



## Open Research Online

### Citation

Antonini, Alessio (2022). On Epistemic Comparability and Challenges on Data Reuse: The Experience of READ-IT. In: DH2022 book of abstracts, 25-29 Jul 2022, Japan.

### URL

<https://oro.open.ac.uk/82366/>

### License

None Specified

### Policy

This document has been downloaded from Open Research Online, The Open University's repository of research publications. This version is being made available in accordance with Open Research Online policies available from [Open Research Online \(ORO\) Policies](#)

### Versions

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding

# On Epistemic Comparability and Challenges on Data Reuse

The Experience of READ-IT

Alessio Antonini (alessio.antonini@open.ac.uk), The Open University, United Kingdom

The EU JPI Reading Europe Advanced-Data Investigation Tools (READ-IT) project (<https://readit-project.eu>) proposed a novel approach to a pan-European research agenda on reading based on shared tools and data interoperability "by design" (Our-Vial & Antonini, 2021). These tools were developed in parallel with and support a wide diverse range of case studies (Vignale et al, 2019). The toolkit includes a) the Reading Experience Ontology (REO) (Vignale et al, 2020; Antonini et al, 2021), b) the text-annotation tool (<https://read-it.hum.uu.nl>) and c) the platform for managing reading testimonies (<https://read-it.in-two.com>). However, while reflecting on using the data at scale, a gap in the models emerged concerning the epistemology of sources, i.e. the conditions about testimonies of reading were generated. This contribution reports on this gap, its implications for future visions for research interoperability and our approach to address it.

The focus of READ-IT on data interoperability led to good results toward a shared definition of the phenomena and response to reading. However, the information needed to interpret testimonies extends beyond readers and reading, e.g., to the situation that led to the generation of the testimony, the socio-economical context, and the temporal distance between reading and the testimony. The result is an epistemological barrier to data reuse, comparability, and integration of experiential studies. The following two real examples expose the hidden implications of building new reading case studies on existing data from, e.g., linked open data service.

**Case 1 - Exploration and reuse.** The discovery and available data on reading could exploit either provenance or metadata or, more specifically, aspects of the reading experience. For instance, PoKUS (<https://pokus.ffzg.unizg.hr/en/>) – a study on reading in Croatia – aims to reuse READ-IT data about reading memories (REO concept of “Memory” specialisation of “State of Mind” [5]). The collected memory annotations [7] include evidence from WWI diaries, readers of Russian periodicals, Soviet Czech school diaries, social media, interviews, or online reading groups. While discussing with researchers about their specific case studies, it was clear that, e.g., social pressure and censorship on Czech students limited the insight that could be extracted by their carefully crafted school diaries or soldier diaries written as part of a veteran support programme, reporting reading years after they occurred.

**Case 2 – Heterogeneous comparative studies.** Having a shared definition of reading enables comparative studies among heterogeneous sources. For instance, an ongoing comparative study on popular reading brings together a) long, matured reviews of books from online reading groups, with b) short impulsive comments on webcomics "issues". The design of heterogeneous comparative studies also requires careful considerations (Benatti et al, 2021) about, e.g., the practices of the audiences, maturity of the experience or type of information provided, connected by the form of prompting.

To summarise, data interoperability is a precondition, but it is not sufficient: a common language of the phenomenon does not exhaust the information needed for interpretation. Specifically, in the vision of step further toward research interoperability - as the synergy between studies in terms of agenda and findings contributing to the knowledge of a common phenomenon – a computer-readable epistemology of research case studies should be included as a part of their output.

In both cases, the data do not reflect these considerations, emerged only by talking directly with the involved researchers. To address this issue, we designed two un-planned ontologies: Experience & Observation (E&O) ([https://github.com/modellingDH/odp\\_experience](https://github.com/modellingDH/odp_experience)) and Profiles, Groups & Communities (PGC) (<https://github.com/modellingDH/profile-group-community-odp>).

E&O is currently in the READ-IT contribution platform. The first application of E&O is to document the different modalities we use in crowdsourcing of experience of reading, e.g. through postcards, webforms and chatbot conversations

(<https://readit-project.eu/contact/contribute-to-read-it/>). Secondly, E&O describes the relations between activities and prompting of reading experiences, typical of the different sources (e.g., topics of the questions asked, time from reading). This use of E&O helped identify nine recurring patterns exposing new, unexpected similarities and differences between case studies that can be evaluated through objective metrics (Antonini et al, under review).

PGC complements the characterisation of readers introduced by REO, with the difference between reader and status of the reader at the time of reading, by introducing the missing social dimension. Specifically, PGC addresses the specific reader profiles in terms of, e.g., linguistic competencies or core values, their belonging to social bodies (religious groups, political parties) or community of practice (as reading groups or professions such as editors or scholars).

This contribution addresses the epistemic gap in how data about experiential research, which is a barrier in moving beyond data interoperability toward a conceptual integration between different research case studies. This gap is addressed by introducing two models used to explicit essential information about research case studies and sources to complement and give context to research data. Both gap and solution are discussed based on actual case studies on reading.

## References

Enter your references here:

- Our-Vial, B. and Antonini, A.** (2021) Reconnecting with the Evolving Journey of Reading. Cultural Practice, Re-Connect, 2021, Last access September 2021  
<https://www.culturalpractice.org/article/reconnecting-with-the-evolving-journey-of-reading>.
- Vignale, F., Benatti, F. and Antonini, A.** (2019). Reading in Europe - Challenge and Case Studies of READ-IT Project. In: Digital Humanities Conference 2019, 9-12 Jul 2019, Utrecht, Netherland.
- Vignale, F., Antonini, A. and Gravier, G.** (2020). The Reading Experience Ontology (REO): Reusing and Extending CIDOC CRM. In proceedings of DH2020. 2020, Ottawa, Ca.

**Antonini, A., Suárez-Figueroa, M. C., Adamou, A., Benatti, F., Vignale, F., Gravier, G. and Lupi, L.** (2021). Understanding the phenomenology of reading through modelling. *Semantic Web*, 12(2), 191-217.

**Benatti, F., Norrick-Rühl, C. and Antonini, A.** (2021) Reading Popular Culture Offline and Online: Outlining a Comparative Study of Reading Experiences Between Webcomics and Twenty-First Century Book Club Choices. *EADH2021*, September 2021.

**Antonini, A., Adamou, A., Suárez-Figueroa, M. C. and Benatti, F.** (under review) Experiential Observations: an Ontology Pattern-based Study on Capturing the Potential Content within Evidences of Experiences. Special issue on Cultural Heritage 2021, *Semantic Web – Interoperability, Usability, Applicability* an IOS Press Journal (under review),

URL: <http://semantic-web-journal.net/content/experiential-observations-ontology-pattern-based-study-capturing-potential-content-within>.