

# Reading in Europe - Challenge and Case Studies of READ-IT Project

## ABSTRACT

This paper aims to present the READ-IT project and the first set of case studies collected by DH and HSS researchers. Use-cases are key in the project's strategy as they are essentials both to the definition and the validation of READ-IT data model and framework. The case studies include different sources, such as social media, students' diaries and letters, from the 18th up to today, in Czech, French, German, Italian and Dutch. Each of them is supported by a specific dataset and a specific research question. In this context, this original validation process must be able to demonstrate the relevance, robustness and ability of both the general concept and the data model to process a wide variety of sources. Then, this model should be transferable to other DH projects where the experiential dimension is present.

---

This paper aims to present the READ-IT project and the first set of case studies collected by DH and HSS researchers. Case studies in reading are one of the first deliverables and results of a joint construction between DH, ICT, LIS and HSS scholars. Case studies occupy a central place in the definition of READ-IT data model and tools, guiding the identification on common issues, dimension of analysis and sources for validating and testing both the conceptual framework and the database. The case studies include different sources, such as social media, students' diaries and letters, from the 18th up to today, in Czech, French, German, Italian and Dutch.

The importance of books and reading is unquestionable in modern society, but at the same time there are still unaddressed questions. Up to now, we can study the circulation of books and therefore the idea they convey, we can identify the factors that may facilitate or impeding the reception of such ideas in different cultural groups, but we still cannot comprehend the impact of reading in our history and in our society. Regarding history of reading practices, knowledge has significantly increased over the last decades about what, where and when people read [Murray 2018]. Nevertheless, two major questions remain unanswered: why and how do people read?

We still lack a systematic approach and tools to study the experience of reading, what are the effects on readers and their lives, what are the outcomes of reading and what is affecting the reading experience of the general public. The two main research questions indicated above can also be decomposed in a series of subordinated questions:

- a. What kind of transaction exists between a reader and a text?
- b. What role does the environment play in this transaction?
- c. Is it possible to list and model the emotions caused by reading?
- d. Have these emotions changed throughout time and space in Europe?
- e. Is it possible to sketch out the portrait of something like the “European reader”?

Through a unique large-scale, user-friendly, open access, semantically-enriched investigation tool to identify and share groundbreaking evidence about 18th-21st century Cultural Heritage of reading in Europe, READ-IT (Reading Europe Advanced Data Investigation Tool) project wants to address these questions. It is a 3-year (2018-2020) transnational, interdisciplinary R&D project funded by the Joint Programming Initiative for Cultural Heritage. READ-IT consists of a robust consortium of 5 academic partners from 4 European countries (Institute of Czech Literature, Academy of Sciences, Prague; Open University, London, UK including the SME IN2; Utrecht University-DH Lab, Netherlands; CNRS-IRISA, Rennes and Le Mans Université-3LAM, France).

The interdisciplinary collaboration between digital humanists, human & social sciences scholars and computer scientists investigates innovative ways of gathering new resources through crowdsourcing and web-crawling as well as linking and reusing pre-existing datasets. READ-IT thus aims to ensure the sustainable and reusable aggregation of qualitative data allowing an in-depth analysis of the Cultural Heritage of reading. The corpus encompassed is a rich ‘human archive’ in multiple media and languages depicting a transaction between reading subjects and reading material from the 18th century up today, to web scraping and social media crowdsourced evidence of reading experiences.

With regard to the work plan of READ-IT, the collection of use cases is considered to be the first significant milestone. Use cases collected in READ-IT are challenging the previous approaches adopted in projects such as [UK-Reading Experience Database](#) (2006 to date), the ANR-funded “[Reading in Europe: Contemporary Issues in Historical and Comparative Perspectives](#)” project (2014-2017) and the [Listening Experience Database](#) project (2012 to present) by going beyond the current state of the art of use cases and by requiring a much deeper analysis of sources.

Specifically, in READ-IT we collected the following seven use cases:

1. **Studying Contemporary Digital Reading Experiences Through Social Media.** How can we document contemporary digital reading experiences? Are there any “new” ways of describing reading experiences out there? [Kovač and van der Weel 2018] Can READ-IT help us preserve a collection of present-day heritage for the future? How can we find and collect data about such digital experiences?
2. **Self-reflection.** Which readers (in terms of cognitive profiles and/or sociocultural variables) tend to report self-referencing as part of their reading experiences? Which texts prompt self-referencing more than others? [Kuzmičová 2014]
3. **The places where we read.** What text genres are read in what physical environments (and on which devices)? What physical environments are conducive to quality reading experiences as perceived/valued by readers? [Kuzmičová 2016]
4. **Four use cases** examining

- a. the semi-automated extraction of evidence of reading from challenging sources written in inflected languages (Czech, German, Italian, Russian) that have undergone significant historical change, using evidence in non-Latin alphabets, some in manuscript form.
5. whether such extracted evidence can be interpreted by historians of reading both at scale and in detail [Gibbs and Cohen 2011, Towheed et al. 2015] to address questions such as the influence of the state and of censorship, the development of reading in educational contexts, the emergence of a European identity in 19th-century readers.

A preliminary analysis of use cases highlighted a challenging set of requirements to be considered in the modelling of READ-IT data model, such as the ability to study the embodiment of the reader in a physical context, the ability to identify the relation between reading and the personal life of the reader, the patterns emerging from the relation between the content, the type of sources and the socio-political context.

This examination of a variety of case study facilitates a dialogue between the ICT and DH scholars who create the underlying data model and the HSS researchers who adopt it [Flanders 2013]. This rigorous critique supports the development of a conceptual framework of reading, enables greater interoperability of data sources across different use cases, and allows scholars to address both macroscopic and microscopic questions [Hitchcock, Tim. 2014] in terms of geographical scale, time frame and types of sources which are at the core of READ-IT vision.

## References

- Flanders, Julia. 2013. 'The Literary, the Humanistic, the Digital: Toward a Research Agenda for Digital Literary Studies'. In *Literary Studies in the Digital Age*, edited by Kenneth M. Price and Ray Siemens. Modern Language Association of America. <http://dlsanthology.commons.mla.org/the-literary-the-humanistic-the-digital/>.
- Gibbs, Frederick W., and Daniel J. Cohen. 2011. 'A Conversation with Data: Prospecting Victorian Words and Ideas'. *Victorian Studies* 54 (1): 69–77.
- Hitchcock, Tim. 2014. 'Big Data, Small Data and Meaning'. 9 November 2014. [http://historyonics.blogspot.co.uk/2014\\_11\\_01\\_archive.html](http://historyonics.blogspot.co.uk/2014_11_01_archive.html).
- Kovač, Miha, and Adriaan van der Weel. 2018. 'Reading in a Post-Textual Era'. *First Monday* 23 (10). <https://doi.org/10.5210/fm.v23i10.9416>.
- Kuzmičová, Anežka. 2016. 'Does It Matter Where You Read? Situating Narrative in Physical Environment'. *Communication Theory* 26 (3): 290–308. <https://doi.org/10.1111/comt.12084>.
- Kuzmicova, Anezka. 2014. 'Literary Narrative and Mental Imagery: A View from Embodied Cognition'. *Style* 48 (3): 275-293.
- Murray, Simone. 2018. *The Digital Literary Sphere*. Baltimore: Johns Hopkins University Press. <https://jhupbooks.press.jhu.edu/content/digital-literary-sphere>.
- Towheed, Shafquat, Francesca Benatti, and Edmund G. C. King. 2015. 'Readers and Reading in the First World War'. *The Yearbook of English Studies* 45: 239–61. <https://doi.org/10.5699/yearenglstud.45.2015.0239>.