As We May Interact: Challenges and Opportunities for
Next-Generation Human-Information Interaction

Beat Signer
Web & Information Systems Engineering Lab
Vrije Universiteit Brussel
Brussels, Belgium
bsigner@vub.be

ABSTRACT

Long before the advent of personal computing, Vannevar Bush envisioned the Memex as a solution to address information overload by enhancing the management and refining of information through associative trails. While other hypertext pioneers like Douglas Engelbart and Ted Nelson introduced advanced hypertext concepts to create more flexible document structures and augment the human intellect, some of their original ideas are still absent in our daily interaction with documents and information systems. Today, many digital document formats mimic paper documents without fully leveraging the opportunities offered by digital media and documents are often organised in hierarchical file structures. In this keynote, we explore how cross-media technologies, such as the resource-selector-link (RSL) hypermedia metamodel, can be used to organise and interact with information across digital and physical spaces. While emerging wearable mixed reality (MR) headsets offer new possibilities to augment the human intellect, we discuss how hypermedia research, in combination with other technologies, could play a major role in providing the necessary linked data and hypertext infrastructure for this augmentation process. We outline the challenges and opportunities for next-generation multimodal human-information interaction enabled by flexible cross-media information spaces and document structures in combination with upcoming mixed and virtual reality solutions.

CCS CONCEPTS
• Human-centered computing → Hypertext / hypermedia; Mixed / augmented reality; Human computer interaction (HCI); • Information systems;

KEYWORDS
Hypertext, cross-media technologies, human-information interaction, mixed reality, multimodal interaction

ACM Reference Format:

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored.

For all other uses, contact the owner/author(s).
HUMAN ’24, September 10, 2024, Poznań, Poland © 2024 Copyright held by the owner/author(s).

Figure 1: Next-generation human-information interaction (image created with the assistance of DALL-E 3)

While Vannevar Bush’s visionary Memex [3] aimed to extend a user’s memory, early hypertext pioneers like Ted Nelson with Project Xanadu [8] and Douglas Engelbart with his oN-Line System (NLS) [5] focused on augmenting human intellect through non-linear and non-hierarchical information structures. Since then, various hypertext research has been conducted and applied to different domains, including hypertext as knowledge representation, interface or infrastructure [1]. Additionally, hypertext has been proposed as a method of inquiry for arbitrary systems with regard to human augmentation and the potential benefits of different forms of first-class structures [2].

Over the last two decades, we have conducted research on cross-media information spaces [14] to organise, refind, augment and interact with information. Our data-centric approach enables the structuring and linking of information across digital and physical spaces based on the resource-selector-link (RSL) hypermedia metamodel [17] offering bidirectional context-aware navigational as well as structural links, versioning and user management. While the RSL hypermedia metamodel was initially developed for interactive paper solutions [13], it has since been applied to the domains of personal cross-media information management (PIM) [18], non-linear cross-media narratives [19], Internet of Things (IoT) applications [10], next-generation presentation solutions [9] and personalised technology-enhanced learning environments [7]. Further, we are currently exploring fluid document formats [12] and the use of hypertext concepts in desktop environments [16].
Recent major advances in wearable mixed reality headsets, such as the Microsoft HoloLens 2, Varjo XR-4, Apple Vision Pro or Meta Quest 3, present exciting opportunities for natural human-information interaction to augment the human intellect. While there is a focus on the necessary MR hardware solutions and individual mixed reality applications, we believe a major challenge will be providing the necessary hypertext infrastructure to support general augmentation and human-information interaction in future mixed reality environments as indicated in Figure 1.

Another key challenge in designing general future mixed reality solutions for human-information interaction is ensuring independence from specific hardware platforms. Fortunately, the development of the W3C’s WebXR Device API [6] is a shared effort that enables web-based mixed reality solutions to be executed in any web browser supporting the WebXR Device API. These browser-based mixed reality applications can then be executed in full-screen mode on mixed reality headsets or other devices like tablets or smartphones with video see-through mixed reality capabilities. To anchor any hypertext structures in real-world environments, various vision- or marker-based approaches can be used to identify individual physical objects or classes of objects. These objects might then be mapped to a Uniform Resource Identifier (URI) for use in a hypertext infrastructure. Further, anchors can be defined with fixed positions in space rather than being bound to specific objects. Another consideration is where to store the data and metadata used to augment real-world objects. Again, web-based solutions can be employed for this purpose. In addition, the collaborative authoring and sharing of information for mixed media environments could, for instance, be realised based on de-centralised Solid Pods [4]. Similar to the W3C’s Web of Things (WoT) Thing Description [11], we might also use a web-based approach to describe passive physical things. By primarily using web-based technologies, mixed reality solutions become just another method to define anchors and interact with existing cross-media information spaces and hypertext infrastructures, including the implementation of our RSL hypermedia metamodel [17].

In addition to a hypertext infrastructure for the general linking between digital or physical artefacts and the definition of complex structures on top of these entities, there are opportunities for new forms of multimodal human-information interaction paradigms in the resulting mixed media environments. For example, a user could point to individual physical artefacts to see potential digital augmentations or links to other digital and physical objects as well as superimposed overlay structures. Furthermore, users can not only receive multimodal feedback while navigating information in these environments but also create new content or associations between digital and physical resources. For instance, a user might use a pointing gesture from one physical artefact to another to create a new link between the objects and attach some additional semantics to the link via visual or voice input. Users can seamlessly move between environments, and a resource in a mixed reality environment might serve as an entry point into a virtual reality (VR) environment. In addition, it is possible to physically augment digital information as demonstrated in dynamic data visualisation [15].

Nowadays, the need for general hypertext infrastructures is greater than ever to enable next-generation multimodal reality user interfaces and advanced forms of human-information interaction. While web-based technologies like the Web of Things, WebXR and the Solid Project could contribute to such an infrastructure, we might also leverage earlier hypertext research to build more flexible collaborative hypermedia and cross-media information spaces based on navigational as well as structural context-dependent hypertext. By adopting a data- and hypertext-centric approach, we can enable sophisticated cross-media and mixed-reality solutions to augment the human mind.

**REFERENCES**


