

Information and Knowledge Organisation in Digital Humanities

Koraljka Golub and Ying-Hsang Liu

ASIS&T & DCMI Webinar, 16 December 2021

Contents

1. The role of knowledge organisation in DH
2. Examples
3. Looking ahead



1. Role of KO in DH



What is KO (Knowledge Organisation)?

- “...is about describing, representing, filing and organising documents and document representations as well as subjects and concepts both by humans and by computer programs” (Hjørland [2016a](#))
- The practices of KO are omnipresent in people’s lives, from the shopping list to the legal code
- In order to support the organisation of information resources, the KO field creates standards and guidelines to create representations of information objects such as those held by LAMs → metadata
 - Conceptual models (IFLA LRM, CIDOC-CRM, RiC)
 - Cataloguing standards (RDA, CCO, DACS)
 - Data values with controlled vocabularies (LCSH, DDC etc.)



KO

- A major discipline within Information Studies
 - Applications in numerous areas because the need to organise data, information or knowledge is omnipresent
 - However, in many domains of human endeavour, information is being organised *ad hoc*
- resulting in systems that underperform and even effectively prevent access to data, information and knowledge



KO -- a hidden theme in DH?

- In recent years, we have been witnessing a significant growth in the number of edited volumes in DH
 - E.g., digital methods, theoretical discussions of DH
- What about the field of KO in DH? What do we know about it?
 - What kinds of perspectives and approaches to organising information in DH do we have internationally?
 - Does DH resort to KO for best solutions?
 - Does KO successfully address information needs of DH?



Why KO in DH?

- *To help ensure that the best solutions are found for KO in DH, it is important to bring the two communities of research and practice together, to explore potential solutions and jointly address challenges*
- *Bringing together the understanding of the KO community and ambitions of the DH community can help provide strategic solutions to the challenge of managing information, thereby*
 - *supporting better information retrieval systems*
 - *improving the stewardship of data*
 - *creating more venues for information access, and*
 - *discovering more meaningful information.*



Other closely related fields

- Information Retrieval (IR)
- Information Behaviour
- Human Computer Interaction

→ All contribute to understanding the complexities surrounding KO, IR and information use and should be used complementarily to improve information access



2 major KO arenas in DH

1. Within cultural heritage (CH)
 - KO standards mentioned earlier
 - Adoption of standards to the Semantic Web
 - Making data FAIR; LOD
 - Metadata linking and aggregation
 - → making CH easily discoverable and openly available to all

2. Within academic research
 - XML to support KO at more granular levels than document-level common in LAMs: TEI markup for research in textual scholarship
 - Automated methods common for both KO and DH from early days
 - punch cards for library records and for Roberto Busa's early concordances; Luhn's KWIC index
 - Today: automatic topic identification both in LAM's and for textual scholarship
 - Research data and other research output curation



Contents

1. The role of knowledge organisation in DH
2. Examples
3. Looking ahead



2. Examples



“Information and Knowledge Organisation in Digital Humanities: Global Perspectives”

- Routledge, 2021
- Edited by Koraljka Golub and Ying-Hsang Liu
- Open Access
- <https://www.routledge.com/Information-and-Knowledge-Organisation-in-Digital-Humanities-Global-Perspectives/Golub-Liu/p/book/9780367675516#>



TOC

- An open call for proposals, followed by a review and selection process
- Final 12 chapters and 1 introductory chapter
- PART I. Modelling and Metadata (6 chapters)
 - Modelling cultural heritage data
 - Conceptual models
 - Metadata aggregation
 - Metadata enrichment
 - Need to move from organising data to organising knowledge
- PART I is the largest part of the book, reflecting the fact that metadata is the dominant area of research within KO for DH



- PART II: Information Management (3 chapters)
 - Management of in-copyright texts
 - Management of lexicographical resources
 - Management of DH research outputs
- PART III: Platforms and Techniques (3 chapters)
 - Specific platforms needed to support DH research
 - Recogito GIS platform
 - Data analysis techniques
 - Browsing visualisation techniques for user interfaces to cultural heritage collections



International & cross-sectoral

- International: 41 authors affiliated with organisations in 16 countries on 4 continents: Asia (China, Israel, Japan, Sri Lanka), Australia, Europe (Belgium, Croatia, France, Germany, Greece, Norway, Portugal, Sweden, Switzerland, United Kingdom) and North America (United States of America).
- Cross-sectoral:
 - 8 authors are affiliated with a cultural heritage institution, a heritage board and the European Commission
 - 2 authors are IT developers



But only a snapshot...

- Only a fraction of international perspectives on a sample of KO DH topics
- Themes that could be more represented:
 - Automated methods and techniques such as entity linking in natural language processing or deep learning models for semantic representations
 - User perspectives in user interface design and in evaluation
 - In general, *more of humanities-driven KO research* rather than vice versa
 - Perspectives from Africa and South America are sadly absent



Why this book? For whom?

- To address the gap between KO and DH
 - A guide for university teachers, researchers and working professionals interested in the role of KO in DH
 - Each chapter = 1. an introductory overview of the topic under discussion + 2. a case study + 3. reflections and suggestions for future work
- For those who work with DH but are unfamiliar with KO and vice versa
- A starting point for discussions on how KO best interacts with DH



In summary...

- The chapters provide snapshots of how information can be organised in various contexts in DH
- Organising CH in digital environments is addressed in topics such as the creation and adoption of conceptual models and metadata standards (2–6), the incorporation of LOD (4-5), ways of enriching metadata (5 and 6) and the aggregation and interoperability of metadata across cultural heritage collections (2–6)
- Managing DH resources and DH documents for preservation and reuse (8–10) and examples of (semi-)automated approaches to support KO for improved access, discovery and navigation of materials (5, 8, 11–13)
- An invitation to consider the nature of knowledge production within the humanities, and to actively work towards representing fundamental epistemic elements such as uncertainty, interpretation, context and narrative rather than import data-based KO from systems and technologies adopted from outside the field (7)



Contents

1. The role of knowledge organisation in DH
2. Examples
3. Looking ahead



3. Looking ahead



Two core areas of KO for DH...

1. Information discovery (search and access)
 - Insufficient adoption of KOS in IR, leading to suboptimal IR for humanities
 - Insufficient evaluation of semantically enriched CH data and LOD in IR contexts
 - Insufficient evaluation of user interfaces
 - DH scholars' information needs have to be thoroughly and continuously researched to inform all of these KO processes, models, standards and guidelines as well as IR → user studies and participatory KO

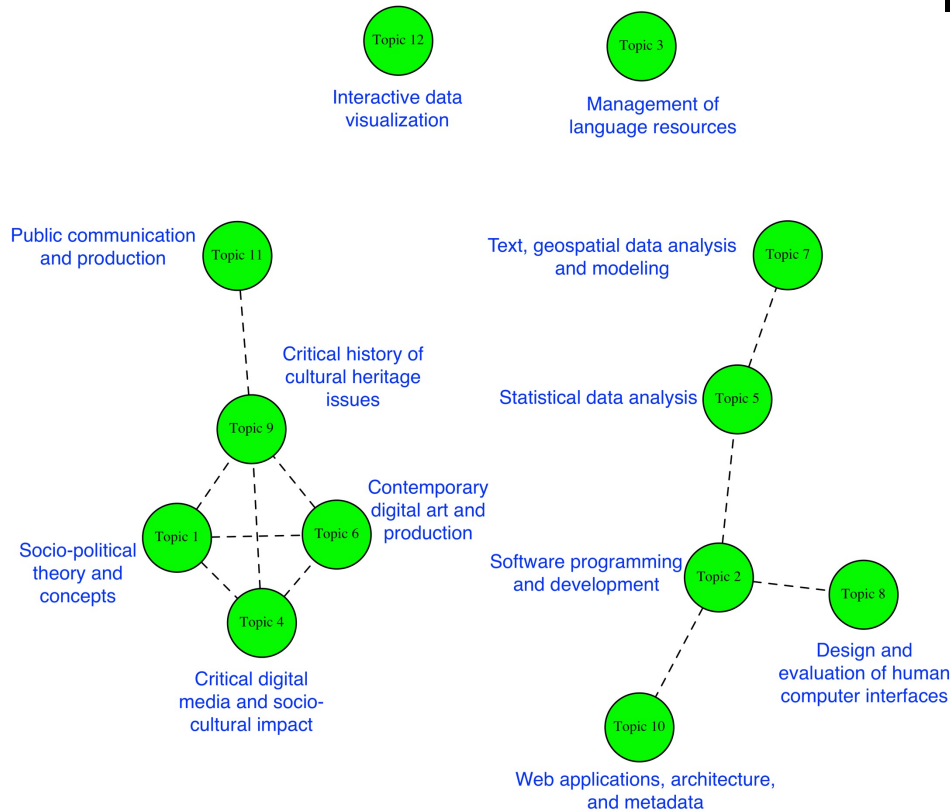


2. Information representation

- Essential in KO (formal and subject metadata)
 - Practical endeavour but also philosophical (e.g. Olson)
- Wake-up call by Oldman: DH research practices and the ways humanities scholars interact with their sources do not match with data-based databases devoid of context
- Interoperability requires harmonized conceptual models that at the same time need to cater for special object types and different granularity levels
- In turn, conceptual models need to be reflected in guidelines, which should be implemented by KO professionals addressing DH needs
- Need for criticality when applying automated methods -- limited by knowledge representations hard-wired into the automated technologies



Force-directed layouts for visualizing topic models



Interdisciplinary Research

- Digital Humanities (DH) course descriptions analysis by iDHCC (Walsh et al., 2021)
- Where is KO within DH?
 - Web applications, architecture and metadata
 - Design and evaluation of human computer interfaces
 - Management of language resources
 - Interactive data visualization

Walsh, J. A., Cobb, P. J., de Fremery, W., **Golub, K.**, Keah, H., Kim, J., Kiplang'at, J., **Liu, Y. H.**, Mahony, S., Oh, S. G., Sula, C. A., Underwood, T., & Wang, X. (2021). Digital humanities in the iSchool. *Journal of the Association for Information Science and Technology*. <https://doi.org/10.1002/asi.24535>



Concluding remarks

- The full potential of information discovery and information representation in the field of DH cannot be realised without **genuine transdisciplinary collaborations** that actively engage with the development of DH technologies, practices and outputs
- Calling for a future in which **DH research is more interdisciplinary, cutting across KO, IR, HCI, IB and other related fields and disciplines.** We need to harness these complementary perspectives in order to provide the best, **evidence-based KO solutions** which address the complexities of DH research and, in turn, feed back into KO research.



Contents

1. The role of knowledge organisation in DH
2. Examples
3. Looking ahead



References

- Golub, K. & Liu, Y.-H. 2021. Information and Knowledge Organisation in Digital Humanities: Global Perspectives. Routledge. Available at <https://www.routledge.com/Information-and-Knowledge-Organisation-in-Digital-Humanities-Global-Perspectives/Golub-Liu/p/book/9780367675516#>
- Golub, K., Kamal, A. M. & Vekselius, J. 2021. Knowledge Organisation for Digital Humanities: An Introduction. In Golub & Liu (Eds.), Information and Knowledge Organisation in Digital Humanities: Global Perspectives. Routledge. Available at <https://www.routledge.com/Information-and-Knowledge-Organisation-in-Digital-Humanities-Global-Perspectives/Golub-Liu/p/book/9780367675516#>
- Hjørland, Birger. 2016a. "Knowledge organization (KO)." *Knowledge Organization* 43 (6): 475–85. Also available in *ISKO Encyclopedia of Knowledge Organization*, edited by Birger Hjørland and Claudio Gnoli. https://www.isko.org/cyclo/knowledge_organization.
- Walsh, J. A., Cobb, P. J., de Fremery, W., Golub, K., Keah, H., Kim, J., Kiplang'at, J., Liu, Y. H., Mahony, S., Oh, S. G., Sula, C. A., Underwood, T., & Wang, X. (2021). Digital humanities in the iSchool. *Journal of the Association for Information Science and Technology*. <https://doi.org/10.1002/asi.24535>



Thank you for listening

- Any questions?
- Any concrete suggestions or examples on how to further bring KO and DH together?
- Would you like to share any experiences of successful or less successful KO in DH?



i | **ilnstitute**



Lnu.se