

Digital Data Handling at UWA

Sylvia Melzer^{1,2}, Stefan Thiemann¹, Ralf Möller²

¹Universität Hamburg & ²Universität zu Lübeck

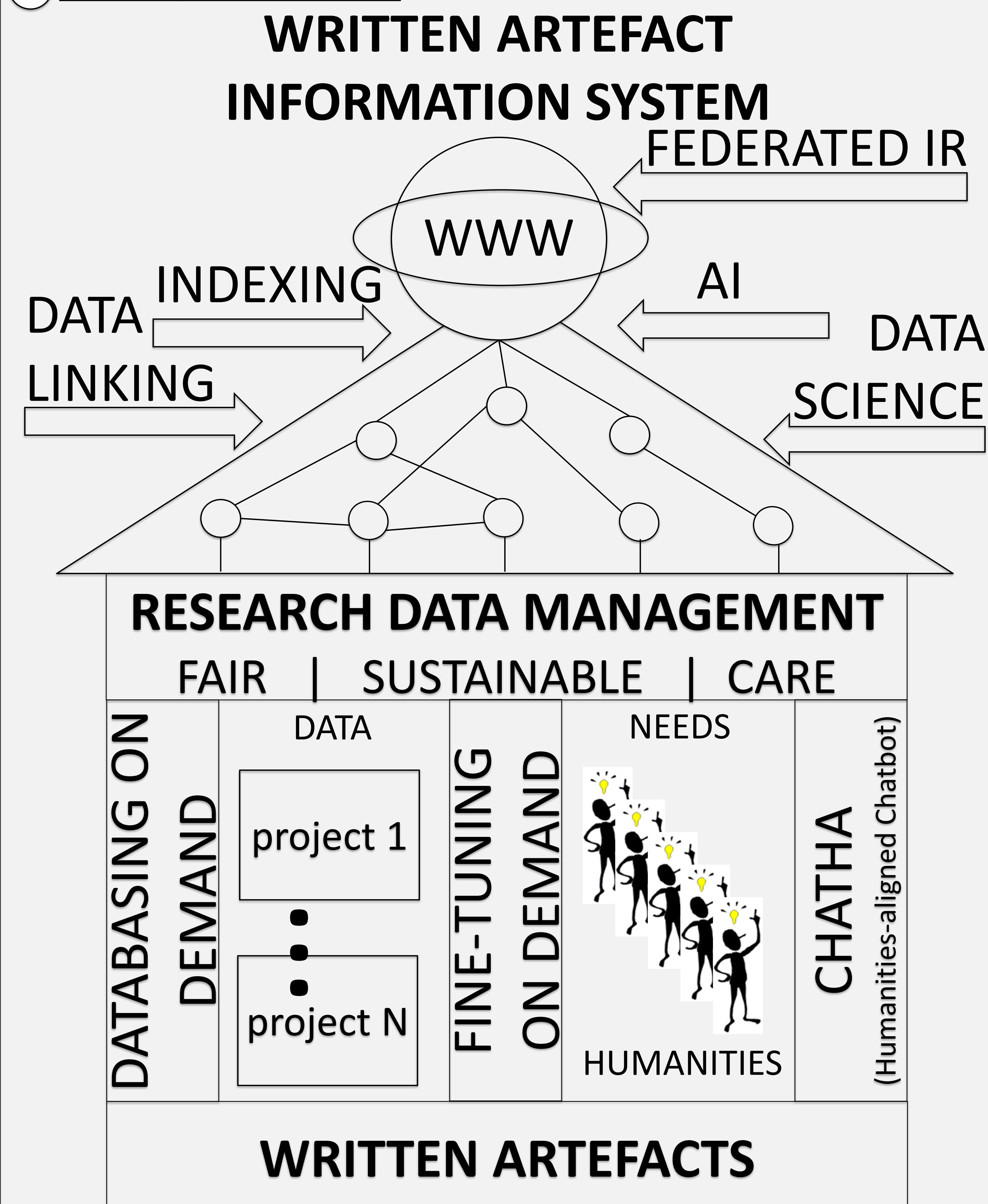
1 Abstract

Around 150 researcher, 40 academic disciplines, 60 research projects, and 11 research fields exist at the Cluster of Excellence UWA (Understanding Written Artefact). The size of the research field shows, it is a challenge to reconcile the many researchers with their different needs. The Research Field Data Linking has taken on the task of establishing research data management in the humanities in a sustainable, FAIR and CAREful way; developing IT approaches to support the reusability of research data and; making newly developed AI algorithms and other computer science approaches easy to use for humanities researchers; and adapting these algorithms to the research questions in the humanities.

2 Research questions

1. How can the management of research data in the humanities be integrated in a FAIR and CAREful way?
2. How can the research data be archived in such a sustainable way that the reusability of this data does not become a new challenge?
3. How can AI, data linking or other computer science approaches be integrated into humanities processes so that they can be applied by non-IT experts to generate new knowledge?
4. How can already collected humanities data generate added value in new IT approaches, such as Transformer models?

3 Developed Methods



4 KEY FINDINGS, IMPACT

- Archive research data in the research data repository (RDR) <https://www.fdr.uni-hamburg.de/> provides FAIR and sustainable data with CARE.
- We developed the generic approaches such as databasing on demand to make reusability not a new challenge for humanities scholars.
- New software for data processing are willing to accept by humanities scholars if it seamless integrates into the traditional research process.
- New AI approaches are accepted and by humanities scholars if they are easy to use and require minimal IT knowledge.
- If the data is stored in databases or in the RDR, these databases can be merged into a WRITTEN ARTEFACT INFORMATION SYSTEM to generate new knowledge, cf. <https://uhh-tamilex.github.io/bookbinding/>.
- We developed the fine-tuning on demand approach which uses research data from the humanities archived in the RDR to fine-tune transformer models.