Collection Computing, Computational Collections. Data as performative knowledge and live heritage

Chris Tanasescu (Margento)

UCLouvain

ADOCHS KBR 2021
DATA—THE PREMISE & THE CONCLUSION; DATA = “MORE LINKS”

• Data come in clusters: “a mass noun.” (OED)

• There is a rhetoric of data and data visualization that develops on more traditional visual rhetoric and multimodal writing (cf. Gallagher and DeVoss 2020). Data are always cooked, no such thing as “raw data” (Gitelman 2013), there is a semantics as well as a rhetoric informing data in general.

• Data do not just exist, they are generated (Manovich 1998).

• Data are a combination of what is “given” and the ways in which the latter is organized. Data = ways of reasoning. Euclid’s use of the term data involves both the clear and comprehended premise, and the demonstrated conclusion (Taisbak cited in Lavin 2021).

• “[T]he Givens hang together in chains, the purpose of any proposition being to produce more links to them.” (ibid.)
GIVEN VS TAKEN. DATA VS CAPTA

• Johanna Drucker advanced the term *capta* as more appropriate than data; the presumably given actually needs to be taken (Drucker 2011).

• Print collections = Big Data + Great Unread; data vs. capta is most relevant in the context; and so is the form[at] issue…

• Yet please refer also to the recent critique of Drucker’s concept. E.g., data are not only the given, they are “situated knowledge” (Lavin 2021).

• That situatedness is rhetorical, political (the DH approach) or (non)tractable (Comp Sci);

• In Comp Sci therefore there is such a thing as “raw data.”

• And yet, the DH contribution to Comp Sci/AI is the question of ethics with regards to data and AI (cf. Tobias Blanke 2020).
DATA-CAPTA: INTERMEDIAL, NETWORKED, PERFORMATIVE. LIVING/LIVE ARCHIVES

- The situatedness of data is multi-faceted; overarchingly epistemological; and therefore, data are situated and/as intermedial [in the acceptance of Larrue and Vitali-Rosati 2019]; from whence the challenge of digitization and the relevance of projects such as ADOCHS.
- Data since intermedial are networked (producing ever “more links”) and performative.
- In a previous KBR-ULB-UGent talk part of the Digital Heritage Seminar (Digital Research Lab & Camille), I talked about live archiving and living/live archives (Tanasescu 2021).
- There are archives and collections that contain live material but also collections that are assembled and/or per/re/de-formed live and that consequently turn out (a)live themselves (cf. MARGENTO, #GraphPoem @ DHSI 2021, Rui Torres, ARQUIVO VIVO É ANARQUIVO!)
#GraphPoem: Data-commoning Webformance

- #GraphPoem (started back in 2010 and applying [network] graph theory in poetry) has a performance component involving interactive coding and dataset assemblage: data-commoning (Tanasescu et al. 2020, Tanasescu and Tanasescu 2022).

- Commoning refers to community in/as performance (Tanasescu 2016). In digital space (the term as defined by Kennedy 2015), performance = data as intermediation = community as action. Hence data-commoning.

- Data-commoning involving subversive approach to web-based corporate monopoly and/or control, thus de/re-forming the Web by performing community-based webs: webformance (Tanasescu et al. 2020, Tanasescu and Tanasescu 2022).

- Digitization thus involves the [re]digitalization of communities by means of, and as, data.
Collection Computing; Performative Knowledge

• In events like #GraphPoem or ARQUIVO VIVO É ANARQUIVO!, data-capta are therefore “situated” on multiple levels (commoning, ‘our’ webs vs the Web, intermediality [and hyper-platformity], etc.);

• Our redefinition of data-capta in digit(al)izing and intermedial contexts thus has to evolve from “situated knowledge” (Lavin 2021) to “performative knowledge;”

• Collections [files, users, AI] are digit(al)ized & computed (assembled and analyzed) by means of, and as part of, the data-commoning webformance; within the digit(al)ization process this is the collection computing part.

• In “US” Poets Foreign Poets (MARGENTO eds., 2018, Fractalia Press) collection computing consisted of performing corpora as networks by taking advantage of electronic literature’s intermediality and then computationally analyzing the networks by resituating print and Jupyter Notebooks as inextricably and dynamically conjoined within such projects.

Chris Tanasescu @ ADOCHS KBR 2021
Computational #GraphPoem Performances


DHSI E-Poetry Event #GraphPoem, by MARGENTO

All those connected to DHSI and its 2021 edition are invited to be part of the EPoetry event #GraphPoem by MARGENTO at 9:30 AM Pacific Time on June 11 by contributing text files or weblinks to a collectively assembled dataset and/or run a script plotting the latter into a real-time evolving network. The Graph Poem is an ongoing transnational project combining natural language processing and graph theory-based approaches to poetry, with academic, DH-literary, and performative outputs.

#GraphPoem will have two main components viewable to anybody accessing the following online venues at the time of the event: a livestreamed performance on Margento’s Facebook page and the bot @GraphPoem tweeting text-nodes selected from the evolving graph by a network analysis algorithm and fed into the performance.
#GraphPoem: Computational Collections

- In a #GraphPoem performance the intermediality of data works in a hyper-platform way: JupyterHub collective coding feeding into bot content on Twitter feeding into Facebook livestreamed jam sessions and back;
- In the process, collection computing turns out to be made possible by a computing collection. That is, it is the collection that computes; the collection computes the collection.
- Data therefore not only get digit(al)ized, but also softwarized. The up and coming paradigms of Softwarized Networks, Edge Computing, and Network Intelligence are already paving the way to digit(al)ization as softwarization.
- In a post-human (data-commoning webperformative) context, networked data (“graph poems”), humans, and AI overlap (Tanasescu et al 2020). Data are therefore the other in the process of becoming ourselves [HCI networks] as we articulate them further and further into otherness.
- The advancement into ‘otherness’ is performed by the other per se, the softwarized collection expanding with ever “more links”…
The H Shaping the D within DH; the Literary as the D...

• **DH** as **digital humanities** but also **data/digital heritage** (Tanasescu 2021a).

• The humanities and literature specifically can prove quite relevant to data/digit(al)ization and AI more widely. E.g. NLP tools developed for (or trained on) poetry data can prove the best for similar tasks on any other kinds of data (cf. Tanasescu 2018).

• **Literature** can be seen as **the** digital genre or even as **the digital** per se (Tanasescu 2021b).

• Addressing the automation of “hermeneutic inventio” (Ramsay cited in Tanasescu 2021b) results in advances in NLP generally. So does developing “digital hermeneutic visualization” (Kleymann & Stange 2021). Network (data) visualization can involve the actual combo of close and distant reading, and of quantitative and qualitative analysis (Birkholz & Budke 2021).
Generat(l)ive Collections; Live Heritage

• In both “US” Poets Foreign Poets [the ‘book’] and #GraphPoem [the performances] or Torres’s above-cited artist’s interventions the [performativity of the] collection computes and thus generates the collection (“performativity” again in Larrue & Vitali-Rosati’s acceptance, yet applied to performance as well). Softwarized networked corpora expand automatically according to the algorithms they are informed by.

• While locating, tapping into, and digitizing, items of various heritages, the softwarized collection per/re/de-forms and thus (re)generates and enhance the latter, which thus emerge as live heritage. Since AI-ed and automatically expanding, such heritage behaves as “technical organism[s]” (Klobucar 2021, Stefans 2017, Tanasescu et al. 2020) capable of the continuation of life “by means other than life” (Simondon cited in Stefans).

• Heritage (a)live therefore represents data as performative knowledge.

Chris Tanasescu @ ADOCHS KBR 2021
Selected References


- Blanke, Tobias. 2020. "Researching the Digital Human—Where are the Humanities in Artificial Intelligence?" Video, https://www.youtube.com/watch?v=mcR-iHz-m5E&t=1s.


- Tanasescu, Chris; Burny, Nicolas; Vanderdonckt, Jean; Inkpen, Diana; Kesarwani, Vaibhav. 2021. “Closeness and Betweenness Centralities for Multiplexes and their #GraphPoem Applications.” Presentation at DS2 Conference, UCLouvain. Video. https://www.youtube.com/watch?v=25b02xj-74


Chris Tanasescu @ ADOCHS KBR 2021
Thank you 😊!

- Q & A
- #GraphPoem @MARGENTO_
- @DHUCLouvain1 [https://sites.uclouvain.be/chairealtissia/]