human activity which are concerned with the understanding of communication of the material and intellectual culture. Such domains include, but are not limited to, museums, art galleries, heritage sites, interpretative centres, cultural heritage research institutes, arts and humanities subjects within higher education institutions, the broader educational sector, and tourism.

*Dependency relationship*

A dependent relationship between the properties of elements within digital models, such that a change in one property will necessitate change in the dependent properties. (For instance, a change in the height of a door will necessitate a corresponding change in the height of the doorframe.)

*Intellectual transparency*

The provision of information, presented in any medium or format, to allow users to understand the nature and scope of “knowledge claim” made by a computer-based visualisation outcome.
Paradata

Information about human processes of understanding and interpretation of data objects. Examples of paradata include descriptions stored within a structured dataset of how evidence was used to interpret an artefact, or a comment on methodological premises within a research publication. It is closely related, but somewhat different in emphasis, to “contextual metadata”, which tend to communicate interpretations of an artefact or collection, rather than the process through which one or more artefacts were processed or interpreted.

Research sources

All information, digital and non-digital, considered during, or directly influencing, the creation of the computer-based visualisation outcomes.

Subject community

A group of researchers generally defined by a discipline (e.g. Archaeology, Classics, Sinology, Egyptology) and sharing a broadly-defined understanding of what constitute valid research questions, methods and outputs within their subject area.

Sustainability strategy

A strategy to ensure that some meaningful record of computer-based visualisation processes and outcomes is preserved for future generations.

Editor: Hugh Denard, King’s College London, 7 February 2009
A multimedia approach to the diffusion, communication, and exploitation of Cultural Heritage (CH) is a well-established trend worldwide. Several studies demonstrate that the use of new and combined media enhances how culture is experienced. The benefit is in terms of both number of people who can have access to knowledge and the quality of the diffusion of the knowledge itself. In this regard, CH uses augmented-, virtual-, and mixed-reality technologies for different purposes, including education, exhibition enhancement, exploration, reconstruction, and virtual museums. These technologies enable user-centred presentation and make cultural heritage digitally accessible, especially when physical access is constrained. A number of surveys of these emerging technologies have been conducted; however, they are either not domain specific or lack a holistic perspective in that they do not cover all the aspects of the technology. A review of these technologies from a cultural heritage perspective is therefore warranted. Accordingly, our article surveys the state-of-the-art in augmented-, virtual-, and mixed-reality systems as a whole and from a cultural heritage perspective. In addition, we identify specific application areas in digital cultural heritage and make suggestions as to which technology is most appropriate in each case. Finally, the article predicts future research directions for augmented and virtual reality, with a particular focus on interaction interfaces and explores the implications for the cultural heritage domain.

CCS Concepts: • Computing methodologies → Mixed/augmented reality; Virtual reality;

Additional Key Words and Phrases: Cultural heritage, augmented reality, virtual reality, mixed reality

ACM Reference format:
https://doi.org/10.1145/3145534
1 INTRODUCTION

Cultural Computing (CC) is an emerging field that applies computer technology and scientific methods to culture, arts, and the social sciences to represent, enhance, extend, and transform creative products and processes (Haydar et al. 2011; Wang 2009). Advancements in computer technology have made the acquisition, recording, and manipulation of three-dimensional (3D) data technically achievable (Portalés et al. 2009) with techniques such as reverse engineering and computer graphics being used for analysing, studying, preserving, and visualising Cultural Heritage (CH) assets (Barsanti et al. 2015). Since the mid-2000s, the use of enabling technologies in Cultural Heritage has been extended to immersive technologies—a collective term for augmented-, virtual-, and mixed-reality technologies, which provide sensory experiences through various combinations of real and digital content.

Cultural Heritage, as a domain, benefits significantly from the use of these technologies. Users are able to experience cultural artefacts in a completely new way. While there are a number of general surveys of immersive-reality technologies (Adhani and Awang 2012; Anthes et al. 2016; Arth et al. 2015; Azuma et al. 2001; Azuma 1997; Carmigniani et al. 2011; Costanza et al. 2009; Papagiannakis et al. 2008; Sanna and Manuri 2016; Van Krevelen and Poelman 2010; Zhou et al. 2008; Zhou and Deng 2009), there has been little attempt to collate and analyse the available literature on their application to the Cultural Heritage domain specifically. In addition, there is no comprehensive review of the research challenges or future directions in this area. Such a review is called for given that recent literature provides a plethora of new applications aimed at enhancing the perception of art through digital content and new interaction mechanisms. Our survey fills this niche and is intended to help researchers, practitioners, art curators, and developers understand the benefits and potential hurdles of applying immersive reality to Digital Cultural Heritage.

The impetus to exploit different forms of digitization in the CH domain dates back decades. It has even been made explicit in EU commission policies that the democratization of goods that have value for all humanity should be ensured through digitization, accessibility, and interoperability to enable sharing of both information and responsibilities aimed at conserving cultural identity and awareness.

Digitization enables the spread of knowledge and the use of innovative immersive reality tools could further facilitate the access to CH in a more appealing and innovative way. The only surveys specifically from a CH perspective cover virtual museums (Styliani et al. 2009), virtual reality for tourism (Guttentag 2010), mobile AR applications for CH communication (Casella and Coelho 2013), and the challenges of AR for CH (Kounavis et al. 2012; Noh et al. 2009; Rigby and Smith 2013). A more holistic view of the field is therefore warranted. This review provides practitioners with all factors that need to be considered when determining technology adoption and the relevant technical requirements for a range of CH applications. Hence, the main objectives of this review are as follows:

- to outline state-of-the-art research and applications of augmented, virtual, and mixed reality for the CH domain;
- to reveal areas of research concentration and deficiency in this field, thereby highlighting limitations of existing technology and impediments to future research;
- to provide a framework for comparing state-of-the-art systems and to understand which solutions are most appropriate for a given application.

Thus, in this article we survey the essential aspects and the current state-of-the-art in augmented, virtual, and mixed reality from a CH perspective and describe research performed to develop applications and systems. We further summarise the adopted technologies and application areas of these studies and suggest future research directions. The remainder of the article is organized as follows: Section 2 describes the reality-virtuality continuum and provides the most accepted definitions of augmented, virtual, and mixed reality. Then Section 3 provides a detailed discussion of the enabling technologies of these immersive reality approaches from a CH
Augmented, Virtual and Mixed Reality

Section 4 evaluates the major CH-related works and identifies application areas, with a focus on the past decade, and provides technical requirements for the identified areas. Current issues and future research directions are outlined in Section 5. Finally, Section 6 provides a concluding summary.

2 THE REALITY-VIRTUALITY CONTINUUM

The reality-virtuality continuum describes the span between real and virtual environments, with Augmented Reality (AR) and Augmented Virtuality (AV) in between (Milgram and Kishino 1994). AR is close to the real world and AV is close to a virtual environment (Milgram et al. 1995), as shown in Figure 1.

Augmented reality’s most accepted definition was provided by Azuma (1997) as “a system that combines real and virtual content, provides a real-time interactive environment, and registers in 3D.” According to Milgram and Kishino (1994), AR completes reality without completely replacing it. AR studies performed in the past decades, however, have shaped the definition of AR as a system that enhances our view of the real world by adding virtual and computer-generated information (Casella and Coelho 2013; Haydar et al. 2011; Liarokapis et al. 2005; Rolland and Fuchs 2000; Vlahakis et al. 2001). An AR system typically has the following characteristics (Azuma et al. 2001): (i) It combines real-world and virtual objects, (ii) runs in real time, and (iii) allows interaction between users and virtual objects (Liarokapis 2007). Beyond this, Azuma (1997) extends the concept of AR to systems with the potential to remove objects from a real environment using graphic overlays—some scholars classify this as Mediated Reality. In general, both augmented reality and mediated reality aim to enhance our perception of and interaction with the real environment by adding virtual information and providing intuitive interaction metaphors. However, the former adds virtual information over the real-world view and displays an augmented view, whereas the latter overlays synthetic content to cover or virtually erase the real-world view or some part of it. Since it is similar to AR, mediated reality can be placed close to the real environment in the continuum.

Virtual Reality (VR), on the other hand, when fully exploited, completely immerses users in a synthetic world without any possibility of seeing the real environment, except through computer-generated representations (Carmigniani et al. 2011). VR provides synthetic content to the senses in such a way that visual perception, hearing, and touch approach the experience of an actual environment (Zhao 2009).

The third approach, AV, augments the virtual world with live scenes from the real world. Mixed Reality (MR) covers the continuum from AR to AV and aims at blending the real and virtual environments in different ways. It is thus a broad category covering various forms of AR and AV in a single technology.

While there is no generally accepted collective term for all these technologies, we will use Immersive Reality when referring to any or all of VR, AR, and MR. It should be recognised, however, that the technologies and applications discussed in this survey support varying degrees of immersion. For instance, Fishtank VR, in which a user views a conventional display using stereoscopic glasses, and Mobile AR, where a cellphone is used as a handheld AR display, can both be categorised at best as semi-immersive. We have chosen to include...
such semi-immersive approaches in this review due to their prevalence, with due recognition of their obvious limitations in evoking a sense of presence.

Providing a fine distinction among AR, AV, VR, and MR is beyond the scope of this survey article. However, we provide the following simple working definitions for the continuum:

- **Augmented Reality**: aims at enhancing our perception and understanding of the real world by superimposing virtual information on our view of the real world.
- **Augmented Virtuality**: aims at augmenting the virtual world with scenes from the real world.
- **Virtual Reality**: aims at enhancing our presence and interaction with a computer-generated environment without a means to interact with or see the real world.
- **Mixed Reality**: aims at blending real and virtual environments.

3 IMMERSIVE REALITIES AND CULTURAL COMPUTING

A number of studies demonstrate the viability of augmented-, virtual-, and mixed-reality adoption for different application areas in CH (Barsanti et al. 2015; Chrysanthi et al. 2012; Dow et al. 2005; Kang 2013; Pietroni et al. 2013).

In terms of the adoption of AR in Cultural Computing, this began as early as 2001 with the ARCHEOGUIDE project (Vlahakis et al. 2001), and Arcese et al. (2011) predict the further spread of AR in the CH sector given its appropriate fit. Based on these investigations and other applications developed by researchers, such as Zoellner et al. (2009b), Kim et al. (2009), Colizzi et al. (2010), Damala et al. (2012), Rattanarungrot et al. (2014), and D’Auria et al. (2015), the three major application areas of AR in CH are enhancing visitors’ experience, heritage reconstruction, and heritage data management and exploration.

Even though the adoption of VR in a wide spectrum of application domains began soon after the term “virtual reality” was introduced in 1989, there has since been a variety of interpretations of the term (Zhou and Deng 2009). The technological and immersive aspects of VR have contributed to the diversity of definitions. However, mediating among the technology and immersion-centered assertions, Carrozzino and Bergamasco (2010) properly defined VR as a complex technology that creates a digital environment with which users may interact and which they feel completely imersed within. Thus, immersion and interaction are essential aspects of a VR experience. In a narrower sense, since visual information tends to override all the other senses, immersion implies that the visual aspect of the experience is the ultimate sensory effect of VR. Ideally immersion, however, also includes a simulation of acoustic, haptic, smell, taste, and motion senses. A perfect virtual reality experience affects all of our senses and allows us to interact with the virtual environment naturally—as we would with our surrounding real environment. Though VR aims at enhancing one’s presence in a virtual environment, which is a cumulative effect of immersion and interaction, it does not necessarily imply that the digital environment is a representation of a fictitious world. Instead, researchers in the CC domain have exploited VR and 3D data acquisition techniques such as photogrammetry and laser scanning to build applications that are used for a variety of CH purposes, such as virtual museum, virtual reconstruction, virtual exploration, and Cultural Heritage education (Barsanti et al. 2015; Christou et al. 2006; Gaitatzes et al. 2001; Haydar et al. 2011; Mourkoussis et al. 2002; Pietroni et al. 2013). Section 4 discusses these application areas, in detail.

Mixed reality is an environment where real and virtual content coexist and interact in real time. The aspects of augmented and virtual reality merge to achieve this. MR is not just an alternative to augmented or virtual reality. Rather, it is a unique perspective that enriches humans’ perception of both real and virtual environments. Flexibility, immersion, interaction, coexistence, and enhancement are the essential aspects of a mixed reality experience. It is achieved by adopting the technological aspects of both AR and VR. Thus, an MR experience, regardless of the domain, provides a real-virtual environment, where users feel immersed and their perception of the real world is enhanced. Mixed-reality systems in the CH domain include the studies by Hall et al. (2001), Galani (2003), Benko et al. (2004), Magnenat-Thalmann et al. (2004), Magnenat-Thalmann and Papagiannakis

Regardless of the domain, the essential aspects of augmented-, virtual-, and mixed-reality applications are as follows:

- Tracking and registration
- Virtual environment modelling
- Computers, display, and devices for input and tracking
- Interaction interfaces

Interested readers can refer to Billinghurst et al. (2015), which provides a general tutorial of Augmented Reality not limited to any one domain. A particular immersive environment system is formed by making different choices for these components, and certain pre-packaged options are available as part of existing development tools, which serve to accelerate system development.

3.1 Tracking and Registration

Although both AR and VR applications seek to track the user’s viewpoint, their ultimate purpose is different. AR needs tracking to superimpose virtual content over real environment views, while in VR the purpose is to correct the perspective of displayed virtual content. Unlike AR, tracking is not a must in VR applications, unless the experience is intended to be immersive. For instance, a desktop or mobile non-immersive VR system can display virtual content without tracking the user’s pose. As with AR, tracking in Mixed Reality is needed to seamlessly register virtual content and real-world views in real time and correct the perspective to enhance users’ presence in the real-virtual environment. It is important to distinguish between calibration and tracking; the former refers to determining an initial viewpoint and camera properties, while the latter refers to continuous re-evaluation of poses to accurately align assets (Rigby and Smith 2013). The practical effectiveness of registration is highly dependent on a tracking method’s speed and accuracy.

There is a broad divide in tracking between techniques that rely on a camera as opposed to using physical sensors. For augmented reality applications in the CH domain, tracking is usually achieved by camera-based techniques (marker-based, markerless, or infrared) (Bay et al. 2005; Seo et al. 2010; Zoellner et al. 2009a), sometimes supplemented by sensor-based electromagnetic or hybrid tracking methods. There are also many ways to achieve positional tracking in VR, but they tend to rely more on sensor-based electromagnetic, acoustic, inertial, and hybrid tracking. One exception to this trend is the widespread use of camera-based infrared (IR) tracking. MR applications use similar methods to achieve tracking.

3.1.1 Camera Based.

Marker-Based Tracking. Marker-based tracking uses a digital camera, vision algorithms, and easily recognisable landmarks placed in indoor or outdoor environments—these fiducial markers could be passive (printed markers) or active (IR emitting), with the latter discussed in more detail later. Most of the existing AR applications use passive markers. However, such a tracking approach is less suitable indoors, because markers generally require good lighting condition, although such lighting conditions can be controlled. More importantly due to CH fragility, markers may not be usable due to the possibility of damage. Nevertheless, placing markers in indoor conditions is technically affordable; for instance, the ARCO project (Wojciechowski et al. 2004) uses markers to display and remove virtual objects (3D models) into and from the AR environment. Users are able to interact with the virtual objects using the markers. In another project, Mobile Augment Reality for Cultural Heritage (MARCH), Choudary et al. (2009), visual makers—in the form of coloured patches—are used to superimpose virtual objects over digital cave images. The marker-based tracking methods employed in these projects are in indoor conditions. However, the former uses fiducial markers while the latter uses visual markers.
Markerless Tracking. Vision-based tracking (also called markerless), generally, tracks camera pose by detecting and recognising geometric features in the real environment to establish 3D world and 2D image coordinate correspondences. This approach can provide realistic real-time camera pose tracking. However, rendering virtual objects over the real environment could be slow due to the large amount of processing required (Papagiannakis et al. 2008). Unlike marker-based techniques—which are dependent on easily recognisable markers, markerless tracking depends on distinguishable geometrical features, such as building corners and edges.

In computer vision, most tracking techniques can be divided into two classes: feature based (Cucchiara and Del Bimbo 2014) and model based (Uchiyama and Marchand 2012). The underlying concept of feature-based methods is to find a correspondence between 2D image features and their 3D world frame coordinates. Model-based techniques, instead, explicitly use a model of the features of tracked objects such as a CAD model or 2D templates of the object based on distinguishable features. The tracking phase is based on lines, edges, or shapes present in the model.

This tracking approach can be used for both indoor and outdoor AR applications. However, it is not always feasible if the site lacks suitable features and the model-based approach requires a database of images for each object in the real environment taken from different viewpoints. Moreover, markerless tracking is more prone to failure under conditions where the motion frequency of a camera is high—geometric features may not be detected at all or virtual objects could be misregistered.

More recently, the Kinect has been used to establish 3D world and 2D image correspondences to determine camera pose, thereby demonstrating that the combination of depth and image correspondence can provide reliable estimates of camera pose (Bostanci et al. 2015). When compared with marker-based approaches, markerless tracking has the potential of being used for both indoor and outdoor AR applications as long as the database of images of the real environment is in place. However, this approach suffers from significant processing requirements, which often introduces registration delay.

Infrared Tracking. Optical IR tracking is a method of estimating in real time the pose of a given target by tracking the position and orientation of either active or passive IR markers. The two basic characteristics that differentiate this tracking are that it always uses IR markers and is not affected by lighting conditions. Active markers are IR emitting diodes that periodically flash IR light, whereas passive markers consist of retro-reflective materials that reflect back the incoming IR lights towards the source. Usually, multiple cameras illuminate the tracking space with IR light, thereby allowing the 3D location of multiple targets to be measured. Here, it is worth distinguishing between measuring the position of a target and measuring the pose of a target. With a single marker attached to a target only its position can be tracked. Multiple markers are needed to track both position and orientation. IR tracking has low latency; however, it does not function if the line of sight between the IR source and retro-reflector is obscured. Such systems can also be affected by ambient IR radiation present in the tracking space. Haydar et al. (2011) and Barsanti et al. (2015) use IR tracking in their respective VR cultural heritage systems—the latter combines optical and inertial tracking methods to obtain robust pose tracking performance.

3.1.2 Sensor Based.

Electromagnetic Tracking. Electromagnetic tracking relies on measuring the intensity of the magnetic field between a base station and a measurement point, in various directions and orientations. This tracking system has low latency and high responsiveness, but it is subject to interference from other magnetic fields near the tracking space. However, this can be mitigated by installing the tracking system in a controlled environment.

Acoustic Tracking. Acoustic tracking estimates the pose of a viewpoint by calculating the time taken for ultrasonic sound waves to travel from a target (emitter) to a sensor, which is usually kept stable in the tracking space. Ultrasonic emitters are attached to the HMD and interaction devices if both the viewpoint and interactions are being tracked. When multiple sensors and emitters are present in the tracking space, the time difference
between the ultrasonic waves travelling through synchronised sensors and emitters provides an estimate of the orientation of the sensors relative to the emitters. Unfortunately acoustic trackers have low update rates as a result of the relatively slow speed of sound. Moreover, this tracking system is prone to measurement errors caused by ambient noise. Acoustic tracking systems provide a better accuracy when fused with other tracking methods. For instance, Hernández et al. (2007), combine acoustic and inertial tracking for a cultural heritage VR application—the authors present an immersive VR system that allows users to physically walk and track their pose while they are exploring a virtual environment.

**Inertial Tracking.** Inertial tracking is a navigation system that uses gyroscopes and accelerometers to measure the rotation and motion of a given target, thereby enabling the calculation of pose and velocity. The accelerometer measures linear acceleration to calculate the position of a target relative to some initial point. The gyroscope, on the other hand, measures angular velocity to calculate the angular rotation of a target relative to some initial orientation. Hence, the pose of a target is the integration of the measurements from the accelerometer and the gyroscope. This tracking method is inexpensive and can provide high update rates with low latency. However, it suffers from positional drift as a result of the accumulation of small measurement errors from the accelerometer and the gyroscope. Thus, relying on inertial tracking alone to estimate the position is problematic. An alternative is to fuse it with other tracking methods to obtain better positional accuracy; for instance, Hernández et al. (2007) combine acoustic and inertial tracking methods, and Barsanti et al. (2015) combine optical and inertial tracking.

3.1.3 **Hybrid Tracking.** A fusion of the aforementioned tracking methods can yield better results than when each of them are employed separately. For instance, inertial tracking suffers from positional drift but provides better accuracy for orientation measurement, and marker-based and IR tracking are affected if markers are occluded. During such situations, the data from the inertial tracker are used to estimate position until camera-based tracking is synced to the marker again. In particular, inertial tracking is often combined with the other tracking methods. Also relevant is the work of Bostanci et al. (2015), which uses Kinect to establish 3D world and 2D image correspondences and, from them, determine camera pose.

There is also a trend to combine GPS and camera-based tracking, which is a good solution in cases where the POIs are very close to each other (e.g., in a big city). With the help of the picture taken by the camera and the GPS coordinates, the device can recognize attractions in a more flexible and reliable way (Attila and Edit 2012). Additional insights about this approach are reported by Geiger et al. (2014).

Typical applications, in the CH domain, that use hybrid tracking include Vlahakis et al. (2001), Schnädelbach et al. (2002), and Miyashita et al. (2008). For instance, the ARCHEOGUIDE application (Vlahakis et al. 2001) combines markerless tracking and GPS to determine viewpoint pose.

3.2 **Virtual Environment Modelling**

In a broader sense, virtual environment modelling is the process of simulating real objects and their state in a digital space, the behavioral rules that the objects obey, and the relationships and interactions between them (Zhao 2009). To this end, there are several types of model data and modelling methods.

3.2.1 **Model Data Types.** Data acquisition methods and the aspects associated with real-world objects are the two broad perspectives used to classify data types. From a data acquisition perspective, there are three types of model data, namely actual measurement, mathematical measurement, and artificial construction (Zhao 2009). Actual measurement refers to the model data acquired through the processes of 2D and 3D scanning and any other process that involves the use of data capturing equipment. For instance, Barsanti et al. (2015) use photogrammetry to acquire the 3D data of ancient Egyptian artefacts—a wooden sarcophagus and heart scarab—and model them for VR visualisation. Mathematical measurement refers to the use of mathematical models, abstractions, and experimental analyses to generate model data of the real environment. The model data from both actual and mathematical measurement represent the real word in digital space, although they use different techniques.
to acquire the model data. Artificial construction, however, refers to model data generated by human imagination, where the world represented by the model data is completely fictitious. Since the virtual environment in most CH-based VR applications are representations of the real world, actual measurement techniques, such as photogrammetry and laser scanning, and mathematical measurement methods are most often employed.

In terms of real-world associations, model data types can be categorised into spatial structure data; physical, behavioural, and dynamic properties; and motion data (Zhao 2009). Spatial structure data refer to the geometric state of real objects; physical property data describe the physical processes and changes of real objects; behavioural property data represent the behavioural processes of real objects; and both dynamic and motion data describe the real objects’ deformation, collision, motion, and so on. Despite this range, in practice, VR systems in the CH domain tend to focus primarily on spatial structure data to represent the geometrical aspects of artefacts and use actual and mathematical measurement methods for data acquisition. A practical example is the Mont’e Prama project (Rodriguez et al. 2015), which employs 3D high-quality scans, enriched with information overlays on both mobile and museum setups.

### 3.2.2 Modelling Methods

Modeling methods can be classified according to the perception modalities of the intended user and aspects of the simulated objects in the VR environment. Accordingly, from a sensory perspective, modelling methods are classified into visual, auditory, and haptic. From the simulated object perspective, on the other hand, the modelling methods are categorised into scene appearance, physics-based behaviour, and real-virtual combined modelling (Zhao 2009). Of these, scene appearance and real-virtual combined modelling methods are common in cultural heritage VR applications, because the former focuses on representing the geometric aspects of real-world objects, and the latter refers to interfusing the computer-generated content and real world scenery to improve the efficiency and flexibility of VR modelling. During actual modelling, there are three guiding factors for determining which model data type and modelling method to employ—complexity of objects in the real world, the users’ intended modality, and the expected degree of model fidelity. Often multiple modelling methods and model data acquisition techniques are combined to generate model data to satisfy the required model fidelity.

MR applications provide a blend of current and historical (theorised) views of CH, as demonstrated by Magnenat-Thalmann and Papagiannakis (2005), Oliva et al. (2015), and Okura et al. (2015). From a technical point of view, the representation of heritage in an MR environment thus requires two distinct forms of 3D data—current and historical (Addison and Gaiani 2000). The complementary combination of these forms is referred to as “real-virtual.”

### 3.3 Devices

In general, the main devices required for augmented-, virtual-, and mixed-reality systems are displays, computers, tracking cameras, and input devices.

#### 3.3.1 Display

Presenting virtual content is perhaps the most essential aspect of immersive technologies. Presentation devices are classified according to the kind of virtual content they are designed to display—visual, auditory, or tactile. However, to date, existing CH-related applications, have focused on visual presentation. There are five types of displays. The first, Head-Mounted-Displays (HMD), can be used for AR, VR, and MR experiences. The HMDs in AR can either be optical-see-through or video-see-through. Optical-see-through allows users to see part of the real environment through the lenses, while the video-see-through HMD supplies a view from video feeds supplied by multiple wearable cameras. Optical-see-through HMDs have to overlay real space to display the augmented view—users see synthetic content and the real environment coexisting in a virtual space. In the case of video-see-through HMDs, on the other hand, a computing device processes the images coming through the cameras mounted on the HMD, augments the scene with virtual information, and renders the blended images and this approach is therefore more demanding in terms of computation. Since the user sees the real environment
through the cameras mounted on the HMD, video-see-through HMDs can trick human perception into believing that virtual and real environments coexist by introducing deliberate delay before rendering the blended image, thereby properly registering virtual information over the real environment (Rolland and Fuchs 2000). Such control over the registration process is extremely difficult with optical-see-through HMDs, because the user can see the real environment through the lenses, firsthand. In any case, the introduced latency must be very low, otherwise users will notice the time gap. HMDs in VR, on the other hand, are not see-through. These displays have been used in a wide spectrum of VR applications to present 3D virtual scenes to users. Such HMDs are connected to a computer for real-time and realistic rendering of virtual scenes. A user’s pose is tracked to correct the perspective of displayed images.

The second type of display, Spatial AR (SAR), layers virtual information directly on the real environment, either by projection using video-projectors (Carmigniani et al. 2011) or through holography, such as with the Microsoft HoloLens. Both methods rely on robust low-latency markerless tracking. A recent AR project in the CH domain that use projected displays is the Revealing Flashlight presented by Ridel et al. (2014). Applications of Holographic AR, on the other hand, are only now beginning to emerge due to the nascent technology.

The third type of display, hand-held devices (HHD), can be used for AR, VR, and MR experiences. It combines a digital camera, inertial and GPS sensors, and a portable display. These displays, when used for AR and MR experiences, use video-see-through approaches to superimpose virtual content over real environment views. Most AR research in the CH domain focuses on handheld displays (Angelopoulou et al. 2011; Casella and Coelho 2013; Kang 2013). Handheld displays are also suitable for non-immersive VR systems. Recent advances in mobile technology, such as Samsung’s Gear VR, have made it even more suitable for Immersive Reality.

The fourth type of display, a desktop screen and projection, is mainly composed of a workbench, projector, and computer. These display systems are common in visualisation environments for non-immersive and semi-immersive VR experiences. With the addition of stereo glasses, desktop displays can provide 3D scene viewing functionality for multiple users. To correct the perspective, tracking methods can be employed to track pose, though tracking is not very often utilized in non-immersive and semi-immersive settings. Gesture-based and device-based interfaces are commonly implemented to allow interaction with the displayed virtual scenes—for instance, the Etruscanning project of Pietroni et al. (2013) uses projector-based display and gesture-based natural interaction to allow users to interact with digital content aimed at experiencing a virtual reconstruction of the Etruscan Regolini Galassi tomb.

The fifth type of display, a Cave Automatic Virtual Environment (CAVE) and related technologies, is a polyhedral projection display technology that allows multiple users to experience fully-immersive and vivid 3D scenes. Multiple projection displays or screen walls—typically three to six—are conjoined to make up a cave-like cube, in which users are situated to experience enhanced presence in fully immersive 3D virtual environment. The VR systems presented by Gaitatzes et al. (2001) and Christou et al. (2006) are typical examples of CAVES in CH.

3.3.2 Computer. Computing devices are used in AR, VR, and MR to run the required software tools. From a hardware perspective, a state-of-the-art system is generally needed to generate and render realistic virtual scenes in real time. As little as a decade ago, it was common to use laptops and bulky bags that users had to carry when using the application in situ. These days, mobile devices, such as smart phones and tablets, are equipped with much better processing units and memory than high-end laptops from a few years ago. More than this, mobile devices typically incorporate a high-resolution display, inertial and touch tracking, and multiple cameras in a single small portable package, making them well suited to outdoor AR application.

3.3.3 Tracking Devices and Cameras. Cameras are used for AR and MR applications that depend on marker-based or markerless tracking. Camera and tracking devices are used in combination if a hybrid tracking approach is required. In general, the commonly used tracking devices are electromagnetic, acoustic, and inertial sensors. For instance, a relatively recent CH application—AR Teleport—by Kang (2013) exploits the smart-phone’s inbuilt inertial sensors and camera to track pose. This application is designed to allow the user to travel from present
sites to the past using rich interactions, such as jumping, blowing air, swiping with a finger, and touching buttons on a phone.

### 3.3.4 Input Devices

A range of input devices are available. To shift interaction interfaces from desktop-based Graphical User Interfaces (GUI) to more intuitive and natural ones, speech, gaze, and gesture sensors—including wearable devices, such as gloves and wireless wristbands—will substitute for conventional input devices. However, the choice of input device should depend both on the domain of the application and the system. For instance, the TOOTEKO AR application presented by D’Agnano et al. (2015) uses Near-Field-Communication (NFC) sensors attached to a 3D printed replica of an artefact as input device, which returns audio content when touched by users. In the case of AR applications that use mobile devices, input and interaction can exploit the touch-screen, microphone, and tracking sensors. More generally, the common input devices for interaction and input in VR applications are data gloves, gesture sensors, joysticks, mice, wands, gamepads, and some wearable haptic sensors. For instance, the Etruscanning project, presented by Pietroni et al. (2013), uses a Kinect sensor, and Barsanti et al. (2015) use the Leap Motion to allow users to interact with virtual scenes through motion sensing.

### 3.4 Interaction Interfaces

Interaction between users and virtual information is one of the essential aspects of immersive reality across domains. Research in the fields of Tangible User Interfaces (TUI), augmented reality, and Human Computer Interaction (HCI) aim to provide intuitive and natural interaction interfaces (Billinghurst et al. 2008; Hürst and Van Wezel 2013; Kang 2013; Kato et al. 2000; Liarokapis et al. 2005; Vlahakis et al. 2001).

Interaction also has a defining impact on the sense of presence. Although there is a range of domain-specific definitions, from a VR perspective, presence is the perception of being physically present in a non-physical world. Enhancing a user’s presence in a virtual environment, which is an essential experiential aspect of VR, is a cumulative effect of immersion and interaction. The former refers to the sense of being surrounded by a virtual environment, whereas the latter is the possible range of users’ interaction with the virtual environment. Therefore, when VR applications become sufficiently immersive and completely embed users within virtual environments, and natural interaction interfaces become a seamless metaphor for interacting with virtual surroundings, a person’s perception can be tricked so that they believe themselves to be in a separate, but realistic world. Despite the undeniable fact that immersion has taken the lion’s share of VR development, interaction plays a significant role as well.

In general, there are six types of interfaces for augmented, virtual and mixed reality systems: tangible, collaborative, device-based, sensor-based, hybrid, and multimodal interfaces.

#### 3.4.1 Tangible

A tangible interface affords interaction that exploits direct manipulation of information through physical objects, and AR’s ability to combine computer-generated content and physical environments (Ishii 2008; Shaer and Hornecker 2010). When its full potential is realised, tangible AR interfaces can support direct augmentation of tangible interfaces. Thus, the same physical object becomes both display and interaction metaphor. Here it is important to distinguish between using a physical object to interact with virtual information displayed separately elsewhere, and augmenting the physical objects with virtual information and interacting with the augmented view through the same object, which fully integrates TUI and AR. In the narrower sense, applications that use physical input devices and mobile AR applications could be considered tangible AR. The use of touch screens make this interface common in the CH domain. However, the broader case where physical objects are augmented and used as interaction metaphors is not a common approach in the CH domain as it requires physical contact with the artefacts to interact with virtual information, which is often not possible due to the fragility and size of the artefacts. However, there are some studies that do investigate this: For instance, the TOOTEKO AR application presented by D’Agnano et al. (2015) uses a tactile 3D printed object as a replica of
an actual artefact. The replica is augmented with audio content and users can touch different parts of the tactile surface and get varying audio feedback.

3.4.2 Collaborative. Collaborative interfaces make use of multiple displays such as see-through HMD and SAR to support remote, face-to-face, and shared activities (Carmigniani et al. 2011). When used for face-to-face collaboration, such interfaces rely on tabletop settings to project virtual information or on see-through HMDs. In both cases, users should be able to see the virtual information from their own perspective. On the other hand, when this interface is employed for remote collaboration, participants can wear a see-through HMD and remotely collaborate in a common virtual space. Reitmayr and Schmalstieg (2004) present such a collaborative AR application using a see-through HMD. Their system is used for collaborative navigation and information browsing at historical sites in an urban environment, thereby providing multiple features so that users can follow, guide, and meet other users based on proximity.

3.4.3 Device Based. Any interaction interface that uses GUIs, haptic interfaces, and conventional devices, such as mouse, gamepad, joystick, wand, and so on, to allow users to interact with the virtual environment, is defined as a device-based interface. Arguably, sensors are a kind of device, but it is important to distinguish between devices and sensors on the basis of their characteristic of demanding touch-based manipulation. The former requires users to physically manipulate the device to function, whereas the latter senses users’ natural interactions, such as gesture, speech, and gaze, without physical contact. For example, the interface for “Reviving the past” Gaitatzes et al. (2001), uses a hand-held navigation tool called Wanda, which is a tracked device that resembles a traditional three-button mouse but with additional features of a joystick and spatial position tracking.

3.4.4 Sensor Based. In general, sensor-based interfaces employ sensing devices to understand natural interaction modes. The flow of interaction commands is not explicitly forwarded from user to system; rather, the system actively perceives the users’ intention through sensors. Common sensors include motion tracking, gaze tracking, and speech recognition. The Etruscan project (Pietroni et al. 2013), and a VR system that presents the “path of the dead,” an important ritual in ancient Egypt (Barsanti et al. 2015), use sensor-based interfaces. The former uses the Kinect sensor to sense simple gestures such as turning one’s hands right and left and spreading the arms. Whereas the latter uses the Leap Motion sensor to allow users to interact with the displayed virtual scenes through simple hand movements such as grabbing.

3.4.5 Hybrid. A hybrid interface integrates a variety of different, but complementary, interfaces and a range of interaction devices (Zhou et al. 2008). Such interfaces should automatically accommodate a changing set of devices and the interaction techniques associated with them (Zhou et al. 2008). As a result, users can specify new modes and operations at runtime.

When used by AR applications, hybrid interfaces provide the possibility of collaboration among multiple users in the same way as collaborative interfaces. For instance, Benko et al. (2004) present a collaborative mixed reality system integrating a tracked handheld display, see-through HMDs, and multi-touch and multi-user projected displays for archaeological excavation, where users employ a tracked glove, speech commands, and a multi-touch sensitive surface to interact multimodally with the system and collaborate to navigate, search, and view data. The basic difference between collaborative and hybrid interfaces is their purpose and the variety of devices and methods supported: hybrid interfaces may be single user, where, by definition, collaborative interfaces cannot be.

Inevitably, the hybrid interface is the most commonly used one in CH-related VR systems, because it unites the benefits of sensor-based and device-based mechanisms. Accordingly, a combination of sensors and input devices is used to communicate a user’s interaction commands to the VR system. In a hybrid interface, sensors are used to track the user’s pose for rendering user-centred perspectives, while input devices, typically, are used to interact with the displayed virtual content. The VR application presented by Hernández et al. (2007) uses a hybrid interface combining a wireless pointing device and inertial and acoustic tracking sensors. The tracking
sensors are used for two tasks—to determine user’s pose and to allow users to interact with the displayed virtual environment by physically walking in the digital space.

3.4.6 Multimodal. A multimodal interface is a fusion of two or more natural interaction modes. Thus, multimodal interfaces use a combination of sensing devices to perceive humans’ natural interaction modalities. It is worth distinguishing between multimodal VR experiences and multimodal interfaces. A multimodal VR experience refers to the realism of virtual reality in terms of presence as a result of the effects of the virtual environment on the visual, auditory, and touch senses. Though a multimodal VR experience is implicit in a multimodal interface, the latter refers explicitly to the use of multiple sensors to perceive the commonly used natural interaction modes, such as speech, gaze, and gesture. It is easier to find literatures on multimodal VR than on multimodal interfaces. However, as the technology advances, multimodal interfaces will likely appear in a wider range of domains.

3.5 Systems

Based on intended flexibility, Carmigniani et al. (2011) categorises AR systems into five types: fixed indoor, fixed outdoor, mobile indoor, mobile outdoor, and mobile indoor/outdoor. However, considering AR applications in the CH domain over the past decade, a simpler categorisation into indoor and outdoor AR is warranted. Virtual reality systems, on the other hand, can be classified, based on the intended experience, into non-immersive, semi-immersive, and fully immersive. These systems are implemented by combining various tracking methods, input devices, displays, and interfaces.

3.5.1 Indoor AR. Indoor AR makes use of either marker-based or markerless tracking, see-through HMDs, spatial or handheld displays, and tangible, collaborative, hybrid or multimodal interfaces. Indoor systems do not need GPS, but if the display is an HMD, then the system might use inertial sensors to track the user’s viewpoint. For instance, Kim et al. (2009) employ markerless tracking for an indoor tour system, and Choudary et al. (2009) use visual tracking and a handheld display to enhance CH discovery. AR studies, in the cultural computing domain, that employ indoor systems, include Kim et al. (2009), Seo et al. (2010), Ridel et al. (2014), and Bostanci et al. (2015).

3.5.2 Outdoor AR. Outdoor AR relies heavily on markerless and hybrid tracking, handheld displays, and tangible interfaces. Optical-see-through HMDs and collaborative interface are used in some cases. AR studies, in the cultural computing domain, that use such systems include Vlahakis et al. (2001), Reitmayr and Schmalstieg (2004), Zoellner et al. (2009a), Seo et al. (2010), Angelopoulos et al. (2011), Mohammed-Amin et al. (2012), Kang (2013), Han et al. (2013), Caggianese et al. (2014), and D’Agnano et al. (2015).

3.5.3 Non-immersive VR. Non-immersive systems, as the name suggests, are the least immersive versions of VR experience. Such systems do not need a pose tracking method at all. The virtual environment is viewed through a desktop or handheld display. Interaction with the virtual environment can occur via device-based interfaces. A sense of presence in such virtual environments is not expected. Zara (2004) uses such a system for a web-based visualisation of CH.

3.5.4 Semi-immersive VR. Semi-immersive VR systems are more akin to a flight simulator. They often consist of a large, concave screen, a projection system, and a monitor and are more similar to large-screen movie experiences. Semi-immersive systems are a common system in museums, because they can accommodate large number of users simultaneously. Tracking is not required if the experience is intended for multiple users. However, if a single person is using the system, then tracking the user’s pose might be useful to correct the perspective of the displayed virtual images. The Etruscan project presented by Pietroni et al. (2013) is a typical example of a semi-immersive VR system implemented in the CH domain.
3.5.5 **Fully Immersive VR.** Telepresence, which is a state of being fully immersed in a virtual environment, is the ultimate effect of immersion and interaction and VR systems that support this are called fully immersive. Immersing users inside a virtual environment is achieved by displaying a virtual scene from the user’s perspective on HMDs and CAVEs. The ability to see one’s surrounding physical environment is one of the aspects that differentiates AR from VR. However, this issue also comes into play with fully immersive VR systems depending on the display device—in the case of HMD-based VR experiences, one cannot see one’s body, whereas a CAVE-based experience allows seeing one’s body and even others situated in the CAVE. Natural interaction and being situated inside a virtual environment are the essential aspects of telepresence and both HMD-based and CAVE-based VR systems are viable approaches. Interaction during a fully immersive VR experience is best achieved by employing hybrid and multimodal interfaces as device-based interfaces may break user’s immersion, because users will have to focus to some extent on the interaction devices. Fully immersive VR experiences that have been observed in CH domains include those presented by Gaitatzes et al. (2001), Christou et al. (2006), and Barsanti et al. (2015).

### 3.6 Commercial and Open Source Development Tools

There have been a number of software frameworks created specifically to support immersive reality development, and this section provides an overview of those more suited to the CH domain. The first discriminant is the choice of the Operating System (OS). This is not trivial, since not all the available frameworks are suitable for the most widely adopted Operating Systems (Android and iOS), and to reach the majority of users, the platform has to be taken into account. There are certain points of overlap between AR and VR, since some existing development platforms are suitable for both experiences.

#### 3.6.1 AR Development Toolkits.

The number of development tools is increasing almost daily, and this review serves only as a snapshot of the most commonly used current frameworks, of which Wikitude, Layar, and Vuforia are commercial and PanicAR, DroidAR, and ARToolkit are free. Wikitude is a commercial framework released in 2008 that exploits both location-based and vision-based tracking. For a description of its use in a museum environment, see Caggianese et al. (2014). Layar is the most widely used solution for location-based services. Being able to store POIs in a remote database (DB) and retrieve associated information based on user location make this framework particularly appropriate for outdoor way-finding experiences (Haugstvedt and Krogstie 2012a).

After the removal of Metaio from the market, which for years was the most powerful tool for developing vision-based AR applications, Vuforia has become the toolkit of choice for the vast majority of developers. Empler et al. (2013) present a framework for the visualization of 3D artefacts with Vuforia in archaeological contexts. Its integration with Unity3D enables well-rendered 3D models and rapid and easy cross-platform development. It supports a variety of 2D and 3D target types, including Image Targets, 3D Multi-Target configurations, and a form of addressable Fiduciary Marker known as a Frame Marker. Additional features of the SDK include localized Occlusion Detection using “Virtual Buttons,” runtime image target selection, and the ability to create and reconfigure target sets programmatically at runtime.

Moving on to free or opensource solutions, ARToolKit is a vision-based AR library that includes features such as camera position/orientation tracking, easy camera calibration code, and cross-platform development.
Table 1. A Comparison between the Most Commonly Adopted AR Frameworks in the Field of CH

<table>
<thead>
<tr>
<th>SDK</th>
<th>Purpose</th>
<th>Tracking</th>
<th>Platforms</th>
<th>Graphics</th>
<th>Cloud, Computing</th>
<th>Tracking Sensors</th>
<th>License</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wikitude</td>
<td>Indoor, Outdoor</td>
<td>Inertial, Markerless, Model based</td>
<td>iOS, Android</td>
<td>3DUnity support, 2D images, text, 3D Models (proprietary format)</td>
<td>yes</td>
<td>Camera, GPS, IMU</td>
<td>free and commercial</td>
</tr>
<tr>
<td>Layar</td>
<td>Mainly for Outdoor</td>
<td>Inertial</td>
<td>iOS, Android</td>
<td>2D images 3D models (proprietary format)</td>
<td>yes</td>
<td>GPS, IMU</td>
<td>commercial</td>
</tr>
<tr>
<td>Vuforia</td>
<td>Indoor</td>
<td>Markerless Model based</td>
<td>iOS, Android</td>
<td>3DUnity, OpenGL 3D models</td>
<td>yes</td>
<td>Camera</td>
<td>free and commercial</td>
</tr>
<tr>
<td>PanicAR</td>
<td>Only outdoor</td>
<td>Inertial</td>
<td>iOS</td>
<td>2DImaged Labels</td>
<td>no</td>
<td>GPS, IMU</td>
<td>for free</td>
</tr>
<tr>
<td>DroiAR</td>
<td>Only outdoor</td>
<td>Inertial</td>
<td>Android</td>
<td>2D Images Labels</td>
<td>no</td>
<td>GPS, IMU</td>
<td>free</td>
</tr>
<tr>
<td>ARToolKit</td>
<td>2DImages, Markers</td>
<td>Markerless</td>
<td>iOS, Android</td>
<td>3DUnity, Android</td>
<td>no</td>
<td>Camera</td>
<td>GPL</td>
</tr>
</tbody>
</table>

Distributed with complete source code, it was initially designed to run on personal computers, making the use of this SDK for the mobile development not preferable (Choudary et al. 2009).

PanicAR, distributed with a free licence, is specifically designed for iOS development and is based on sensor tracking. Kounavis et al. (2012) show its use for location-based AR to enhance the tourism experience in an outdoor scenario.

Finally, DroidAR was designed to create AR applications for Android OS with both location- and vision-based approaches. Again the source code is freely available. A test of this tool appears in the work of Quattrini et al. (2016). From our own tests, Vuforia provides the most reliable tracking in terms of rapidity and stability. In contrast, for outdoor scenarios Layar, unfortunately, has some weaknesses, especially in terms of accuracy. Table 1 shows the features and weaknesses of the listed tools.

Recently, AR toolkits based on visual-inertial odometry tracking have been gaining attention. In particular, the Google Tango project has been used for some applications in the CH domain (Lee 2017). The ARkit by Apple is also promising. Unlike Google Tango it does not requires additional hardware, but, being still in its infancy, has yet to be applied to CH.

3.6.2 VR Development Toolkits. With the mass market sale of simple VR devices (e.g., Google Cardboard, Gear VR, HTC Vive) accompanied by supporting applications, VR has become more publicly accessible and affordable.

Game Engines have become the de facto approach for implementing VR systems, due to the range of support they offer, including management of complex 3D models, interoperability of file formats, rendering, animation,
Table 2. A Comparison between the Most Commonly Adopted VR Game Engines in the Field of CH

<table>
<thead>
<tr>
<th>SDK</th>
<th>License</th>
<th>Dev.Platform</th>
<th>Mobile Platform</th>
<th>Visual Editor</th>
<th>VR Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unity 3D</td>
<td>Proprietary</td>
<td>Windows, OSX,</td>
<td>Windows Phone, iOS,</td>
<td>Yes</td>
<td>Oculus Rift, Gear VR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>eucalyptus</td>
<td>Android, Tizen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OpenSceneGraph</td>
<td>Open Source</td>
<td>Linux, Windows</td>
<td>–</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Unreal</td>
<td>free and commercial</td>
<td>Windows, Mac OS X, Linux, iOS, Android</td>
<td>Yes</td>
<td>HTC Vive, Oculus Rift, Google VR, Samsung Gear VR</td>
<td></td>
</tr>
<tr>
<td>CryENGINE</td>
<td>free and commercial</td>
<td>Windows, Mac OS X, Linux, iOS, Android</td>
<td>Yes</td>
<td>PlayStation, XBox, HTC Vive, Oculus Rift, Google VR, Samsung Gear VR</td>
<td></td>
</tr>
</tbody>
</table>

and interaction. Unfortunately, they have the drawback of being complex and represent a significant hurdle for inexperienced programmers.

The most popular game engines for VR are Unity 3D, OpenSceneGraph, Unreal Engine 4, and CryENGINE. Unity 3D is perhaps the most developer-friendly platform and is the sole one that allows fully cross-platform development. The most commonly cited drawback is that it does not allow real-time modelling. Bruno et al. (2010) demonstrate its use in the DCH domain.

OpenSceneGraph is widely used for VR, scientific visualization, visual simulation, modelling, games, and mobile applications (Baglivo et al. 2013). Although, as a high performance 3D graphics toolkit, it is more oriented towards desktop and web-based rather than mobile applications.

Unreal Engine 4 includes outstanding graphical features, and it is probably the best tool for achieving realistic results. Enables one to deploy projects to Windows PC, PlayStation 4, Xbox One, Mac OS X, iOS, Android, VR (including but not limited to SteamVR/HTC Vive, Oculus Rift, PlayStation VR, Google VR/Daydream, OSVR and Samsung Gear VR), Linux, SteamOS, and HTML5. Unreal Editor can run on Windows, OS X, and Linux.

CryENGINE is also worth mentioning, even though it is not widely used and requires expert developers. Table 2 summarizes the main features of these game engines. A more detailed cross comparison of these tools for the CH domain is provided by Herrmann and Pastorelli (2014).

3.7 Summary

Regardless of the domain, the essential aspects and enabling technologies of immersive reality applications are as follows: tracking and registration methods; virtual environment modelling; computer, display, input, and tracking devices; interaction interfaces; and systems.

AR applications in the CH domain frequently use marker-based, markerless, and hybrid tracking approaches. Optical-see-through HMD, handheld, and spatial/projected displays are the common choices for displaying...
augmented views, whereas tangible and collaborative interfaces are used more often to interact with virtual information, though alternatives do exist. AR systems in CH are more commonly outdoor than indoor. Recent advances in computer technology, however, provide the necessary enabling technologies for a combination of indoor and outdoor use. For instance, HoloLens is an optical-see-through HMD and a holographic computer, which allows users to interact with virtual content via gaze, gesture, and speech.

VR applications in the CH domain use electromagnetic, inertial, acoustic, IR, and hybrid tracking approaches with IR and inertial tracking most frequent. HMD, desktop, and CAVE displays are the common choices for displaying the virtual environment. Device-based, sensor-based, and hybrid interfaces are used most often to interact with the virtual environment, though multimodal interfaces are more intuitive and natural. The most common VR systems employed in cultural heritage are semi-immersive and fully immersive. Recent advances in HMD, tracking sensors, and computer graphics technologies enable very realistic modelling and real-time rendering of virtual environments, but this has yet to be widely adopted in CH.

4 CULTURAL HERITAGE APPLICATIONS

On the whole, cultural heritage sites and artefacts gain significant added value from enrichment through digital media. Nevertheless, many art curators believe that the use of technology relegates art to the background (Cameron and Kenderdine 2007). This attitude seems to stem from either a cultural or generational source. First, there is widespread skepticism among those not comfortable with technology about the benefits of mobile technologies. Second, and more importantly, there are issues with the way technology is used. The trend in multimedia applications is towards the show and glamour of innovation rather than a focus on solving specific problems with digital (Pierdicca et al. 2016a).

However, there is general agreement that visual CH tools suitable for users unskilled in multimedia technologies are important for CH dissemination (Cignoni and Scopigno 2008). This is particularly the case for younger participants. Generally, museum installations that do not introduce new technologies are rightly or wrongly regarded as less interesting and attract fewer visitors (Gerval and Le Ru 2015). Learning experiences in museums that rely only on labels and descriptions may be informative but they are not interactive (Lu et al. 2014). The creation of an intelligent environment that is responsive to human presence adapts dynamically and supports mobile technology makes the visit path more appealing, opening up new avenues in both tangible and intangible CH (Manovich 2006).

Tangible CH refers to physical artefacts of a society, including artistic creations, built heritage, and other physical products that are imbued with cultural significance. Conversely, Intangible Cultural Heritage indicates non-physical practices, representations, expressions, knowledge, and skills that are recognized as a vital component of Cultural Heritage (Ahmad 2006).

Immersive reality systems have proven to be a viable solution in this regard, allowing navigation, interaction, and discovery in different settings and with a variety of purposes. In archaeology, for example, the problem of the dissemination of heritage is often related to communicating goods that are either seriously damaged or definitively lost. Technologies can serve an X-Ray-like function to show what is concealed under the ground or to augment an environment with virtual reconstructions of lost heritage (Clini et al. 2016).

From these considerations arises the need for a classification of Cultural Heritage application areas to better understand where AR, VR, and MR can offer successful solutions. Accordingly, we classify the purpose of immersive reality in CH as education, exhibition enhancement, exploration, reconstruction, and virtual museums.

- Education aims at enabling users to learn the historical aspects of tangible and intangible CH.
- Exhibition enhancement is intended to improve the visitor experience at physical museums and heritage sites, typically through tour guidance.
- Exploration supports users in visualizing and exploring historical and current views of CHs to discover, interpret, and acquire new insight and knowledge.
Reconstruction aims at enabling users to visualise and interact with reconstructed historical views of tangible and intangible CHs. Two characteristics differentiate this from exploration: It does not solely target experts and the visualisation and interaction do not necessarily extend to discovery of new insights.

Virtual museums simulate and present tangible and intangible CHs in digital museum form to the public.

Some of these categories overlap. For instance, a reconstruction application might also allow a user to learn the history of the reconstructed CH. Thus, education and reconstruction purposes coexist under such conditions. Similarly, a virtual museum exhibit might very well be housed within and enhance a physical museum, thus combining with exhibition enhancement. Despite such characteristics, the central objective of a given surveyed application generally favours a particular purpose over others, and we use this as the basis for deciding the primary category.

The relationship between categories is depicted in the Venn diagram (Heberle et al. 2015) of Figure 2, which shows the technologies that dominate in overlapping categories as well as a statistical overview of the distribution of developed systems in the CH domain.

Considering the objective of these application areas, we next discuss technological suitability and technical requirements. Figure 5 summarises these requirements in tabular form. We also categorise a number of augmented-, virtual-, and mixed-reality applications on the basis of these themes. See Tables 3, 5, and 4.

(1) **Education.** In some sense every research area that deals with the dissemination and diffusion of CH must consider education. However, our focus here is on tools and applications where learning is the primary aim (Bacca et al. 2014). As an example, museum designers have recently turned to leveraging immersive realities’ capacity for spatial and temporal representation, narrative, and interactivity, real-time personalized scaffolds, and collaboration to create meaningful learning experiences in medicine and human biology (Matuk 2016). By enhancing a sense of place, for instance, by improving the visit or way-finding in a virtual environment, the learning activity can be significantly improved (Chang et al. 2015). Gargalakos et al. (2011) discuss how playful learning can cross boundaries among schools, museums, and science centers by involving participants in extended episodes of digital interaction with the exhibition. This approach provides significantly improved learning outcomes, increasing students’ curiosity, their willingness to share their experiences, and their eagerness to use new technologies and acquire knowledge. In a similar vein, the work by Invitto et al. (2014) considers various interventions and studies related to new technologies and new scientific languages based on the learning objective. The idea is to enhance the usability of the MAUS Museum through an AR application and Virtual Reality projections, related to the natural sciences (Plankton 3D and Tarbosaurus 3D). There is thus evidence that educating users on the historical aspects of both tangible and intangible CHs requires presenting the content in an entertaining environment. To this end, the system should be immersive and interactive. Tangible educational CH needs users’ active interaction with the displayed content. An interface in such cases must be as natural and intuitive as possible. Users’ inexperience with such applications should not be a constraint that prevents delivering the historical aspects as intended. A user’s age, background, and knowledge of the domain may differ, and the system should adapt accordingly. Intangible CHs, however, do not need to rely on adaptive interfaces to the same extent, because ready-made audio-visual content is presented, and user interaction is often limited to playing and pausing the content. For tangible CH, HMDs and CAVEs with a high-resolution display and realistic rendering capability can achieve the required immersion. Tracking is mandatory to enhance the immersion and interaction. A combination of inertial and IR sensors can provide the user’s pose. Interface-related tracking is better achieved by natural interaction such as gesture, gaze, and speech sensors. Tracking is not required if tangible CHs are presented, but it can enhance the experience. Both VR and MR can be used to achieve educational support in a fully immersive environment. AR may not be a

---

suitable alternative as it overlays virtual and real-world views, while the focus of educational applications tends to be historical.

(2) **Exhibition enhancement.** Enhancing a visitor’s experience can take place indoors or outdoors, or sometimes both, based on the location of CH assets. In all cases, a virtual element, such as a description, guide map, or virtual-human character, is superimposed over the users’ current view of the real world. The number, and the quality, of applied research articles in this field is high, since AR can provide a variety of solutions to help museums fulfill their role and goals (Choi 2014). Regardless of the type of installation, there is evidence that visitor interest grows when such immersive solutions are adopted (Chang et al. 2014). In the user study conducted during the ARCO project (Sylaiou et al. 2010), this same trend was evident. In fact, in many cases an immersive reality approach enables new media and storytelling that represents the major highlight of a user’s experience (Pescarin et al. 2012). The work of Liestøl (2014) in uncovering...
Table 3. The Surveyed MR Applications in the CH Domain and Their Purpose and the Enabling Technologies They Adopted

<table>
<thead>
<tr>
<th>Application</th>
<th>Purpose</th>
<th>Tracking</th>
<th>Display</th>
<th>Interface</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schädelbach et al. (2002)</td>
<td>Reconstruction</td>
<td>Hybrid (rotation sensors and GPS)</td>
<td>Custom-built (Tripod-mounted display)</td>
<td>Tangible (Device-based)</td>
<td>Non-immersive</td>
</tr>
<tr>
<td>Magnenat-Thalmann et al. (2004)</td>
<td>Reconstruction</td>
<td>Markerless</td>
<td>Mobile (laptop)</td>
<td>Tangible</td>
<td>Non-immersive</td>
</tr>
<tr>
<td>Dow et al. (2005)</td>
<td>Education</td>
<td>Hybrid (GPS and IMU)</td>
<td>Mobile (Audio presentation device and tablet)</td>
<td>Hybrid (Tangible, device-based)</td>
<td>Non-immersive</td>
</tr>
<tr>
<td>Liarokapis et al. (2007)</td>
<td>Virtual museum</td>
<td>Hybrid (Marker-based and sensors)</td>
<td>Mobile (laptop)</td>
<td>Multimodal (audio-visual)</td>
<td>Non-immersive</td>
</tr>
<tr>
<td>Chrysanthi et al. (2012)</td>
<td>Education</td>
<td>Marker-based</td>
<td>Plasma screen</td>
<td>Tangible</td>
<td>Non-immersive</td>
</tr>
<tr>
<td>Durand et al. (2014)</td>
<td>Reconstruction</td>
<td>Sensor-based (IMU)</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Non-immersive</td>
</tr>
<tr>
<td>Okura et al. (2015)</td>
<td>Reconstruction</td>
<td>Hybrid (GPS and Markerless)</td>
<td>Mobile</td>
<td>Hybrid (Tangible, device-based)</td>
<td>Non-immersive</td>
</tr>
<tr>
<td>Benko et al. (2004)</td>
<td>Exploration</td>
<td>Hybrid (inertial, electromagnetic and acoustic)</td>
<td>Combo (HMD, projected table, large screen)</td>
<td>Hybrid (Multimodal and collaborative)</td>
<td>Fully immersive</td>
</tr>
<tr>
<td>Magnenat-Thalmann and Papagiannakis (2005)</td>
<td>Reconstruction</td>
<td>Markerless</td>
<td>HMD</td>
<td>Hybrid (Tangible, device-base)</td>
<td>Fully immersive</td>
</tr>
</tbody>
</table>

the Appian way and Ozden et al. (2014) on user interaction modules for two Istanbul museums are good examples. In another case (Petridis et al. 2013), the user experience was made more immersive, engaging, and interactive at the Herbert Museum and Art Gallery. In Sdegno et al. (2015), painted architecture by Paolo Veronese was brought to life thanks to a 3D reconstruction, while in Pierdicca et al. (2015b) and Clini et al. (2014) the famous painting "La Città Ideale" was augmented with digital information without requiring artificial markers. The use of mobile devices is increasing in the cultural and museum sectors, as are the number of apps (e.g., on Google Play and the Apple Store) and many Museums’ Apps are good examples of technological integration with experience design.

In general, both AR and MR can be employed for exhibition enhancement. VR cannot be used since it blocks the real-word views. See-though HMDs, however, can deliver immersive experiences of both indoor and outdoor sites, but the virtual elements should not distract visitors’ view of the real world, because the aim is enhancing a visit experience at physical museums and CH sites, not substituting it with virtual views. Therefore, the rendering should be vivid and realistic, the tracking must be robust, and the registration must be fast, especially, if optical-see-through HMDs are used. Otherwise, a user’s experience will be unpleasant. In the case of indoor systems, a combination of markerless and sensor-based tracking methods can be employed. Marker-based and other approaches that need physical attachment to a CH asset should
<table>
<thead>
<tr>
<th>Application</th>
<th>Purpose</th>
<th>Tracking</th>
<th>Display</th>
<th>Interface</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wojciechowski et al. (2004)</td>
<td>Virtual museum</td>
<td>No pose tracking</td>
<td>Desktop screen</td>
<td>Web-based (Mouse and keyboard)</td>
<td>Non-immersive</td>
</tr>
<tr>
<td>Laycock et al. (2008)</td>
<td>Reconstruction</td>
<td>No pose tracking</td>
<td>Desktop screen</td>
<td>Device-based</td>
<td>Non-immersive</td>
</tr>
<tr>
<td>Richards-Rissetto et al. (2014)</td>
<td>Education</td>
<td>No pose tracking</td>
<td>Screen/wall</td>
<td>Device-based</td>
<td>Non-immersive</td>
</tr>
<tr>
<td>Baldassini and Gaiani (2014)</td>
<td>Education</td>
<td>No pose tracking</td>
<td>Desktop</td>
<td>Device-based</td>
<td>Non-immersive</td>
</tr>
<tr>
<td>Bruno et al. (2010)</td>
<td>Virtual museum</td>
<td>No pose tracking</td>
<td>Stereoscopic screen</td>
<td>Device-based</td>
<td>Semi-immersive</td>
</tr>
<tr>
<td>Haydar et al. (2011)</td>
<td>Virtual museum</td>
<td>Optical</td>
<td>Large Screen with LCD glasses</td>
<td>Device-based</td>
<td>Semi-immersive</td>
</tr>
<tr>
<td>Pietroni et al. (2013)</td>
<td>Reconstruction</td>
<td>No pose tracking</td>
<td>Screen/wall</td>
<td>Sensor-based (through natural gesture)</td>
<td>Semi-immersive</td>
</tr>
<tr>
<td>Richards-Rissetto et al. (2014)</td>
<td>Education</td>
<td>No pose tracking</td>
<td>3D stereo display</td>
<td>Sensor-based</td>
<td>Semi-immersive</td>
</tr>
<tr>
<td>Hsieh et al. (2014)</td>
<td>Virtual museum</td>
<td>No pose tracking</td>
<td>Projector</td>
<td>Hybrid (Device-based and sensor-based)</td>
<td>Semi-immersive</td>
</tr>
<tr>
<td>Marton et al. (2014)</td>
<td>Exploration</td>
<td>Custom-designed (Image-based)</td>
<td>Back-projected screen</td>
<td>Device-based</td>
<td>Semi-immersive</td>
</tr>
<tr>
<td>Bustillo et al. (2015)</td>
<td>Education</td>
<td>Hybrid (optical and markerless)</td>
<td>Projector</td>
<td>Device-based</td>
<td>Semi-immersive</td>
</tr>
<tr>
<td>Gaitatzes et al. (2001)</td>
<td>Education</td>
<td>Electromagnetic</td>
<td>CAVE</td>
<td>Device-based</td>
<td>Fully immersive</td>
</tr>
<tr>
<td>Acevedo et al. (2001)</td>
<td>Exploration</td>
<td>Hybrid (IMU, optical/electromagnetic)</td>
<td>CAVE</td>
<td>Device-based</td>
<td>Fully immersive</td>
</tr>
<tr>
<td>Gutierrez et al. (2004)</td>
<td>Reconstruction</td>
<td>No pose tracking</td>
<td>CAVE</td>
<td>Device-based (radio-based remote control)</td>
<td>Fully immersive</td>
</tr>
<tr>
<td>Christou et al. (2006)</td>
<td>Virtual museum</td>
<td>Electromagnetic</td>
<td>CAVE</td>
<td>Multimodal</td>
<td>Fully immersive</td>
</tr>
<tr>
<td>Hernández et al. (2007)</td>
<td>Virtual museum</td>
<td>Hybrid (Acoustic and Inertial)</td>
<td>HMD</td>
<td>Sensor-based</td>
<td>Fully immersive</td>
</tr>
<tr>
<td>Haydar et al. (2011)</td>
<td>Virtual museum</td>
<td>Optical</td>
<td>HMD</td>
<td>Device-based</td>
<td>Fully immersive</td>
</tr>
<tr>
<td>Barsanti et al. (2015)</td>
<td>Virtual museum</td>
<td>Hybrid (IMU and optical)</td>
<td>HMD</td>
<td>Sensor-based</td>
<td>Fully immersive</td>
</tr>
<tr>
<td>Katsouri et al. (2015)</td>
<td>Virtual museum</td>
<td>Optical</td>
<td>CAVE</td>
<td>Device-based</td>
<td>Fully immersive</td>
</tr>
</tbody>
</table>
be avoided, because such practices damage the historical value of CH assets. Regarding outdoor systems, a combination of location-based, sensor-based, and markerless methods can be used to achieve tracking. If users can approach the CH asset, then a markerless method is a suitable choice; otherwise, long-range optical sensors and GPS localization are more appropriate. Most of the time, visitors tend to attend museums and CH sites in groups. Hence, the interaction interface should be collaborative and intuitive so that users can experience the visit from their own perspective and at the same time collaborate with co-located visitors. This can be extended to accommodate remote collaboration between spatially distributed visitors.

(3) Exploration. Exploration-focused applications primarily focus on the historical and current aspects of tangible archaeological CHs, especially, to allow users to discover, explore, visualize, and manipulate the content, thereby leading to knowledge creation and new insights. Users of such applications are assumed to have expertise in the domain; therefore, the system can assume prior knowledge of domain-specific visualization. Exploration-focused applications need hybrid tracking, a combination of complementary displays, collaborative and multimodal interfaces, and distributed and immersive environments. Single-technology tracking is not sufficient due to its current lack of accuracy. As the users are assumed to be experts in the domain, the tracking method can focus on accuracy over user experience, however, this should not be at a cost of compromising users’ comfort. To this end, a combination of sensor-based methods can fit the purpose of indoor environments, while location and sensor-based methods can achieve outdoor tracking. If the exploration tasks are at distributed locations, then the pose readings from all locations must be synchronised; otherwise, users cannot collaborate seamlessly. A suitable combination of displays could consist of HMDs and table-top projectors for indoor, and HMDs for outdoor, settings. Table-top projectors and CAVEs can accommodate multiple users at a time, and HMDs can display user-centred perspectives. Interactions with the displayed content rely heavily on the accuracy of interface-related tracking methods. For this, sensor-based input devices and natural interaction mode sensors should be used in combination. MR and AR are the best choices for exploration purposes, because users can see both the real-world and virtual views. This feature is especially invaluable in archaeological settings. The possibility of enriching reality with computer-generated information, providing innovative information access at CH sites, has been noted in recent research. Verykokou et al. (2014), for instance, visualize a part of the Middle Stoa in the Ancient Agora of Athens (see Figure 3(a)) and users have the opportunity to see what this building looked like in ancient times, as its three-dimensional model is displayed on the camera view of their device.

Fig. 3. Examples of AR applications for archaeological purposes. Ancient architecture visualized in its original location thanks to the use of location-based AR.
projected onto the modern-day ruins. Related examples of AR for exploration can be found in the work of Etxeberria et al. (2012), Pierdicca et al. (2015a), Stancoc et al. (2012), Deliyiannis and Papaioannou (2014), and Empler (2015). From this brief analysis, three main points arise: First, AR applications in this domain are restricted to sensor-based AR (see Section 3.1.2), and interaction with virtual archaeological content in museums using AR has not spread (VR provides more possibilities in this regard). Second, AR is mainly used to visualize lost or posited artefacts. Third, for archaeology the use of geomatics applications is unavoidable (Portalés et al. 2009). Virtual reconstructions (see Figure 3(b)), in fact, must rely on accurate data sources (Pierdicca et al. 2016b; Quattrini et al. 2016). Interaction is key to making such experiences more attractive for users and involving them actively in the exploration process, as shown in the well-designed mobile interaction solutions of Wiley and Schulze (2015) and Kang (2013). Since the mid-2000s, archaeology has benefited from the widespread availability of digital 3D models (Comes et al. 2014), allowing developers to represent difficult to reach environments, for example, in underwater conditions (Haydar et al. 2011). The challenge, discussed further in Section 5 is twofold: on the one hand, describing a known workflow that moves from data acquisition to the visualization in an immersive environment and, on the other hand, making these data portable and suitable for different devices or platforms. Besides, these technologies can better accommodate collaborative and multimodal interaction. VR systems cannot meet these requirements to the same extent. However, if collaboration is not needed, then HMD-based VR can suffice.

Reconstruction. Applications for reconstruction display reconstructed views of tangible and intangible CHs. Such applications allow users to visualise CH assets that existed only in the past or that partially exist. Reconstructed assets can be presented in three forms: tangible, intangible, and a blend of both. AR and MR are best suited to tangible and a blend of tangible and intangible, because both technologies can superimpose the reconstructed views over their historical location. Additional information beyond the virtual reconstruction itself can also be overlaid (Saggio and Borra 2011). To ensure the preservation of artifacts, such as statues or paintings, they must be analyzed to diagnose physical frailties that could result in permanent damage. While such a diagnosis is aided by advancements in digital imaging techniques and computer-aided analysis, the ability to work directly with the artifact in the field remains limited. Several examples of different kind of diagnosis are reported by Colizzi et al. (2010). Of particular interest is the work of Girbacia et al. (2013) on a workflow for the restoration of religious heritage, starting from the reconstruction of statues and extending to their in-place geo-located visualization in AR. Vamoni et al. (2012) describe ARTifact, a tablet-based augmented reality system that enables on-site visual analysis (see Figure 4(a)). Their idea is to use overlaid layers to represent images acquired from different data sources. Another tool (Figure 4(b)), "the revealing flashlight" (Ridel et al. 2014), is intended to distinguish details obscured by aging effects. This system works by projecting an expressive 3D visualization that highlights features, based on an analysis of previously acquired geometry at multiple scales. The novelty mainly lies in the interaction, which is based on gestures.

VR, on the other hand, is suitable for intangible reconstruction and visualising tangible assets in indoor environments, because this does not rely on displaying reconstructed views over their historical location. A review of such applications with a focus on interaction and gamification is provided by Katersis et al. (2015). In the case of AR and MR, positional tracking can be achieved using a combination of GPS, orientation sensors, and markerless tracking. VR requires tracking to correct perspective, and this can be achieved through orientation sensors attached to HMDs or stereo glasses. Users of a reconstruction application may range from domain experts, students, to the general public, preferably with gamification (Münster et al. 2017; Papaefthymiou et al. 2017). Hence, the system should be inclusive of these groups’ background. To achieve such inclusiveness, more focus should be given to interaction and presentation aspects. As a result, such applications should have a multimodal interface and immersive features. In general, see-through HMDs’ can fit the requirement for AR and MR systems, and CAVEs will do the same for VR.

(5) Virtual museums. In general, virtual museums simulate physical museums and CH sites including their tangible and intangible assets. Much of the time, such assets are inaccessible and fragile. Hence, the simulation must be very realistic and detailed to serve as a replica of artefacts so that users cannot easily discern differences between the originals and their replicas. Such simulations enhance users’ presence in virtual museums, thereby tricking users into feeling as if they are physically present at an actual museum or CH site in situ. This can be extended to represent users as virtual-human characters inside the simulated environment so that users who share this environment can see co-located users. To achieve this, the chosen modelling method should blend pre-rendered scenes and virtual-human characters in real time. Also, such virtual characters should not be mere avatars but close approximations of the actual users to simulate real-life interaction. Hence, the simulation should consider the behavioural and physical properties of users. In addition, the system should be fully-immersive. Users should be able to interact via gesture, gaze, speech, and movement. The enabling sensors for such interfaces should not remind users of their attachment to the real world. Otherwise, presence in the virtual museum will be interrupted. For instance, sensors physically attached to users may require direct manipulation compared to remotely placed sensors, which may result in decreased presence. Also, the interaction should create a perception of physical movement inside a digital environment. In general, HMD-based AV and CAVE-based VR environments can achieve a fully immersive virtual museum. However, large-scale CAVEs are more appropriate, because such environments can accommodate multiple users, and virtual-human characters are unnecessary.

The following sections discuss the surveyed articles from the perspectives of enabling technology, system, and purpose as observed in the survey. Moreover, some suggestions are made based on these observations as to which technologies are most suited to a given purpose. These suggestions differ from those discussed above, because those suggestions are made based on the central objective of the identified application areas, whereas the suggestions below are based on the technologies adopted by the surveyed works.

4.1 AR Applications

Most AR applications are aimed at exhibition enhancement, followed by reconstruction and exploration. Table 5 and Figure 6 show the details of these applications. Hybrid tracking is a relatively common approach, with markerless, sensor-based, and marker-based methods used in that order. In terms of presentation devices, mobile
Table 5. The Surveyed AR Applications in the CH Domain and Their Purpose and the Enabling Technologies They Adopted

<table>
<thead>
<tr>
<th>Application</th>
<th>Purpose</th>
<th>Tracking</th>
<th>Display</th>
<th>Interface</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wojciechowski et al. (2004)</td>
<td>Exhibition</td>
<td>Marker-based</td>
<td>Desktop screen</td>
<td>Hybrid (Tangible and Web-based)</td>
<td>Indoor</td>
</tr>
<tr>
<td>Miyashita et al. (2008)</td>
<td>Enhancement</td>
<td>Hybrid (Markerless and IMU)</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Indoor</td>
</tr>
<tr>
<td>Portalés et al. (2009)</td>
<td>Reconstruction</td>
<td>Markerless</td>
<td>HMD</td>
<td>Natural (Movement-based)</td>
<td>Indoor</td>
</tr>
<tr>
<td>Choudary et al. (2009)</td>
<td>Exploration</td>
<td>Marker-based</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Indoor</td>
</tr>
<tr>
<td>Zoellner et al. (2009b)</td>
<td>Reconstruction</td>
<td>Markerless</td>
<td>Mobile (UMPC and MovableScreen)</td>
<td>Tangible</td>
<td>Indoor</td>
</tr>
<tr>
<td>Kim et al. (2009)</td>
<td>Exhibition</td>
<td>Markerless</td>
<td>Laptop</td>
<td>Tangible</td>
<td>Indoor</td>
</tr>
<tr>
<td>Haydar et al. (2011)</td>
<td>Reconstruction</td>
<td>Marker-based</td>
<td>HMD</td>
<td>Tangible</td>
<td>Indoor</td>
</tr>
<tr>
<td>Damala et al. (2012)</td>
<td>Exhibition</td>
<td>No pose tracking (audio-based AR)</td>
<td>Mobile</td>
<td>Tangible (3D print with tactile surface)</td>
<td>Indoor</td>
</tr>
<tr>
<td>Ridel et al. (2014)</td>
<td>Exploration</td>
<td>Hybrid (electromagnetic and optical)</td>
<td>SAR</td>
<td>Tangible</td>
<td>Indoor</td>
</tr>
<tr>
<td>D’Agnano et al. (2015)</td>
<td>Exhibition</td>
<td>No pose tracking</td>
<td>Mobile</td>
<td>Tangible (AR)</td>
<td>Indoor</td>
</tr>
<tr>
<td>Damala et al. (2016)</td>
<td>Exhibition</td>
<td>Hybrid (IMU and markerless)</td>
<td>Mobile</td>
<td>Tangible (iPhone in a loupe-like wooden case)</td>
<td>Indoor</td>
</tr>
<tr>
<td>Breuss-Schneeweis (2016)</td>
<td>Exhibition</td>
<td>Marker-based</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Indoor</td>
</tr>
<tr>
<td>Invitto et al. (2014)</td>
<td>Education</td>
<td>Hybrid</td>
<td>Desktop</td>
<td>Natural</td>
<td>Indoor</td>
</tr>
<tr>
<td>Chang et al. (2014)</td>
<td>Exhibition</td>
<td>Markerless</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Indoor</td>
</tr>
<tr>
<td>Petridis et al. (2013)</td>
<td>Exhibition</td>
<td>Marker-based</td>
<td>Mobile</td>
<td>Natural</td>
<td>Indoor</td>
</tr>
<tr>
<td>Sdegno et al. (2015)</td>
<td>Exhibition</td>
<td>Marker-based</td>
<td>Combo</td>
<td>Tangible</td>
<td>Indoor</td>
</tr>
<tr>
<td>Pierdicca et al. (2015b)</td>
<td>Exhibition</td>
<td>Markerless</td>
<td>Combo</td>
<td>Tangible</td>
<td>Indoor</td>
</tr>
<tr>
<td>Clini et al. (2014)</td>
<td>Exhibition</td>
<td>Markerless</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Indoor</td>
</tr>
<tr>
<td>Dieck et al. (2016)</td>
<td>Exhibition</td>
<td>Sensor-based</td>
<td>HMD</td>
<td>Natural (Gaze)</td>
<td>Indoor</td>
</tr>
<tr>
<td>Vlahakis et al. (2001)</td>
<td>Exhibition</td>
<td>Hybrid (GPS and compass)</td>
<td>Combo (Mobile and HMD)</td>
<td>Multimodal</td>
<td>Outdoor</td>
</tr>
<tr>
<td>Reitmayr and Schmalstieg (2004)</td>
<td>Exhibition</td>
<td>Sensor-based (GPS)</td>
<td>HMD</td>
<td>Collaborative</td>
<td>Outdoor</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Application</th>
<th>Purpose</th>
<th>Tracking</th>
<th>Display</th>
<th>Interface</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pierdicca et al. (2015a)</td>
<td>Exploration</td>
<td>Sensor-based (GPS)</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Outdoor</td>
</tr>
<tr>
<td>Girbacia et al. (2013)</td>
<td>Reconstruction</td>
<td>Markerless</td>
<td>Mobile</td>
<td>Natural</td>
<td>Outdoor</td>
</tr>
<tr>
<td>Fritz et al. (2005)</td>
<td>Exhibition enhancement</td>
<td>Sensor-based (IMU)</td>
<td>Custom-built binocular-like video see-through</td>
<td>Tangible</td>
<td>Outdoor</td>
</tr>
<tr>
<td>Zoellner et al. (2009a)</td>
<td>Reconstruction</td>
<td>Markerless</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Outdoor</td>
</tr>
<tr>
<td>Haugstvedt and Krogstie (2012b)</td>
<td>Reconstruction</td>
<td>Sensor-based (GPS)</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Outdoor</td>
</tr>
<tr>
<td>Chang et al. (2015)</td>
<td>Education</td>
<td>Hybrid</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Outdoor</td>
</tr>
<tr>
<td>Han et al. (2013)</td>
<td>Reconstruction</td>
<td>Hybrid (GPS and Markerless)</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Outdoor</td>
</tr>
<tr>
<td>Amato et al. (2013)</td>
<td>Exhibition enhancement</td>
<td>Hybrid (GPS, RFID, compass)</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Outdoor</td>
</tr>
<tr>
<td>Empler et al. (2013)</td>
<td>Reconstruction</td>
<td>Markerless</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Outdoor</td>
</tr>
<tr>
<td>Caggianese et al. (2014)</td>
<td>Exhibition enhancement</td>
<td>Hybrid (GPS and IMU)</td>
<td>HMD</td>
<td>Natural</td>
<td>Outdoor (Gesture-based)</td>
</tr>
<tr>
<td>Pacheco et al. (2015)</td>
<td>Reconstruction</td>
<td>Hybrid (GPS and IMU)</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Outdoor</td>
</tr>
<tr>
<td>Huang et al. (2016)</td>
<td>Exhibition enhancement</td>
<td>Hybrid (GPS and IMU)</td>
<td>Combo (Mobile and HMD)</td>
<td>Collaborative</td>
<td>Outdoor</td>
</tr>
<tr>
<td>Canciani et al. (2016)</td>
<td>Reconstruction</td>
<td>Markerless</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Outdoor</td>
</tr>
<tr>
<td>Petrucco and Agostini (2016)</td>
<td>Reconstruction</td>
<td>Hybrid (GPS and IMU)</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Outdoor</td>
</tr>
<tr>
<td>Angelopoulou et al. (2011)</td>
<td>Exhibition enhancement</td>
<td>Marker-based</td>
<td>Mobile</td>
<td>Collaborative</td>
<td>Indoor and outdoor</td>
</tr>
<tr>
<td>D’Auria et al. (2015)</td>
<td>Exhibition enhancement</td>
<td>Hybrid (IMU, GPS)</td>
<td>Mobile with spatial headphone</td>
<td>Natural</td>
<td>Indoor (Audio-based)</td>
</tr>
<tr>
<td>Madsen and Madsen (2016)</td>
<td>Reconstruction</td>
<td>Sensor-based (IMU)</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Indoor and outdoor</td>
</tr>
<tr>
<td>Vanoni et al. (2012)</td>
<td>Reconstruction</td>
<td>Markerless</td>
<td>Mobile</td>
<td>Tangible</td>
<td>Indoor and outdoor</td>
</tr>
</tbody>
</table>

Displays make up the majority followed by HMDs, a combination of diverse displays, desktop, custom-built, and SAR displays. In terms of interfaces, most applications use a tangible interface, followed by natural, collaborative, multimodal, and hybrid interfaces. In addition, most of the surveyed applications were targeted at indoor conditions and a few applications for both indoor and outdoor environments.

There are two notable differences in the environmental settings for AR applications.

4.1.1 Indoor Systems. Most indoor applications focus on exhibition enhancement. The tracking methods used are marker based and markerless. A significant number of indoor applications use mobile devices for display, and tangible interaction interfaces are the most common.

4.1.2 Outdoor Systems. The majority of outdoor applications are aimed at reconstruction. Hybrid tracking is often adopted. In terms of display and interface, mobile devices and tangible interface are common choices, respectively.
In general, indoor applications are more suited to exhibition enhancement experiences since physical museums tend to use such applications for virtual tour guidance more often than outdoor CHs. Outdoor applications are more suitable for a reconstruction approach, because it is then possible to overlay reconstructed historical views over the real world. Moreover, a reconstruction theme is often applied to outdoor sites that have been demolished or worn away.

4.2 VR Applications

In our findings, the majority of VR applications serve virtual museums, followed by education, reconstruction, and exploration purposes, in that order. Table 4 and Figure 7 show the details of these applications. Most do not use any tracking methods at all. This is because these applications are non-immersive or semi-immersive. The remaining applications use hybrid, electromagnetic, and optical tracking methods. In terms of presentation devices, a screen/projector is used by the majority of the applications followed by CAVE and HMD. In terms of interfaces, most of the applications use a device-based interface followed by sensor-based, multimodal, and hybrid interfaces. Applications tend to be semi-immersive.

We discuss these applications from the perspective of their level of immersion because their systems range across non-immersive, semi-immersive, and fully immersive environments.

4.2.1 Non-immersive. The areas of non-immersive applications are education, virtual museums, and reconstruction. These applications do not use any tracking methods. They employ desktop screens for displaying the virtual content, and use device-based interfaces. Therefore, desktop screen and device-based interfaces seem to be sufficient for non-immersive experiences for education, virtual museum, and education themes. Pose tracking is not required for non-immersive systems.
4.2.2 Semi-immersive. Virtual museums and education are the areas of the majority of the semi-immersive applications. However, a few applications serve reconstruction and exploration purposes. Electromagnetic, optical, and hybrid methods are used to track the pose of users and interaction devices in a few applications, but most applications do not use any tracking. This is acceptable given that users see pre-rendered virtual content and most often tracking is only required for interaction. Moreover, tracking may be unnecessary if a gamepad or mouse is used. In terms of presentation devices, back-projected screens and 3D stereo displays are common choices for semi-immersive applications. Most applications use device-based interfaces, with a few using sensor-based and hybrid interfaces. Hence, optical, electromagnetic, and hybrid tracking methods; back-projected screens and 3D stereo displays; and device-based, sensor-based, and hybrid interfaces are viable for semi-immersive systems intended for virtual museums, reconstruction, education, and exploration.
4.2.3 Fully Immersive. A virtual museum is the most frequent application area for fully immersive VR. A few applications achieve education and exploration themes, though. Most applications use hybrid tracking with a few systems employing electromagnetic and optical methods. Fully immersive experiences are achieved by CAVE and HMD displays. Device-based interfaces are widely adopted. However, a few applications use multimodal and sensor-based interfaces. Therefore, fully immersive VR experiences to support virtual museums, education, and exploration are best achieved by adopting electromagnetic, optical, and hybrid tracking methods, CAVE and HMD displays, and device-based, sensor-based, and multimodal interfaces.

4.3 MR Applications

The majority of the surveyed MR applications exhibit a reconstruction purpose followed in order by education, exploration, and virtual museums. Table 3 and Figure 8 show the details of these applications. Most of these applications are designed for non-immersive experiences. Hybrid tracking, often a fusion of GPS and markerless, GPS and IMU, and inertial, electromagnetic, and acoustic tracking methods are used. Mobile displays are commonly used to present visual and audio content. However, some systems also use custom-built HMDs, and combinations of different types of presentation devices to display real-virtual content.

MR applications in the CH domain are not as widespread as AR and VR. This is understandable given that the technological aspects of MR are still in their infancy. However, when robust real-time tracking, 3D registration, realistic virtual environments, natural interfaces, and presentation devices for vivid experiences reach fruition, more MR applications will likely appear in the CH domain. Considering the current systems in the domain, however, hybrid tracking, HMD and mobile display, and tangible interface seem to satisfy the needs for implementing MR in the CH domain, especially when focusing on reconstruction.

5 DISCUSSION: CURRENT ISSUES AND FUTURE DIRECTIONS

This survey provides an exploration of research and examples of the different way in which cultural artefacts can be experienced in an immersive form through the application of AR, MR, and VR technology. The taxonomy provided in Section 4 demonstrates that these technologies are suited in a wide variety of sub-domains. What emerges in the main is the need for curators to provide users with a new perspective on their collections.
Museums, for example, can increase their appeal by augmenting their artifacts or paintings with digital media, archaeological areas can bring to life lost architectures or ruins. However, there are still many hurdles preventing the acceptance and diffusion of immersive technologies in Cultural Heritage. These hurdles are mainly due to (i) technological limitations, (ii) content complexity, and (iii) human factors. First, there are many aspects of immersive technology, such as sensor-based tracking, that could benefit from further attention. Second, the model resolution requirements of CH often exceed the capacity of current technology, particularly with respect to internet retrieval. Third, without careful consideration of human factors as they affect the user experience, immersive technologies are unlikely to experience widespread adoption.

Considering the ongoing research on tracking and registration, realistic rendering, HCI, and CH, we expect further research in the following areas:

(1) **Robust Tracking.** Sensor-based tracking using commercial devices, particularly in an outdoor CH environment, remains error prone and has necessitated hybrid solutions. However, the situation is likely to improve with recent investment in these technologies. In this respect, camera-based approaches are a more mature technology in terms of accuracy and reliability, but there is still no prevailing standard.

(2) **Standardisation.** Despite its advantages, immersive reality has not been widely adopted by art curators and managers. Partly this can be traced to a lack of standardization, which could facilitate rapid, sequential development projects. In AR the only available standard is the Augmented Reality Markup Language (ARML) 2.0, provided by Open Geospatial Consortium (OGS), which is primarily oriented towards location-based services. Proposed alternatives include a service-oriented strategy (Rattanarungrot et al. 2014) or standardization of the entire AR architecture (Sambinelli and Arias 2015). The community could also benefit from a self-documenting standard data format that describes the structure as well as data types and meanings of values for text, 3D models, images, audio, and video. VR systems also lack effective formal or de facto standards. Fragmentation of descriptive, structural and administrative metadata for 3D media causes interoperability issues that hamper the exploitation of 3D models on different platforms. However, in VR the most widespread standard is X3D, a royalty-free ISO standard XML-based file format for representing 3D computer graphics. The adoption of a common representation for scanned models would represent a turning point for researchers dealing to the acquisition and reconstruction of ancient artefacts (Fernández-Palacios et al. 2017). Visualization issues are mainly entrusted to the worldwide adoption of WebGL, which offers the ability to render 3D scenes within any common browser.

(3) **User-Driven Semantics.** To deal with clutter in information rich environments, allow users to focus on particular points of interest and adapt the cultural heritage experience to their preferences, one approach is to exploit semantic web technologies, such as OWL, RDF, and SPARQL. While this approach is not novel in itself (Damala and Stojanovic 2012; Hatala and Wakkary 2005; Kovachev et al. 2014; Matuszka 2015; Van Aart et al. 2010), it does open up possibilities for citizen participation (Ruta et al. 2014).

(4) **Tangible AR.** A number of augmented reality applications use tangible interfaces in a much narrower scope than its potential warrants. We hope to see more research that integrates Tangible User Interfaces and augmented reality so that future applications, irrespective of domain, will be able to augment physical objects with virtual content and enable interaction with this content through the augmented objects.

(5) **Fully Immersive VR.** Fully immersive VR systems are not common for a number of reasons. The expense of CAVE technology being one. Recent advances, however, provide relatively affordable technologies such as the Oculus Rift, Microsoft HoloLens, and the HTC Vive, which are HMDs capable of high-resolution rendering, pose tracking, and natural interaction with virtual content. Thus, fully immersive VR applications will likely appear soon in a wider range of domains. We hope the CH domain will make use of such technologies to realise virtual museums, reconstruction, exploration, and education in a fully immersive virtual environment.


(6) **Multimodal Interfaces.** A multimodal interface is a very intuitive interface and AR, VR, and MR systems can exploit this potential. However, it is extremely difficult to implement such interfaces with the state-of-the-art in HCI. However, as research in sensor technology, speech recognition, and artificial intelligence advance, multimodal interfaces will likely become more prevalent in CH and other domains, thereby allowing users to interact with virtual content through all their senses.

6 **CONCLUSION**

In this article, we have surveyed augmented, virtual, and mixed reality from a cultural heritage perspective focusing on aspects such as tracking and registration, virtual environment modelling, presentation, tracking, and input devices, interaction interfaces, and systems. Moreover, we have categorised a number of CH-related augmented, virtual, and mixed reality applications into the general application areas of education, exhibition enhancement, exploration, reconstruction, and virtual museums. Also, we have discussed the technological requirements to support these areas. Though, the ultimate choice of enabling technology must depend on the experience that an application is intended to provide, we make the following suggestions as to which systems are more viable for a given purpose.

Even though augmented, virtual and mixed reality can all be used to achieve the above-mentioned purposes, our survey shows that augmented reality is preferable for exhibition enhancement. Similarly, virtual reality seems better for virtual museums, and mixed reality most viable for both indoor and outdoor reconstruction applications.

**REFERENCES**


Even though augmented, virtual and mixed reality can all be used to achieve the above-mentioned purposes, our survey shows that augmented reality is preferable for exhibition enhancement. Similarly, virtual reality seems better for virtual museums, and mixed reality most viable for both indoor and outdoor reconstruction applications.


Augmented, Virtual and Mixed Reality • 7:35


Filippo Stanco, Davide Tanasi, Giovanni Gallo, Matteo Buffa, and Beatrice Basile. 2012. Augmented perception of the past. the case of hellenistic syracuse. J. Multimedia 7 (2012), 211–216. DOI: http://dx.doi.org/10.4304/jmm.7.2.211-216


Received June 2017; revised September 2017; accepted September 2017
Lars Elleström*

A medium-centered model of communication

DOI 10.1515/sem-2016-0024

Abstract: The aim of this article is to form a new communication model, which is centered on the intermediate stage of communication, here called medium. The model is intended to be irreducible, to highlight the essential communication entities and their interrelations, and potentially to cover all conceivable kinds of communication of meaning. It is designed to clearly account for both verbal and nonverbal meaning, the different roles played by minds and bodies in communication, and the relation between presemiotic and semiotic media features. As a result, the model also pinpoints fundamental obstacles for communication located in media products themselves, and demonstrates how Shannon’s model of transmission of computable data can be incorporated in a model of human communication of meaning.

Keywords: communication model, intermediality, meaning, medium, multimodality, nonverbal communication

1 Aim

There are many communication models. For almost a century, researchers from various fields have tried to capture the essential factual and theoretical entities of human exchange of information and meaning. Many successful and highly useful attempts have been made to describe and analyze the basic features of communication, and several of the most influential scholarly conceptualizations of the entities present within communication have taken the form of models (see Lanigan 2013). Nevertheless, the principal aim of this article is to expound a new model of communication. Given the plethora of extant theories, why is yet another communication model called for?

First, a new model is necessary because, in reality, existing models only take into account verbal meaning, while ignoring nonverbal. Whereas it is

*Corresponding author: Lars Elleström, Department of Film and Literature, Faculty of Art and the Humanities, Linnaeus University, Växjö 35195, Sweden, E-mail: lars.ellestrom@lnu.se

© 2016 Elleström, published by De Gruyter. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 3.0 License.
common to acknowledge that communication includes not only speech and other verbal media types, these admissions are chiefly nominal and have had no far-reaching consequences for how communication has been theorized. I claim that it is deceptive to assume that the communication of nonverbal meaning can be modelled by way of analogue-to-verbal meaning.

I also argue that in order to remedy these lacks we need a new communication model that methodically identifies the irreducible components of communication, systematically puts them in relation to each other, and – most importantly – thoroughly develops the transitional stage of communication, which is often called the “channel,” “message,” “contact,” or, sometimes, “medium.” I will try to demonstrate that although several existing models capture most of the essential communication entities in highly useful ways, some core features of communication have yet to be properly incorporated. I maintain that these features are best conceptualized in terms of mediality and with the aid of semiotics, thereby offering theoretical tools that are vital for pinpointing basic media similarities and differences and developing the material and mental aspects of mediality to a sufficient degree of complexity.

The task of a communication model is undoubtedly to offer, as lucidly as possible, a theoretical framework for describing, explaining, and analyzing processes of communication. While this must include an understanding of not only how communication is conceivable at all, but why communication is often not possible to fully realize, there are vital sides of the latter aspect that have not yet been properly investigated. The technological notion of noise has offered only limited possibilities to explain communicative limitations beyond physical disturbances; imperative as the concept may be in engineering information data, it becomes rather trivial when applied to communication of meaning. On the other hand, a well-developed notion of medium that includes both presemiotic and semiotic media traits offers the possibility to explain both how various sorts of meaning can be communicated, and why these various sorts of meaning cannot always be realized. Therefore, basic media dissimilarities that are vital for differing communicative capacities must be mapped.

The aim of this article is thus to delineate a model of communication that is centered on the notion of medium. A model should be understood as a clearly outlined cognitive scheme that is both described with the aid of language, and depicted as a diagram. A medium should be understood in a broad way as the intermediate stage of communication; thus the term medium here refers not only to mass media, but also media used in more intimate communication; not only media based on external technological devices, but also media based on corporeality; not only premeditated media, but also casual media; not only media
used for practical purposes, but also artistic media – and so forth. Furthermore, media may be involved in both two-way communication, which some might consider as “true” communication, and one-way communication, which might also be termed “expression”; Werner Wolf, for instance, wrote about “media of expression or communication” (1999: 37).

Seen from a different scholarly perspective, the goal of the article is to put studies of intermediality, which have a focus on interrelations among dissimilar media, within the broad frame of human communication of meaning. I propose that the most relevant means by which to frame the notion of medium is to conceptualize it in terms of communication. In other words, I take the ideas of communication and medium to be interdependent and mutually explanatory.

The outline of the article is as follows. I first deliver a rudimentary critique of a handful of well-known communication models. After that, I suggest a new, medium-centered model of communication. On the basis of a methodical exposition of its entities and their interrelations, I elaborate on some vital implications and possibilities of the new conceptualization. The article concludes with a brief account of possible expansions of the expounded model, and some closing remarks on the model’s distinctive traits.

2 Critique of old communication models

I will now present, and very briefly discuss and criticize, the most fundamental features of four classic communication models in order to establish some standard conceptions, and hence a few points of departure for reconceptualized methods of theorizing about communication. As it is impossible to cover all communication models that have seen the light of day, I have chosen influential models by Shannon (1948), Jakobson (1960), Schramm (1971), and Hall (1980) from areas such as technology, linguistics, communication studies, and cultural studies – the ideas of which can also be traced in later models. While all of these models aim to represent the basic traits of communication, they were created for different purposes, and must hence be compared with some caution – in particular, there is some distance between the area of technology and the other fields. In any case, it is certainly not my aim to dismiss the models as such. I have a deep respect for the scholars who formed them and I will actually use most of their substance; however, I do believe that several aspects must be modified and clarified in order to create a truly comprehensive model of communication of meaning.
2.1 Shannon

In “A mathematical theory of communication” (1948), Claude Shannon distinctly declared that “[t]he fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem” (1948: 379). This clear account reveals that there is a fundamental discrepancy between the engineering problem and a meaning-centered approach to communication. Nevertheless, Shannon’s model (Figure 1) has also been very influential outside of technological circles, as it seems to provide a useful analogy between the engineering perspective and a humanities and social sciences viewpoint. This analogy is based on similar cognitive schemes that are structured around some sort of transmission from a source to a destination.

Although the mathematical substance of Shannon’s groundbreaking treatise will not be discussed here, he also offered useful explanations of the entities within his model. The *information source*, he stated, “produces a message or sequence of messages to be communicated to the receiving terminal” (1948: 380). It is vital to note that “message” here means quantifiable data of various kinds; it has nothing to do with meaning. Next, a *transmitter* is a technical device “which operates on the message in some way to produce a signal suitable for transmission over the channel” (1948: 381). The channel (represented by the tiny square in the middle of the visual diagram) “is merely the medium used to transmit the signal from transmitter to receiver. It may be a pair of wires, a coaxial cable, a band of radio frequencies, a beam of light, etc.” (1948: 381). Shannon furthermore explained that the *receiver* “ordinarily performs the

![Figure 1: Shannon’s communication model (1948: 381).](image-url)
inverse operation of that done by the transmitter, reconstructing the message from the signal” so that the message can reach its destination, which can be understood as a person or a thing (1948: 381).

Later in the article, Shannon elucidated the crucial notion of noise – a purely technical phenomenon, which is present when “the received signal is not necessarily the same as that sent out by the transmitter” (1948: 406). Nevertheless, this aspect of technological communication has also been found useful for later communication models focused on meaning.

Just one year after the initial publication of Shannon’s treatise, Warren Weaver (1998 [1949]) brought out a paper – together with a reissue of Shannon’s article and which, to a large extent, built on Shannon’s ideas – in which he discussed a broad definition of communication as “all the procedures by which one mind may affect another,” including “not only written and oral speech, but also music, the pictorial arts, the theatre, the ballet, and in fact all human behavior” (1998 [1949]: 3). While this excellent delineation of a broad notion of communication perfectly fits my purpose, Weaver unfortunately did not develop his comprehensive perspective at all, instead restricting his extension of Shannon’s model mainly to speech. Later in this article I will come back to some problems caused by the adaptation of an engineering model of communication to the humanities and social sciences.

### 2.2 Jakobson

Roman Jakobson’s communication model (Figure 2), presented in the article “Closing statement: Linguistics and poetics” (1960), crosses the border between linguistics and literary studies. His aim was to investigate language “in all the variety of its functions”; an exploration that “demands a concise survey of the constitutive factors in any speech event, in any act of verbal communication” (1960: 353).

```
ADDRESSER  MESSAGE  ADDRESSEE

CONTEXT
CONTACT
CODE
```

*Figure 2: Jakobson’s communication model (1960: 353).*

Jakobson explained the factors in this succinct way:

The ADDRESSER sends a MESSAGE to the ADDRESSEE. To be operative the message requires a CONTEXT [that is] seizable by the addressee, and either verbal or capable of being
verbalized; a Code fully, or at least partially, common to the addressee and addressee (or in other words, to the encoder and decoder of the message); and, finally, a Contact, a physical channel and psychological connection between the addressee and the addressee, enabling both of them to enter and stay in communication. (Jakobson 1960: 353)

These constitutive factors are arranged in a scheme that, like Shannon’s and all other models discussed in this article, presupposes a movement from left to right. It is clear that the outer ends of Shannon’s and Jakobson’s models seem to roughly correspond to one another. Jakobson’s notion of contact can also be understood to approximately correspond to Shannon’s notion of channel (the contact was described by Jakobson as “a physical channel” [1960: 353]). Furthermore, both researchers used the term message to refer to entities that may seem to have much in common. However, on closer inspection the differences appear to dominate: in Shannon’s model, the message consists of quantifiable data which form an entity that is transmitted; in Jakobson’s model the message appears to be an intermediate (meaningful) entity that is in close relation to the contact. Jakobson’s two factors “context” and “code” are unmistakably linked to meaning production, and do not have any equivalences in Shannon’s model.

Jakobson clearly declared that his research object is language, which makes his investigation less useful for the construction of a broader communication model. Another problem is that his article contains no serious efforts to define or discuss the six constitutive factors of his scheme at any length. Nevertheless, Jakobson’s model stresses elements that are imperative for communication of meaning, and has been prominent within its research fields.

2.3 Schramm

In “The nature of communication between humans” (1971, earlier version 1954), communication scholar Wilbur Schramm suggested that “the communication process consists of information-processing organized around a shared orientation to certain signs” (1971: 22). His point of departure was that “the communication relationship includes three elements” (1971: 15), which together form an ascetic model consisting of the communicator (A), the message (m), and the receiver (B; Figure 3).

![Figure 3: Schramm’s communication model (1971: 23).](image-url)
While Schramm discussed many things in his oft-quoted article, these three elements are central. Interestingly, he argued that “nothing really passes from A to B”; A, the communicator (which must be understood as a person) “encodes a message as best he can in signs,” and then B, the receiver (another person) “reads a message into those signs” (1971: 22). The merit of this approach is that the interpretive activity of the receiver is clearly demarcated. Schramm stated that “it is just as meaningful to say that B acts on the signs, as that they act on B” (1971: 22) – this is why the arrow between m and B points in both directions.

To my mind, the main problem with Schramm’s model is that he (like Jakobson) did not really scrutinize the notion of message. He clearly tied it to the notion of signs, which is a step towards a more complex understanding of the in-between element of communication. At the same time, he almost seems to substitute “message” with “signs” when pinpointing the intermediate element: whereas the “message” is actually circumscribed as two messages (the communicator “encodes a message,” the receiver “reads a message,” and “nothing really passes from A to B”), the “signs” appear to be a more stable intermediate unit according to Schramm’s explanation.

In spite of its merits, Schramm’s communication model becomes rather messy, I think. On the one hand, he explicitly argued that nothing passes from communicator to receiver, and furthermore implied that the receiver’s message is not the same as the communicator’s, although they share the same signs; on the other hand, he stated that “it is messages, not ideas or thoughts, that pass from communicator to receiver” (1971: 9). Consequently, the message is confusingly understood as both something that is situated between communicator and receiver, and something that passes from communicator to receiver. It is my impression that Schramm conflates two entities that should be kept separate in a lucid communication model.

### 2.4 Hall

My last example of influential communication models is that presented by Stuart Hall in “Encoding/decoding” (1980, earlier version 1973); a model that was designed to be used primarily within cultural studies. Hall argued, convincingly I think, that there are a few “determinate moments” in the complex communications system: the moments of “encoding” and “decoding” and the “message form” (1980: 129). Although these three determinate moments are part of “the social relations of the communication process as a whole” (1980: 129) they must necessarily be highlighted, as Hall’s visual diagram of communication (Figure 4) makes clear.
As the diagram shows, the moments of encoding and decoding (notions that we recognize from both Jakobson and Schramm) are also called “meaning structures” 1 and 2. Remarkably, Hall avoided putting entities such as communicator and receiver at the end-points of the diagram, which makes his model special and productive; it is suggested that communication both starts and ends with (non-identical) meaning structures.

The drawback of Hall’s way of modelling communication is that he used an abundance of terms to suggest, rather than specify, the contours of his central notions. The central entity in his model – the third determinate moment – is called “programme as ‘meaningful’ discourse” in the diagram (program should be understood in the sense of television program – television being his main example of communication), and “‘meaningful’ discourse” is only one out of many terms that seem to aim at the same target. Hall wrote about the central entity as the “objects” of communication: “meanings and messages in the form of sign-vehicles” that are organized “through the operation of codes within the syntagmatic chain of a discourse” (1980: 128). Further on in his text, one finds wordings such as “symbolic vehicles constituted within the rules of ‘language’” (1980: 128); “message form” (1980: 129); “encoded messages in the form of a meaningful discourse” (1980: 130); and “discursive form” (1980: 131).

Thus, without being very precise Hall circumscribed semiotic issues with the aid of terms such as sign-vehicle, code, symbolic vehicles, and, later on, iconic signs. I judge this to be a decisive move in the right direction if the aim is to create a broad communication model that is suited to all kinds of communication, including nonverbal. However, Hall’s complex account of communication entails some problems. He actually used the term message in relation to all three determinate moments (1980: 130), which makes his text somewhat obscure. He furthermore focused on television, and yet emphasized that communication is
inscribed into “the discursive rules of language” (1980: 130), and constantly used the terms code and reading of television. Although these terms, together with terms such as text and language, were commonly used during the 1970s and 1980s to denote concepts other than verbal sign-systems, the conceptual framing of Hall’s arguments is clearly primarily linguistic.

In opposition to such an account, I argue that meaning is about much more than language and symbolic codes based on conventions (for more detailed and wide-ranging critical discussions of different notions of code, see Cobley 2013). Signs of similarity (icons) and signs of contiguity (indexes) also create meaning. I agree with Hall only partly when he states, referring to Umberto Eco, that “[i]conic signs are ... coded signs too – even if the codes here work differently from those of other signs” (1980: 131–132). I believe he is right to the extent that there are probably no “pure” icons that are not mixed with symbols in some way; nevertheless, the core of iconic meaning-production (similarity between sign and object) fundamentally differs from the core of symbolic meaning-production (habitual or conventional connection of sign and object). Therefore, iconic (and indexical) meaning-production cannot accurately be understood in terms of codes; all three sign types – icons, indexes, and symbols – must be appreciated as profoundly different ways of creating meaning (also when mixed), and cannot be subsumed under each other.

3 A medium-centered model of communication

I will now suggest a new means of modelling communication of what is traditionally called meaning. My proposed model consists of what I take to be the smallest and fewest possible entities of communication and their essential interrelations. If one of these entities or interrelations is removed, communication is no longer possible; thus, the model is irreducible. I submit that three indispensable and interconnected entities can be discerned:

1. Something being transferred.
2. Two separate places between which the transfer occurs.
3. An intermediate stage that makes the transfer possible.

3.1 The three entities of communication in earlier models

The first entity, “something being transferred” was referred to as the message by Shannon. However, as Shannon’s message is clearly defined as computable data, it cannot be equated to what is supposed to be transferred during the
communication of meaning. Jakobson also used the term *message* to capture the transferred entity, but did not delineate the notion underlying his term. Schramm vacillated between two incompatible arguments: that there is no such thing as an entity being transferred, and that the transferred entity is a “message” – not ideas or thoughts. In another publication, however, he stated that a “concept” must be reproduced in order to obtain communication (Schramm 1955: 133). Hall was also rather vague when he implied that “meaning” is transferred in communication. Instead of clearly stating that communication is about transferring meaning, he emphasized that “meaning structures 1” and “meaning structures 2” may differ; there are degrees of “symmetry” and degrees of “understanding” and “misunderstanding” (1980: 131). In other words: if there is transfer of meaning in communication, this involves transformation of meaning. Hall’s contention is certainly feasible.

While the second entity, “two separate places between which the transfer occurs,” arguably consists of two units, they can only be outlined in relation to each other. Shannon termed them “information source” and “destination,” pointing out that the destination may be either a thing or a person; it might thus be inferred that the source may also be understood as either of these things. However, given that Shannon’s model deals with the transfer of computable data, it is difficult to see how the destination can be a person; in order to deal with the activity of human beings in Shannon’s model, stages that connect computable, physical data and mental significance must be added at the outer ends of it. In order to really reach *into* a person, the raw data of the transferred (so-called) message must be perceived and interpreted. On the other hand, for Jakobson and Schramm, who both reasoned (albeit slightly hesitantly) in terms of messages consisting of meaning rather than computable data, it is less problematic to circumscribe the two places as persons. Whereas Jakobson’s terms are *addressee* and *addressee*, Schramm prefers *communicator* and *receiver*. Finally, Hall avoided outlining the two separate places between which the transfer occurs as persons; in fact, he avoided pointing to such places at all. However, his notion that “meaning structures” are to some extent transferred implies that such meaning structures indeed need to be located at places that are capable of holding “meaning” – which must be understood as the minds of human beings.

The third entity, “an intermediate stage that makes the transfer possible,” might seem to correspond to Shannon’s “channel,” or what he (only once) called “medium.” Shannon’s channel is a material unit, such as a pair of wires or a band of radio frequencies, which are capable of transferring computable data. However, although such units may certainly be involved in the transfer of meaning, they are not the central entities that connect two places in which
meaning can be located. In Shannon’s own words, the channel is merely a medium for transmission of a signal, so I submit that it is not a crucial point of transition. Again, the other three communication models, which were designed to address mental significance, offer better ways of understanding this intermediate stage of meaning communication. Jakobson’s “contact” notably incorporates both a material and a mental aspect; it was described as “a physical channel and psychological connection between the addressee and the addressee” (1960: 353). As noted earlier, Schramm used the term message to represent not only the transferred entity, but also the intermediate stage of communication (the message seems to be understood as something that is both “transferred” and “transferred through”). Importantly, however, Schramm described the transmitting message not only as a material entity – such as “a letter” – but also as “a collection of signs,” thus indicating the capacity of the material to produce mental significance through signs (1971: 15). Hall also emphasized the semiotic nature of the intermediate stage of communication. His term for this entity was ‘meaningful discourse’; however, as noted above, his terminology is generally rather incoherent, resulting in uncertainty about the more precise nature of the intermediate stage.

3.2 The three entities of communication in a medium-centered model

Regarding the first entity, “something being transferred,” there is certainly a point in Schramm’s notion that no ideas or thoughts are transferred in communication. As Hall indicated, transfer of meaning is very likely to entail change of meaning; this modification may be only slight, or more radical. Nevertheless, I claim that communication models cannot do without the notion of something being transferred. If there is no correlation at all between input and output there is simply no communication, given the foundational idea that to communicate is “to share”; thus, a concept of communication without the notion of something being transferred is actually nonsensical. However problematic it may be, the notion of something being transferred must be retained and painstakingly scrutinized, instead of being avoided. Although this is not the place for such an in-depth examination, I think it is clear that the transferred units or features certainly cannot be confined to distinct and consciously intended conceptions, and perhaps not even to “ideas” as Schramm understands them.

My suggestion is to use the term cognitive import to refer to those mental configurations that are the input and output of communication. The notion that I want to suggest using this term is clearly closely related to notions captured by
terms such as *meaning* and *ideas*, although the term *cognitive import* is perhaps less burdened with certain notions that a term such as *meaning* seems to have difficulties getting rid of; meaning is often understood to be a rather rigid concept of verbal, firm, definable, or even logical sense. Cognitive import should be understood as a very broad notion of meaning that is relevant to a wide range of media types. It is imperative to emphasize that although cognitive import is always a result of mind-work, it is not always possible to articulate using language; hence, communication, according to my proposed model, cannot be reduced to communication of verbal or verbalizable meaning.

The second entity, “two separate places between which the transfer occurs,” is most often construed as two persons. However, this straightforward notion is not precise enough for my purposes. As it is imperative to be able to connect mind and body to different entities of the communication model, one must avoid crude notions such as that of Jakobson’s addressee–addressee and Schramm’s communicator–receiver. These notions give the impression that the transfer necessarily occurs between two persons consisting of minds and bodies and with a third, separate, intermediate object in the middle, so to speak; an intermediate object in the form of a “message” that is essentially disconnected from the communicating persons. It is better to follow Hall’s implicit idea that communication occurs between sites that are capable of holding “meaning.” Weaver’s description of communication as something that occurs between “one mind” and “another” is simple and right to the point.

My suggestion is to use the terms *producer’s mind* and *perceiver’s mind* to refer to the mental places in which cognitive import appears; first, there are certain mental configurations in the producer’s mind and then, following the communicative transfer, there are mental configurations in the perceiver’s mind that are at least remotely similar to those in the producer’s mind.

Most researchers referred to in this article have either explicitly or implicitly recognized that the third entity, “an intermediate stage that makes the transfer possible,” is in some way material. As stated succinctly in a recent publication, any act of communication “is made possible by some form of *concrete reification* of the message, which, at its most elementary level, must abide by physical laws to exist and take shape” (Bolchini and Lu 2013: 398). Furthermore, Schramm and Hall clearly discussed the intermediate stage in terms of signs. In line with this, I suggest that the intermediate entity connecting two minds with each other is always in some way material, although it clearly cannot be conceptualized only in terms of materiality. As it connects two minds in terms of a transfer of cognitive import, it must be understood as materiality having the capacity to trigger certain mental responses.
My suggestion is to use the term *media product* to refer to the intermediate stage that enables the transfer of cognitive import from a producer’s to a perceiver’s mind. As the bodies of these two minds may well be used as instruments for the transfer of cognitive import, they are potential mediators of media products. I propose that a media product may be realized by either non-bodily or bodily matter (including matter emanating directly from a body), or a combination of these. This means that the producer’s mind may, for instance, use either non-bodily matter (say, a written letter) or her own body and its immediate extensions (speech and gestures) to realize media products. Furthermore, the perceiver’s body may be used to mediate media products; for instance, the producer may realize a painting on the perceiver’s skin or push her gently to communicate the desire that she moves a bit. In contrast to influential scholars such as Marshall McLuhan who conceptualize media as the “extensions of man” in general (McLuhan 1994 [1964]), I thus define media products as “extensions of mind” in the context of communication. As being a media product should be understood as a function rather than an essential property, virtually any material entity can be used as one.

### 3.3 A medium-centered model

An advantage of using the term *medium* is that its etymology leads to rather neutral basic notions, such as middle and interspace; a possible disadvantage is that it has actually been used to refer to many different entities and phenomena in various research areas. It has certainly been applied before in communication models, but to the best of my knowledge it has not yet been tied to a developed concept that takes into account all those fundamental qualities that are vital for a proper understanding of the intermediate stage of communication. In fact, no such concept has been presented to date (see, for instance, the overview in Crowley 2013).

Before attempting to present the contours of this concept, I will display my newly developed communication model in the form of a visual diagram (Figure 5) and elaborate on some of its implications. Construing this diagram from left to right, the act of communication starts with certain cognitive import in the producer’s mind. Consciously or unconsciously, the producer forms a media product, which may be taken in by some perceiver. The media product makes possible a transfer of cognitive import from the producer’s mind to the perceiver’s; this is a transfer not in the strong sense that the cognitive import as such passes through the media product (which lacks consciousness), but in the sense that there is, in the end, cognitive import in the perceiver’s mind that bears some resemblance to the cognitive import in the producer’s mind.
3.4 The four interrelations of communication entities in a medium-centered model

I will now elaborate on these interrelations, especially the fourth one. The notion of media product, which was only rudimentarily demarcated above, and the question how cognitive import may be transferred through a media product, are essential for any attempt to understand the core of communication.

The first interrelation, “an act of production ‘between’ the producer’s mind and media product,” is always initiated by the producer’s mind and always, to begin with, effectuated by the producer’s body. Sometimes this primary bodily act immediately results in a media product, for instance, when one person begins talking to another person who is standing beside her: the speech emanating from the vocal chords constitute a media product that reaches the perceiver directly. At other times the primary bodily act is linked to subsequent stages of production, and it is not unusual for the primary bodily act to be connected to a broad range of actions and procedures before a media product comes to be present for a perceiver. For instance, talking through a

---

The visual diagram contains the three entities of communication circumscribed above:
1. Something being transferred: cognitive import.
2. Two separate places between which the transfer occurs: producer’s mind and perceiver’s mind.
3. An intermediate stage that makes the transfer possible: media product.

Additionally, it displays four essential interrelations among these entities:
1. An act of production “between” the producer’s mind and media product.
2. An act of perception “between” the media product and the perceiver’s mind.
3. Cognitive import “inside” the producer’s mind and the perceiver’s mind.
4. A transfer of cognitive import “through” the media product.
telephone often requires manual handling of the telephone in addition to the activation of the user’s vocal chords, and always requires constructed, technological devices that are suitable to transmit the initial speech to another place, in which the actual media product is constituted – that is, the speech that can be heard by the perceiver. Similarly, a child drawing a picture for her father who is sitting at the same kitchen table only has to perform, in principle, one primary bodily act in order to create a media product that is immediately available for the perceiver. However, if the father is in another place, additional stages of actions and procedures must be added: the drawing may be posted and physically relocated, or scanned and emailed, after which it appears in a slightly transformed way as a media product that is mediated by a computer screen. The act of production may thus be simple and direct, as well as complex and indirect. It may furthermore include stages of storage.

The second interrelation, “an act of perception ‘between’ the media product and the perceiver’s mind,” is always initiated by the perceiver’s sense organs and always, to some extent, followed by and entangled with interpretation. Interpretation should be understood as all kinds of mental activities that somehow make sense of the sensory input; these activities may be both conscious and unconscious, and are no doubt already present in a basic way when the sense impressions are initially processed. Thus, compared to the potentially extensive act of production, the act of perception is brief and very quickly channeled into interpretation, which of course occurs in the perceiver’s mind. Nevertheless, the type, quality, and form of sensory input provided by the media product, and actually taken in by the perceiver’s sense organs, are absolutely crucial for the interpretation formed by the perceiver’s mind.

A comparison of meaning-oriented communication models, such as the one that I propose here, and Shannon’s model of transmission of quantifiable data, may give the impression that the acts of production and perception are equivalent to the entities and processes that Shannon calls “transmitter”/“signal” and “received signal”/“receiver.” However, the production and perception of media products involve complex and entangled corporeal and cognitive processes that cannot properly be reduced to analogies of mechanical processes. Therefore, Weaver’s suggested extension of Shannon’s model to the area of meaning only works on a rather superficial level. Whereas in an engineering context it may well be true that “The receiver is a sort of inverse transmitter, changing the transmitted signal back into a message” (Weaver 1998 [1949]: 7), the act of perception cannot accurately be described as an inverse act of production, unless one is satisfied with a very rudimentary analogy. The creation of cognitive import, or meaning, requires much more than an unbiased perception of the raw material qualities of the media product; it is not a
process of reproduction of computable data, but rather one of interpreting perceived sensory configurations. Although this problem is not entirely neglected by Weaver, he did not scrutinize it at any length. Furthermore, leaving the claims of Weaver aside, it must be noted that the act of production often actually involves the producer’s acts of perception and interpretation of the emerging media product, which brings the two acts closer to each other. Rather than saying that the act of perception is an inversion of the act of production, which makes the relation between the two acts resemble a purely material transmission, one must say that the act of production and the act of perception are related in an act of material and mental transmission, including transition.

Even if one notes that there is a shallow analogy between Shannon’s model of communication of computable data and meaning-oriented communication models, it is more useful to observe that Shannon’s model might actually be incorporated as part of the act of production in my proposed model. Whereas some acts of production do not require external technological devices, others do. As the elementary examples of person-to-person speech and telephone calls demonstrated, it may be the case that a primary bodily act immediately results in a media product, while at other times the primary bodily act is linked to subsequent stages of production involving external technology which ultimately result in a media product that is presented to the perceiver. Needless to say, Shannon’s model and mathematical theory has proven exceptionally useful for the engineering of the non-corporeal stages of production, while it is largely irrelevant for the corporeal and mental stages of production.

Above, I refuted the idea that the “destination” in Shannon’s model can be understood as a person, and that an act of perception must be added if his model is to be applicable to communication of cognitive import. Consequently, the solution is to understand the entities of Shannon’s model as potential parts of the act of production in my suggested model, with “destination” seen not as a person but as a thing: the media product. This material entity must clearly be perceived and interpreted by the perceiver’s mind if communication of cognitive import is to be achieved.

The third interrelation among the entities of communication, “cognitive import ‘inside’ the producer’s mind and the perceiver’s mind” will only be briefly commented upon here. Clearly, one cannot state, without intricate implications, that there is a certain amount of confinable cognitive import inside a mind, and it is undoubtedly difficult to judge the actual extent of similarity between the two amounts of confined cognitive import in the two minds. Deciding this in a more precise way is probably beyond the reach of known research methods. However, I find the notion that the transferred
cognitive import is only one part of the producer’s and the perceiver’s minds unproblematic; the cognitive import is “inside” the minds in the respect that it is closely interconnected with a multitude of other cognitive entities and processes and, in the end, with the total sum of mental activities in general that surrounds it.

The fourth interrelation, “a transfer of cognitive import ‘through’ the media product,” is central for my arguments in this article. So far, the media product has simply been described as the entity of communication that enables a transfer of cognitive import from a producer’s mind to a perceiver’s mind; a material entity that has the capacity of triggering mental response. In order to give a somewhat more detailed account of this notion, the very capacity itself must be scrutinized.

Of course, the transfer of cognitive import is only partly comparable to other transfers – such as, for instance, the transfer of goods between two cities by train. The cognitive import transfer is not a material transfer, but a mental transfer with the aid of materiality. In one respect it can be compared to teleportation, the transfer of energy or matter between two points without traversing the intermediate space: the cognitive import is indeed transferred between two points (two minds), and, contrary to the transfer of goods, it does not traverse the intermediate space. Nevertheless, as the transfer depends on the media product it may reasonably be said to go “through” it. The media product is actually neither a neutral object of material transfer, like a freight car, nor an intermediate space without effect, as in teleportation; it constitutes a crucial stage of transition, not only transmission.

At this stage of the account it is necessary to introduce a distinction between media products and what may be called technical media or, to be more precise, technical media of distribution of sensory configurations. I have emphasized that media products are material entities; however, media products need technical media in order to actually be realized. Technical media are material devices, either simply present in the producer’s mind’s environment or more or less crafted, that cause media products to physically manifest in the world. They are entities that have the capacity to display media products and make them available for the senses of the perceiver; they distribute sensory configurations (Elleström 2010: 30–37; Elleström 2014b: 47–56). For instance, Harold A. Innis (1950) has emphasized the importance of technical media such as stone, clay, papyrus, and paper for the historical development of communication – more specifically writing – and society at large. More modern technical media include electronic screens, and sound waves produced by loudspeakers. All these different kinds of physical entities are necessary conditions for making media products discernible.
The distinction between media product and technical medium is clearly theoretical, rather than being a distinction between two different kinds of material entities. On the contrary, the technical medium is a prerequisite for the physical existence of a media product, and in a communicative situation the perceiver identifies only one object or phenomenon. The distinction is needed in order to demonstrate the difference—and mutual interdependence—between, for instance, a television program (a media product) and a television set (a technical medium). A technical medium such as a television set (which actually consists of two kinds of technical media: a flat screen that emits photons and loudspeakers that set the air into pulsation) may realize several different media products (many television programs). Conversely, a media product such as a television program can be realized by many types of technical media (not only television sets but also, for instance, computers, which also consist of a screen and loudspeakers). Sometimes the media product and the technical medium are virtually inseparable, as in the case of an oil painting.

Above, I suggested that media products may be realized by either non-bodily or bodily matter, or any mixture of these. In other words, there are external technical media (non-bodily devices such as screens and ink on paper) and there are internal technical media (bodies, parts of bodies, or physical phenomena emanating directly from bodies). I furthermore suggest that all media products can be analyzed in terms of four kinds of basic traits, which might be called media modalities (Elleström 2010).

Three of these modalities are presemiotic, which means that they cover media traits that are involved in signification—the creation of cognitive import in the perceiver’s mind—although they are not semiotic qualities in themselves. The presemiotic traits concern the fundamentals of mediation, which is to say that they are necessary conditions for any media product to be realized in the outer world by a technical medium, and hence for any communication to be brought about.

The three presemiotic modalities are the material modality, the spatiotemporal modality, and the sensorial modality. Media products are all material in the plain sense that they may be, for instance, solid or non-solid, or organic or inorganic, and comparable traits like these belong to the material modality. It is also the case that all media products have spatiotemporal traits, which means that such products that do not have at least either spatial or temporal extension are inconceivable; hence, the spatiotemporal modality consists of comparable media traits such as temporality, stasis, or spatiality. Furthermore, media products must reach the mind through at least one sense; hence, sensory perception is the common denominator of the media traits belonging to the sensorial modality—media products may be visual, auditory, tactile, and so forth.
Of course, these kinds of traits are not unknown to communication researchers. Hall (1980), for instance, discussed the two sensory channels of television; David K. Berlo (1960) highlighted all five external senses; and Schramm at least briefly mentioned that “a message has dimensions in time or space” (1971: 32). However, thorough understanding of the conditions for mediation requires systematic attention to all three presemiotic modalities. It is clear that cognitive import of any sort cannot be freely mediated by any kinds of material, spatiotemporal, and sensorial traits. For instance – to take some blatant examples – complex assertions cannot easily be transferred through the sense of smell, and it is more difficult to effectively transfer detailed series of visual events through a static media product than through a temporal media product.

The fourth modality is the semiotic modality, which covers media traits concerning representation rather than mediation. Whereas the semiotic traits of a media product are less palpable than the presemiotic ones, and in fact are entirely derived from them, they are equally essential for realizing communication. The mediated sensory configurations of a media product do not transfer any cognitive import until the perceiver’s mind comprehends them as signs. In other words: the perceived sense data are meaningless until that are understood to represent something through unconscious or conscious interpretation. This is to say that all objects and phenomena that act as media products have semiotic traits by definition. By far the most successful effort to define the basic ways in which to create meaning in terms of signs is Charles Sanders Peirce’s foundational trichotomy icon, index, and symbol.

In brief, Peirce held that signs (often called *representamens*) stand for objects – a relationship that results in interpretants in the perceiver’s mind (CP 2.228 [c.1897]). This is a mental process, although both representamens and objects may be (or be connected to) external elements or phenomena; however, the interpretant is always in the mind. My notion of cognitive import created in the perceiver’s mind corresponds with Peirce’s notion of interpretant.

Hence, the media product can be understood as an assemblage of representamens that, due to their qualities (material, spatiotemporal, and sensorial traits), represent certain objects (that are available to the perceiver), thus creating interpretants (cognitive import) in the perceiver’s mind. Peirce’s three basic sign types are defined on the basis of the representamen–object relationship and can be understood as fundamental cognitive abilities. Icons stand for (represent) their objects based on similarity; indexes do so based on contiguity; and symbols rely on habits or conventions (CP 2.247–2.249 [c.1903]; Elleström 2014a: 98–113). I take iconicity, indexicality, and symbolicity to be the main media traits within the semiotic modality, which is to say that no communication occurs unless cognitive import is created through at least one of the three sign types (icons, indexes, and symbols).
Again, semiotics is certainly not unknown in communication research. Among the scholars quoted in this article, Schramm clearly related to some basic semiotic features. For instance, he accurately noted that “it is just as meaningful to say that B [the receiver] acts on the signs [the message], as that they act on B” (1971: 22). Indeed, the mind of the perceiver is very active in construing the signs of the media product. In addition, Hall spoke in terms of semiotics, although with a distinct linguistic bias, which I have already criticized. On the whole, I judge that semiotic approaches to communication based on the tradition of Ferdinand de Saussure, which downplay the role of iconicity and indexicality, have been harmful to the development of theory that also embraces non-verbal communication. Peirce’s semiotic framework is much more fruitful as it incorporates sign types that work far outside of the linguistic domain.

Furthermore, my emphasis here is on the notion that a semiotic perspective must be combined with a presemiotic perspective. Communication is equally dependent on the presemiotic media modalities and the semiotic modality. What we take to be represented objects called forth by representamens, or signs (objects such as persons, things, events, actions, feelings, ideas, desires, conditions, and narratives), are results of both the basic features of the media product as such (the mediated material, spatiotemporal, and sensorial traits) and of cognitive activity (resulting in representation). While signification is ultimately about mind-work, in the case of communication this mind-work is fundamentally dependent on the physical appearance of the media product – although some representation is clearly more closely tied to the appearance of the medium, whereas other is more a result of interpretation, and hence the context of the perceiving mind.

As with presemiotic traits, the semiotic traits of a media product offer certain possibilities and set some restrictions. Obviously, cognitive import of any sort cannot be freely created on the basis of just any sign type. For instance, the iconic signs of music can represent complex feelings and motional structures that are largely inaccessible to the symbolic signs of written text; conversely, written symbolic signs can represent arguments, and the appearance of visual objects, with much greater accuracy compared to auditory icons. Flagrant examples like these are only the tip of the iceberg in terms of the (in)capacities of signs based on similarity, contiguity, and habits or conventions, respectively. Therefore, communicative transfer of cognitive import through a media product is made possible – but also fundamentally limited – by the semiotic traits of the medium.

In line with this proposal, it is appropriate to bring the notion of noise back into the discussion. Shannon’s idea that signal disturbances in communication can be conceptualized as noise has been picked up by many researchers engaged in communication of meaning. The basic phenomenon of disruptions
that occur on the way from the producer’s to the perceiver’s mind is clearly relevant to the transfer of cognitive import. Schramm, for instance, noted that noise is “anything in the channel other than what the communicator puts there” (1955: 138). For instance, speech can be disturbed by other sounds and a motion picture can be disrupted because of material decay or censorship. Noise in this sense occurs in both the act of production and the act of perception. In my visual model of communication (Figure 5), this noise is shown as disruptions in the arrow representing transfer of cognitive import – both before and after the transfer through the media product – reflecting the unsatisfactory conditions of production and perception.

The problem with the notion of noise when applied to communication of meaning is that it might imply that complete absence of noise would bring about complete transfer of cognitive import – as in the case of technical transmission of computable data – which is clearly not the case. The technological notion of noise is simply not sufficient to understand communication of cognitive import. According to Hall, “distortions” or “misunderstandings” are also due to, among other things, “the asymmetry between the codes of ‘source’ and ‘receiver’ at the moment of transformation into and out of the discursive form” (1980: 131).

This is definitely a step in the right direction in terms of offering a more complex notion of possible disruptions in the communication of cognitive import. However, it does not provide a more complete view of restraining factors in the transfer of cognitive import. It must also be emphasized that creators of media products generally do not have access to, or do not master, more than a few media types. Consequently, they often cannot possibly form media products that have the capacity to create cognitive import in the perceiver’s mind that is similar to the cognitive import in their own mind. Therefore, I argue that perhaps the most important restraining factors of communication are to be found in the basic presemiotic and semiotic traits of the media products.

Many exceedingly complex factors are clearly involved when the perceiver’s mind forms cognitive import. My proposed model highlights one cluster of crucial factors in particular: media products have partly similar and partly dissimilar material, spatiotemporal, sensorial, and even semiotic traits, and the combination of traits to a large extent – although certainly not completely – determines what kinds of cognitive import can be transferred from the producer’s mind to the perceiver’s mind. Songs, emails, photographs, gestures, films, and advertisements differ in various ways concerning their presemiotic and semiotic traits, and hence can only transfer the same sort of cognitive import to a limited extent. In my diagram (Figure 5), this communicative restriction is shown as disruptions in the arrow representing transfer of cognitive import as it passes through the media product.
4 Possible expansions of the model

Many important aspects of communication could and should be added to my suggested model, which outlines only the smallest and fewest possible entities of communication and their essential interrelations. While I believe that the model is irreducible, it is certainly expandable. Below, I will briefly comment upon some of the most urgent possible developments.

4.1 The formation of the producer’s and the perceiver’s minds

The most important issue that has not been properly addressed is the question of how the producer’s and the perceiver’s minds are formed or constituted by surrounding factors. In addition to its innate basic capacity to perceive and interpret mediated qualities, the mind is inclined to form cognitive import on the basis of acquired knowledge, experiences, beliefs, expectations, preferences, and values – preconceptions that are largely shaped by culture, society, geography, and history. It is clear that this concept is immensely important for the outcome of communication. The perceiver’s mind acts upon the perceived media product on the basis of both its hardwired cognitive capacities and its attained predispositions; evidently, the cognitive import that was stored in the mind before the media product was perceived has a significant effect – to various degrees – on the new cognitive import formed by communication.

As this is a widely recognized fact that has been extensively theorized in various ways, I have bracketed it in my outline of the new model. Others, such as Jakobson and Hall, have given it more attention. Jakobson discussed it in terms of “a CONTEXT [that is] seizable by the addressee, and either verbal or capable of being verbalized” (1960: 353). Context is no doubt important for all kinds of communication, although I think it is a mistake – even for a restricted focus on verbal communication – to say that the context must be verbalizable in order to be relevant. Hall distinctly emphasized the “social relations of the communication process as a whole” and the “frameworks of knowledge” (1980: 129–130), and discussed them in detail. These and other issues that are central to the formation of meaning in a broad context have been minutely scrutinized within the research area of hermeneutics.

4.2 Perceiver’s mind becomes producer’s mind

Another way of developing the model is to highlight the fact that in real communicative situations the perceiver’s mind is very often also a producer’s
mind. On the basis of the cognitive import generated by an initial media product, the perceiver becomes a producer in creating another media product (of the same or another kind) that reaches another perceiver’s mind, thereby forming new cognitive import that is more or less similar to that in earlier producers’ minds. Hence, a communicative chain is formed.

4.3 Perceiver’s mind becomes producer’s mind and producer’s mind becomes perceiver’s mind

When the communicative chain involves only the initial producer and perceiver constantly changing roles and forming new media products (of the same or another kind), we have two-way communication. The creation of new media products in two-way communication is often conceptualized as feedback, which may result in the creation of cognitive import that is either rather constant or significantly developed.

4.4 Producer’s and/or perceiver’s mind are actually several minds

It is furthermore often the case that media products are either produced or perceived, or both, by several minds. For instance, a film is normally both produced and perceived by more than one mind. A plenary talk is produced by one mind but perceived by many. An unsuccessful theater performance may be produced by many minds but perceived (from the right position) by only one.

4.5 Producer’s and perceiver’s mind are one and the same

Finally, it may be the case that the perceiver takes in her own media product. Although I would not say that pure thinking is communication (as suggested by Berlo 1960: 31), perception of one’s own media product created earlier may mean that the mind tries to construe cognitive import on the basis of the media product, rather than on the memory of what one had in mind on the occasion of production. In this case, a transfer of cognitive import actually occurs through a media product from one mind to another, in the sense that the mind, when perceiving the media product, is in a different state compared to that during production. The effort of writing a scholarly text is a good example of this sort of internal communication: sometimes, alas, communication fails when one cannot understand one’s own words written the day before.
5 Conclusions

The irreducible model of communication presented in this article, with its clearly defined entities and interrelations among entities, has several advantages. It avoids the widespread but notoriously indistinct notion of message that generally conflates communicative entities. As a model of the transfer of cognitive import from one mind to another through a media product, it allows for a refined pinpointing of the minds and bodies of the communicators in separate parts of the model (which does not in any way rule out the important fact that mind and body are profoundly interrelated). Its features also make it possible to incorporate Shannon’s model of the transmission of quantifiable data as a possible part of the act of production, instead of using it as a vague analogy for human communication. Furthermore, by avoiding the notion of code as a collective formula for meaning creation, it has the capacity to deal with all sorts of cognitive import, not only verbal or other symbolic significance.

The heart of the proposed model consists of the media product. Although the true complexity of this intermediate entity has only been hinted at, some rudimentary clarifications have shown the way. In order to understand the modes of existence of media products, a distinction between media products and technical distribution media must be made – technical media being the necessary conditions for making media products discernible. Media products are always material, although they have capacities for making a certain mental impact; in other words, they have both presemiotic and semiotic qualities. A methodical, bottom-up conception of material, spatiotemporal, sensorial, and semiotic media traits paves the way for avoiding blunt and misleading dichotomies such as “text” vs. “image” or “verbal” vs. “visual.” Instead, basic media similarities and dissimilarities can be analyzed in a more refined way, which makes it possible to more sharply discern the obstacles of transferring cognitive import among minds – which goes far beyond the more trivial phenomenon of noise.

Finally, the medium-centered model of communication is designed in such a way that it is compatible with communicative features that can be added to the irreducible entities and their interrelations. It can also be developed to account for serial communication, and communication among more than two minds, thus incorporating all conceivable kinds of human communication.

Acknowledgement: This article is dedicated to Jørgen Bruhn, who always urged me to theorize the notion of medium within a communicative frame.
References


Peirce, Charles Sanders. 1932. *The collected papers of Charles Sanders Peirce*, vol. 2, Charles Hartshorne & Paul Weiss (eds.). Cambridge: Harvard University Press. [Reference to Peirce’s papers will be designated CP followed by volume and paragraph number.]


Abstract
A broad variety of media traits are transmedial in the sense that they can, to a certain extent, be transferred among media that differ in fundamental ways. This article presents a new theoretical framework for studying media transformation, which should be understood as the transfer of transmedial characteristics. The goal is to explain how meaningful data are changed or corrupted during transfer among various media. First, I launch a few fundamental theoretical distinctions concerning the creation of meaningful media data. The most fundamental distinction is that between mediation and representation. Whereas mediation is the material prerequisite for representation in media, representation should be understood as a semiotic operation, that is, the creation of meaning in the mind. On the basis of this division, I also distinguish between two kinds of media transformation: transmediation and media representation. The article then continues with a section about the transmedial basis. All media have basic and universal (material, sensorial, spatiotemporal and semiotic) properties that are shared to

1 This article is a condensed version of some of the main contentions in my book, *Media Transformation: The Transfer of Media Characteristics Among Media*. Further conceptual developments, more detailed discussions and plenty of empirical examples are found in the volume published by Palgrave Macmillan in 2014.

2 orcid.org/0000-0002-8664-1970. Linnaeus University. Sweden. lars.ellestrom@lnu.se
some extent. Furthermore, media form compound characteristics (such as narrativity) that are more or less transmedial, which means that they can be transferred among media to some extent. Finally, a model for analyzing media characteristic transfer is outlined.

**Keywords**
Intermediality; semiotics; adaptation; transmediation; media representation (Source: Unesco Thesaurus).
Transferencia de las características de los medios entre medios disimilares

Resumen
Una amplia variedad de rasgos de los medios son transmediales en el sentido en que pueden, en cierta medida, ser transferidos entre medios que difieren en formas fundamentales. Este artículo presenta un nuevo marco teórico para el estudio de la “transformación mediática”, el cual debe entenderse como la transferencia de características transmediales. El objetivo es explicar cómo los datos significativos se cambian o corrompen en la transferencia entre varios medios. En primer lugar, se hacen algunas distinciones teóricas fundamentales con respecto a la creación de datos mediáticos significativos. La distinción más fundamental es la que existe entre la mediación y la representación. Mientras que la mediación es el prerrequisito material para la representación en los medios, la representación debe ser entendida como una operación semiótica, es decir, la creación de significado en la mente. A partir de esta división, también se distingue entre dos tipos de transformación mediática: la transmediación y la representación mediática. A continuación, el artículo dedica una sección a la base transmedial. Todos los medios tienen propiedades básicas y universales (materiales, sensoriales, espaciotemporales y semióticas) que son compartidas en cierta medida. Además, los medios forman características compuestas (como la narratividad) que son más o menos transmediales, lo que significa que pueden transferirse entre los medios hasta cierto punto. Finalmente, se esboza un modelo para analizar la transferencia de características mediáticas.

Palabras clave
Intermedialidad; semiótica; adaptación; transmediación; representación mediática (Fuente: Tesoro de la Unesco).
Transferência das características da mídia entre os meios dissimilares

Resumo
Uma ampla variedade de rasgos da mídia são transmídiais no sentido em que podem, em certa medida, ser transferidos entre os meios que diferem em formas fundamentais. Este artigo apresenta um novo marco teórico para o estudo da “transformação mediática”, e que deve entender-se como a transferência de características transmídiais. O objetivo é explicar como os dados significativos são alterados ou corrompidos durante a transferência entre vários meios. Em primeiro lugar, faço algumas distinções teóricas fundamentais relativas à criação de dados mediáticos significativos. A distinção mais fundamental é a que existe entre a mediação e a representação. Enquanto a mediação é o pré-requisito material para a representação na mídia, a representação deve ser entendida como uma operação semiótica, ou seja, a criação de significado na mente. A partir desta divisão, distingo também dois tipos de transformação da mídia: a transmediação e a representação mediática. O artigo continua depois com uma seção sobre a base transmídia. Todos os meios têm propriedades básicas e universais (materiais, sensoriais, espaciotemporais e semióticas) que são compartidos em certa medida. Além do mais, os meios formam características compostas (tais como a narratividade) que são mais ou menos de caráter transmídia, o que significa que podem transferir-se entre os meios até certo ponto. Finalmente, se esboça um modelo para analisar a transferência de características mediáticas.

Palavras-chave
Intermedialidade; semiótica; adaptação; transmediação; representação mediática (Fonte: Tesouro da Unesco).
Meaningful data are constantly exchanged among people with or without external technical devices and they are also transferred among different kinds of media. We talk and write to each other, create music and pictures and transfer content among an abundance of different media. When commenting on a newspaper photograph, a visual and static picture is transformed into audible words; when making a movie based upon a graphic novel, a visual and static medium based on iconic structures and symbolic words is transformed into a similarly based audiovisual, spatiotemporal medium. In neither case does the transfer take place seamlessly.

Transfer of media characteristics among dissimilar media is an exceedingly widespread phenomenon, fundamental for most communicative situations. The problem is that we do not have, so far, a comprehensive theory for analyzing and understanding the complex interrelations between media transfers’ material and cognitive facets. The most influential recent study that touches upon the field is Jay Bolter and Richard Grusin’s Remediation: Understanding New Media from 1999. While one of its great merits is its many observations of what the authors call remediation, the fundamental notions of media and remediation are only vaguely outlined. Furthermore, the authors refer mainly to visual media. The notion of remediation is a good start, but I believe that we need stern theory to really understand the complicated process of transferring characteristics among media. Such theories must ideally include aspects of media materiality and sensory perception, as well as semiotic and cognitive aspects.

Consequently, the aim of this article is to develop a new theoretical framework for the study of media characteristic transfer, which I call media transformation as a general term. The goal of the framework is to explain what happens when meaningful data are changed or corrupted during transfer among different media. To me, an in-depth understanding of such processes is an acutely important matter with far-reaching consequences for understanding communication.

My approach differs from earlier media transformation studies (Clüver, 1989; Bolter & Grusin; Wolf, 1999, 2002; Rajewsky, 2002). I rely on a bottom-up model of basic media traits. Instead of beginning with a
selection of established media and their interrelations, I start with focusing on fundamental properties that are potentially shared by all media (Elleström, 2010). The similarities and differences among media are fundamental for this approach: to transfer meaningful data among dissimilar media is to transform them, which is equivalent to keeping something, getting rid of something else, and adding something new.

It should be noted that I do not wish to isolate certain media products and label them as transmedial. For me, transmediality is an analytical perspective. All media products can be investigated from both a synchronic perspective, in terms of combination and integration, and from a diachronic perspective, in terms of transfer and transformation. No doubt, certain media products, analyzed diachronically, tend to produce meaning prolifi-
cally vis-à-vis their relations to other, pre-existing media products; however, there are no media products that cannot be treated in terms of media transformation without some profit.

The article begins with a few fundamental theoretical distinctions concerning the creation of meaningful data by media. It then continues with a section about transmedial characteristics and ends with the sketch of a proposed model for analyzing the media characteristic transfer.

**Mediation vs. Representation and Transmediation vs. Media Representation**

My main distinction, which is rarely highlighted in intermedial relations studies, is between *mediation* and *representation*. If these two notions are conflated, it becomes difficult to discern certain important media transformation stages and aspects (Elleström, 2013).

Mediation, as I define it, is a pre-semiotic, physical realization of entities (with material, sensorial, and spatiotemporal qualities, as well as semiotic potential) perceived by human sense receptors within a communication context.
Representation, as discussed here, is the creation of meaning in the perceptual and cognitive acts of reception. To say that a media product represents something is to say that it triggers a certain kind of interpretation. This interpretation may be more or less hardwired in the media product and the way one perceives it, but it never exists independently of the recipient’s cognitive activity. When something represents, it calls forth something else; the representing entity makes something else, the represented, present to the mind. As noted by Charles Sanders Peirce (1960), a sign, or a representamen, stands for an object.\(^3\) Representation, the very essence of the semiotic, is often a result of mediation.

The notion of mediation thus foregrounds the material realization of the medium, whereas the notion of representation highlights the semiotic conception of the medium. This distinction is helpful for analyzing complex relations and processes. In practice, however, mediation and representation are deeply interrelated. Every single representation is based on a specific mediation’s distinctiveness. Furthermore, some kinds of mediation facilitate certain sorts of representation while rendering other sorts impossible. For example, vibrating air emerging from vocal cords and lips, perceived as sound, is well suited for iconic representation of bird song, though such sound cannot possibly form a detailed, three-dimensional iconic representation of a cathedral.

Hence, I use the term mediate to describe the process of a technical medium realizing potentially meaningful sensory configurations: a book page can mediate, say, a poem, a diagram, or a musical score. If equivalent sensory configurations (that is, those that have the capacity to trigger corresponding representations) are mediated for a second (or third, or fourth) time by another kind of technical medium, they are transmediated: the poem that was seen on the page can later be heard when it is transmediated by a voice. In other words, the content of the poem is represented again by a new kind of sensory configuration (not visual, but auditory signs), mediated by another kind of technical medium (not a book page emitting photons, but sound waves generated by vocal cords).

---

\(^3\) See, for instance, CP 2.228–229 [c. 1897].
The concept of transmediation involves not only *re-*mediation, that is, repeated mediation, but also repeated mediation of equivalent sensory configurations by another kind of technical medium (please note that the term remediation, as used here, should not be understood in Bolter and Grusin’s open-ended sense). Hence, the composite term *transmedial remediation* would be more accurate for the concept in question. For the sake of simplicity, I prefer the brief term *transmediation*.

All transmediation involves some degree of transformation: the equivalent sensory configurations and the corresponding representations that they trigger may be only slightly different and clearly recognizable, but they may also be profoundly transformed (for example, musical narratives based on literature differ very much from their sources).

Transmediation is the first kind of media transformation. Also, media representation, the second kind of media transformation, involves modification in the transfer process. Such media representations as, for instance, a news article describing a documentary, or a photograph depicting a dance performance, should be understood as potentially representing both a medium’s form and its content; media representation is at hand when one medium presents another medium to the mind. A medium, something that represents, becomes itself represented.

The distinctive features of transmediation and media representation can be seen in Figure 1. The M circles should be understood as media products or, more specifically, as technical media mediating sensory configurations. The C, which is placed within an M, should be understood as media characteristics represented by the sensory configurations. A circle and its interior is thus a depiction of both the mediation and the ensuing representation. The T arrows represent the transfer acts between two media products: the source medium (M1) and the target medium (M2). M2 is thus a new technical medium mediating more or less different sensory configurations compared to M1.
In the case of transmediation, the target medium (M2) represents the same content (C1) as the source medium (M1); in the case of media representation, the target medium (M2) represents the source medium (M1), which means that the source medium constitutes the media characteristics of the target medium (C2 = M1). As represented content of M2, M1 still represents C1. In other words: In the first media transformation case, the target medium (M2) transmediates (represents again) the source medium characteristics (M1). In the second case, the target medium (M2) represents the source medium (M1). In both cases, the source medium’s characteristics (C1) must be understood to remain the same, to a certain extent, after the transfer from M1 to M2.

Consequently, media representations may often be understood also as transmediations if they include, to some degree, a repeated mediation of equivalent sensory configurations. There is no contradiction between a target medium representing, on the one hand, a source medium, and, on
the other hand, mediating sensory configurations equivalent to those of the source medium. This might be inferred from the diagram in Figure 1. A photograph representing a drawing of three kittens is obviously a medium representing another medium, but it also clearly includes repeated mediation of not only equivalent, but actually very similar (visual) sensory configurations by another technical medium. An auditory, verbal description of a drawing such as “I bought a drawing of three cute kittens” is also a case of media representation, but since it includes repeated mediation of equivalent sensory configurations by another technical medium (the voice is able to produce symbolic signs that represent substantial parts of the objects represented by the iconic signs on the paper: the notion of three kittens), it also includes transmediation.

Both these examples may be understood as comparatively complex instances of media representation and it is clear that if a medium is represented in some detail, the represented source medium characteristics become transmediated by the transfer target. However, a very simple verbal representation such as “I bought a drawing” is a media representation, but not a transmediation; C1 (in this case, the represented kittens) is not represented again. Hence, in pure media representation, only M1, the empty shell of the source medium, so to speak, is transferred to M2. In pure transmediation, only C1, the source medium content, is transferred to M2. Often, both M1 and C1 are transferred, which means that both media representation and transmediation are present. To be strict, then, the diagram representing media representation in Figure 1 actually depicts media representation including transmediation. Pure media representation should be depicted as in Figure 2.

Other distinctions that should be kept in mind are that both transmediation and media representation can involve, on the one hand, specific media products (which has been assumed so far), and, on the other hand, general characteristics of qualified media. Qualified media is a term I use to denote all kinds of abstract media categories—both artistic and non-artistic—that are historically and communicatively situated, meaning that their properties differ depending on time, culture and aesthetic preference. Qualified media
include such categories as music, painting, television programs, and news articles. A qualified medium is made up of a cluster of concrete media products (Elleström, 2010). Media products can represent other specific media products, as well as general characteristics of qualified media, both of which can be transmediated (Elleström, 2014). While a novel may describe a particular piece of music, it may also discuss and, hence, represent music in general. A poem may transmediate characteristics of a specific musical piece, but it may also transmediate general musical characteristics, such as formal traits. Hence, the diagrams in Figure 1 can also be extended to illustrate transmediation and media representation involving general media characteristics, in which case $M$ must be understood as the idea of a qualified medium and $C$ as general media characteristics.

Certain types of complex, specific media product representations are commonly called *ekphrasis*. Whereas an ekphrasis is typically understood to be a poem representing a painting, the notion has been extended substantially during the last decades (Yacobi, 1995; Clüver, 1997; Bruhn, 2000; Sager Eidt, 2008). *Ekphrasis*, in turn, is only the tip of the iceberg of media representation.

The general term for transmediation of media products is “adaptation” (Elleström, 2013). While the archetypal adaptation is a novel-to-film transmediation, the term has not been reserved exclusively for this type of transfer (Elliott, 2003; Hutcheon, 2006; Bernhart, 2008; Urrows, 2008; Schober, 2013). Furthermore, far from all types of transmediation of specific media products tend to be called adaptation. Transmediations from libretti, scores,
scripts and so forth, and transmediations from written, visual and verbal text to oral, auditory and verbal text (aloud readings of texts), or the other way round, to mention only a few examples, are very seldom referred to as “adaptation” (see, however, Groensteen, 1998).

The Transmedial Basis

So far, transmediation and media representation have been discussed without really asking how these phenomena are at all possible. Which features are involved in the transformational processes encompassing several media and how are they related? Initially, it must be restated that no medium can fully transmediate or represent all media. Both qualified media and individual media products have dissimilar basic properties; these differences set the limits for what can be transmediated or represented. In addition, transmediated and represented media characteristics are not equally transmedial; while certain traits are almost universally present in the media landscape, others can only be marginally transformed to fit other media.

I refer to this wide range of media features as the transmedial basis. The questions are: which characteristics can be transmediated or represented by other media, and why? As media characteristics are often results of contextualization and complex interpretive practices, this question can be treated systematically only to a certain extent.

The most elementary transmedial basis consists of what I have called elsewhere the “four modalities of media”—the material, the sensorial, the spatiotemporal and the semiotic (Elleström, 2010). A modality should be understood as a category of related characteristics that are basic in the sense that all media can be described in terms of all four modalities. All individual media products, and all conceptions of qualified media, may be understood as specific combinations of modes of the four modalities. The modes of the modalities do not cover all media characteristics—far from it—but they constitute a sort of a skeleton upon which all media are built.

The four modalities of media and, more specifically, the modes of the four modalities, thus constitute an essential transmedial foundation. The
flat surface, being a mode of the material modality, is an aspect of printed novels, etchings, posters, television news and so forth, and is a prerequisite for comprehensive and close transmediation of, say, a graphic novel to a motion picture. The audible, a mode of the sensorial modality, is an aspect of radio theatre, opera, ordinary conversation, gamelan music and many other types of qualified media. The best way of faithfully representing sound media characteristics is to produce similar sounds.

Temporality, which must be understood as a mode of the spatiotemporal modality, is an aspect of recited poetry, theatre and television commercials, but not of oil paintings and printed tourist brochures. While all media are perceived in time, only some media are temporal in themselves. Transmediation often involves media that are either temporal or non-temporal. Graffiti is easily and faithfully transmediated by still photographs, whereas stills can only partially transmediate dance. However, some qualified media, such as most written, visual literature, are conventionally decoded in fixed sequences, which makes them second-order temporal, so to speak, and hence well-suited for transmediation into temporal media, such as motion pictures.

Iconicity, a mode of the semiotic modality, is a vital aspect of creating meaning in media, such as newspaper advertisements, statistical graphs, rock music, and scholarly figures (such as the ones in this article). Iconic structures create meaning on the ground of resemblance; similarities can be established over both sensorial and spatiotemporal borders. For example, visual traits may depict auditory or cognitive phenomena, and static structures may depict temporal phenomena; that is, a graph may depict both changing pitch and altering financial status. In general, iconicity interacts with the two other main modes of the semiotic modality: indexicality (meaning created by contiguity) and symbolicity (meaning created by conventions). This well-known trichotomy originates in Peirce.4

The modes of the modalities are clearly necessary for identifying media similarities and differences and, consequently, essential for delineating

---

4 See, for instance, CP 2.304 [1901].
processes of transmediation and representation, though they do not definitively determine their limits. It may be the case that shared modality modes facilitate extensive transmediation and representation, while some media are very difficult to transmediate or represent if the target medium does not possess vital modality modes. Nevertheless, due to the brain’s cross-modal capacities, transmediation and media representation over modality mode borders are, to a certain extent, possible, common and, indeed, productive.

To summarize, it is the material, sensorial, spatiotemporal and semiotic differences between source medium and target medium that allow for inventive alterations of media products into new creations. By the same token, modal differences make it impossible to transfer vital information without transforming it, as in news reports that include chains of interconnected media.

While the modes of the four media modalities are basic and universal transmedial characteristics, it is not the modality modes that are transferred in the processes of media transformation; rather, they are deeply integrated fundamentals that are required for forming what I call “compound media characteristics” (Elleström, 2014). Many media products share modality modes, but each individual media product has a distinct set of compound media characteristics created by the arrangement of all details in the full medial expression, as discerned and construed by the media product’s perceiver. Compound media characteristics should be understood as features of media products that are apprehended and formed when a structuring and interpreting mind makes sense of the mediated sensory configurations.

Compound media characteristics may be strongly linked to certain modality modes, to a specific media product, or to a qualified medium; they may also be transmedial to a considerable degree, meaning that they can successfully be transferred among many kinds of qualified media. However, compound media characteristics can never be fully transmedial: the modal differences among dissimilar media always make a difference. Nevertheless, the mind has the capacity of recognizing similarities that bridge media variances.
Compound media characteristics that can be transferred among media can roughly be divided into aspects of form and content, which should be understood as a coarse distinction between comprehensive media characteristics and more confined elements.

Form includes all kinds of structure, manifested sensorially in what can be seen, heard or otherwise perceived, or as cognitive configurations. There is a wide range of aspects or derivations of structure: pattern, rhythm, balance, proportion, relation, repetition, similarity, and contrast, to mention only a few. In spite of their (sometimes) inaccessible nature, these notions have extraordinary longevity, which must mean that they correspond to basic perceptual inclinations and fulfil vital cognitive needs.

Other formal compound media characteristics are the intricate qualities of style and perspective. Lately, Werner Wolf has provided several volumes on formal transmedial characteristics, such as description and metareference (Wolf & Bernhart, 2007; Bantleon, Thoss & Wolf, 2009). One complex transmedial characteristic that I have investigated myself is irony (Elleström, 2002).

When it comes to content, some compound media characteristics are directly perceptible from the material interface of the media product as the appearance of symbolic and iconic microstructures, such as visual or auditory words and sentences, and visual or auditory iconic details. Other compound media characteristics that have the nature of content are products of cognitive activity triggered by processes of representation and conceived as situations, spaces, places, persons, animals, objects, motifs and the like. All these characteristics are transmedial to a certain extent.

Narration, which includes aspects of both form and content, is one of the most important compound media characteristics. While narration is traditionally associated with literature and motion pictures, it has increasingly gained the status of a fundamental cognitive notion. Narration can be said to be an offspring of structured perception and spatial thinking. To narrate, and to interpret in terms of narration, is a way of creating meaning
in sequential form. Narration is not limited to specific material or sensorial modes. Our two most cognitively developed senses, sight and hearing, are both well suited to narration, and all types of spatiotemporal configurations may display traits that can be connected to narrative sequences. Naturally, however, media that are temporal on the material level, such as movies and music, and those that are based on conventionally sequential sign systems, such as oral and printed literature, have an advantage when it comes to forming developed narratives. Furthermore, media that rely on advanced, symbolic sign systems (primarily verbal language), such as literature and motion pictures, are well suited to outline complex narratives. Certainly, many kinds of narratives can be transferred among media. The phenomenon has been investigated by Marie-Laure Ryan (2004), among others, who has explored what she calls “transmedial narratology”.

Plots and stories are two kinds of narrative sequential structures that can be more or less fully transferred among media. In addition, the storyworld, which includes an elaborated virtual space and should count as form, can at least partly be transferred among different media, as can portions of narratives, such as relations among particular characters or other figurations. Linda Hutcheon (2006) has listed a number of features that are certainly compound media characteristics and should be understood as transmedial narrative content: characters, motivations, consequences, events, symbols, and themes, among others.

It is clearly impossible to create an exhaustive list of transmedial compound media characteristics. Furthermore, media characteristic complexity makes neat classification very difficult. Nevertheless, the notion of compound media characteristics cannot be dispensed with if the idea of transfer among media shall be fully understood.

**A Basic Formula for the Transfer of Media Characteristics**

On the basis of Figure 1, I would like to propose a formula for the rudimentary traits of media transformation; this involves recapitulating the central issues presented so far. Since “trans” means “across” or “beyond”, the term
must be understood to represent a spatiotemporal notion: compound media characteristics are transferred from one place to another. First, we read a novel in a book and then we see a motion picture and recognize it as more or less the same story. First, there is a sculpture placed on the square and then we see photos and read newspaper descriptions of it. Hence, media transformation can be captured in the formula “A compound transmedial Characteristics is Transferred from a source Medium to a target Medium”, or “C is T from M1 to M2”.

The transfer is either a transmediation or a media representation or a combination of the two. In either case, it involves some sort of transformation. However, in what follows, I will mainly discuss transmediation in light of one practical example to make the rudimentary aspect survey as clear as possible: William M. Thackeray’s 1844 novel The Luck of Barry Lyndon being transmediated into Stanley Kubrick’s 1975 motion picture Barry Lyndon. This is a typical example of adaptation that I will refer to briefly in order to illuminate some standard transmediation traits.

Vis-à-vis Barry Lyndon, how should C, T, M1 and M2 be understood? M1 is the source medium, the “first place”, and M2 is the target medium, the “second place”. When thinking of media characteristic transfer in the most straightforward way, M1 and M2 are two particular media products, as in the case of Barry Lyndon, where the novel’s central, compound characteristics are transmediated by the movie. However, transmediation also occurs in cases where either M1 or M2 is a qualified medium or a submedium (genre) rather than a particular media product, which may be illustrated with the same example. One of the traits of the movie is voice-over narration. When seeing Barry Lyndon as a version of the novel, this voice-over is part of the general transmediation of The Luck of Barry Lyndon, yet one may also say that the voice-over is simply a novelistic trait of the motion picture. Indeed, there are many movies with voice-over that cannot be understood as transmediations of particular literary works. Nevertheless, they can be understood within the framework of the formula “C is T from M1 to M2”, the difference being that M1 is a genre of written literature, a qualified medium, and M2 stands for particular media products. The formula
might then be “C is T from MQ1 to MP2”, where MQ means qualified medium and MP means a specific media product.

Furthermore, both M1 and M2 may be qualified media or genres (“C is T from MQ1 to MQ2”). It makes perfect sense to talk about, say, “novelistic traits in movies”, and indeed the transformation of The Luck of Barry Lyndon to Barry Lyndon may be understood as a particular instance of novelistic traits in movies.

As explained earlier, C represents transmedial compound media characteristics, such as form, structure, rhythm, narration, material, theme, motif, and so forth. Obviously, characteristics that are not transmedial cannot be transmediated by other media and not all kinds of transmedial characteristics can be transmediated by all media. In Barry Lyndon, a certain narrative form (the rise and fall of an ambitious man), many verbal microstructures, several characters, themes, and motifs, and probably many other characteristics can be said to be transmediated from the novel.

However, the book and the cinema or television screen are entirely different technical media, so the transmediation necessarily involves several modal changes: the auditory mode is added, the novel’s conventional sequentiality is transformed to the material temporality of the movie, the degree of iconicity of the visual surface increases dramatically, and so forth. Yet, vital narrative form aspects survive the transmediation from conventional sequentiality to material temporality: the verbal micro-structures are easy to transfer from the visual to the auditory; many of the qualities of the characters in the novel can also be expressed in the movie, since verbal language and images overlap to a large degree when it comes to what they can represent.

Finally, T is the transfer; it includes both transmediation and media transformation, of course, although the focus is on transmediation for the moment. I will here pay attention to three further aspects of T that are valid for both transmediation and media transformation.
The first aspect concerns the “thickness” of the $T$ arrow and includes differences between what might perhaps be called more and less complete transfers, strong and weak transfers, and so forth. Conversely, one may perceive differences between degrees of transformation; a less complete transfer is likely to include a higher degree of transformation.

As we have seen in the *Barry Lyndon* example, a particular transmedial may involve several compound media characteristics. The more characteristics it involves, the closer the target medium is to the source medium. If very few transmedial characteristics are involved, it might be questioned whether there is a point in treating it as a transmediation at all. As noted, transmedial characteristics are, as a rule, more or less modified by the modal changes involved in a media transformation, which certainly has an effect on the perceived transfer strength.

The second aspect concerns the “direction” of the $T$ arrow. In the straightforward standard transfer, the arrow points from $M1$ to $M2$ (as in Figures 1 and 2). This is how media representation must always be understood: one particular media product ($M2$) represents other media products or qualified media ($M1$). There is no question about what represents and what is being represented. In this respect, transmediation is more complicated. When both $M1$ and $M2$ are qualified media or submedia (genres), there is not always a point in saying that the compound media characteristics are definitely transferred from one place to another; the truth might rather be that they “circulate”, or go back and forth between $MQ1$ and $MQ2$ as in the development of forms and motifs in modern literature and film. The $T$ arrow sometimes points in many directions.

The third aspect concerns the “extension” of the $T$ arrow. Individual transfers must also be seen as parts of more far-reaching and complex networks involving many specific media products (MP3, MP4, and so on). The $T$ arrow may be part of arrow chains, perhaps with weak and strong links or thin and thick branches. There may also be several source media that are transformed to one new target medium. The number of potential
transmediation variations is probably endless. Figure 3 illustrates one example: two qualified media that are transmediated to one media product, as when an advertisement borrows traits from both concrete poetry and comics, or when a photograph has the appearance of both a classicist painting and a scene from theatre drama.

Figure 3. An example of transmediation

After this rather formalistic presentation of distinctions, diagrams, and formulas not intended to capture the phenomenon of media transformation in strict and endless subdivisions, but rather to make possible methodical analyses of a multi-faceted area that no doubt escapes neat classifications, I want to emphasize that there is necessarily a strong subjectivity element in all media transformation discussions. Hermeneutics can never be escaped. When finding traces of other media in media, whether they are specific media products or qualified media, it sometimes simply makes sense to say that some media should be treated as source media because they are recognizable in other media, which may then be treated as target media. Ultimately, theoretical analysis is nothing without interpretation.
References


