UNIL | Université de Lausanne
Humanités numériques

computational humanities

Digitized, digitalized, or digitally transformed humanities?

Michael Piotrowski
<michael.piotrowski@unil.ch>
🐦 @true_mxp

HCH19, Heidelberg, July 17, 2019

Computational humanities

> *"Digital humanities" is simply a formula that has come to identify a large field. I use it only because everybody uses it and everybody will use it. But, frankly, I don't like it. I think it means nothing, whereas "quantitative" and "computational" mean something.*
>
> **–Franco Moretti** *(Hackler and Kirsten 2016, p. 5)*

► So, I guess he'd applaud the choice of "computational humanities" for this summer school…

► … unfortunately, actual usage doesn't differ much from that of "digital humanities."

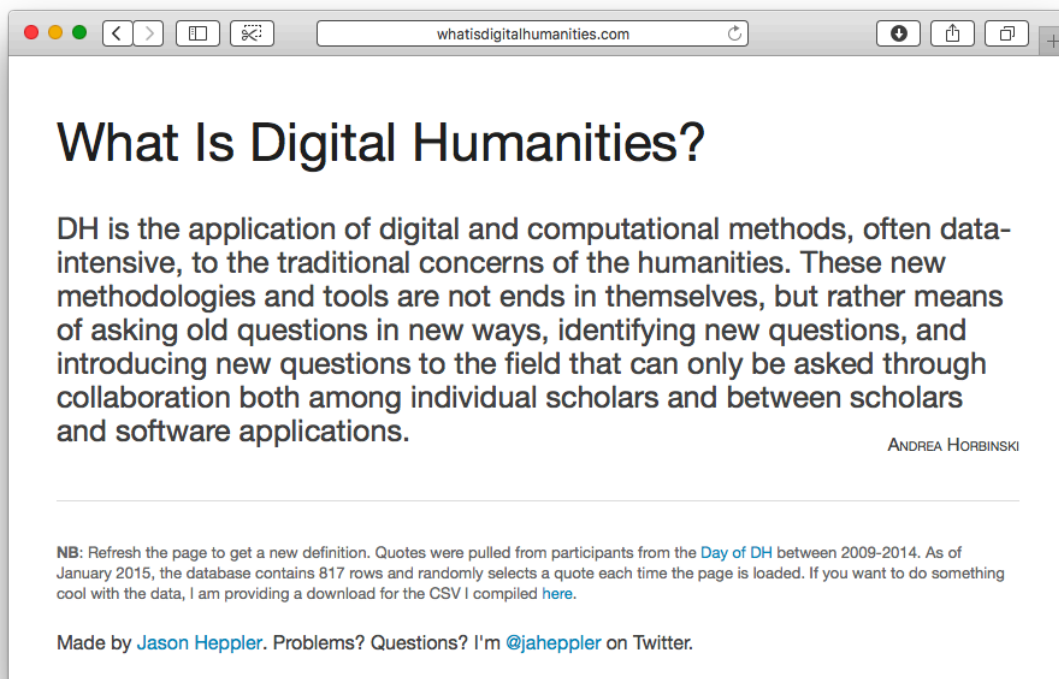# Computational humanities

Random example:

> *[...] is an emerging area of research that refers to the investigation of humanities, arts, and social science research questions through advanced computing technologies. A social scientist who **analyses a large dataset** of census information using a supercomputer is engaging in [computational humanities], as is a historian who investigates **historical texts** using **visual analytic software**. The use of computational methods is more than a bridging of disciplines – it is an increasingly necessary methodological approach for researchers who wish to investigate the **numerous human activities that result in digital information**. [ichass.illinois.edu]*

Whatever you call it...

- ► Is it about "big data"?
- ► Is it about scanning and digitizing?
- ► Is it about visualizations?
- ► Is it about "using tools"?
- ► Is it about studying digital traces?
- ► Is it about accelerating research in the humanities?
- ► Is it the "humanities for the 21st century"?
- ➜ **Digitized, digitalized, or digitally transformed humanities?**

Je sais même pas ce que ça veut dire...

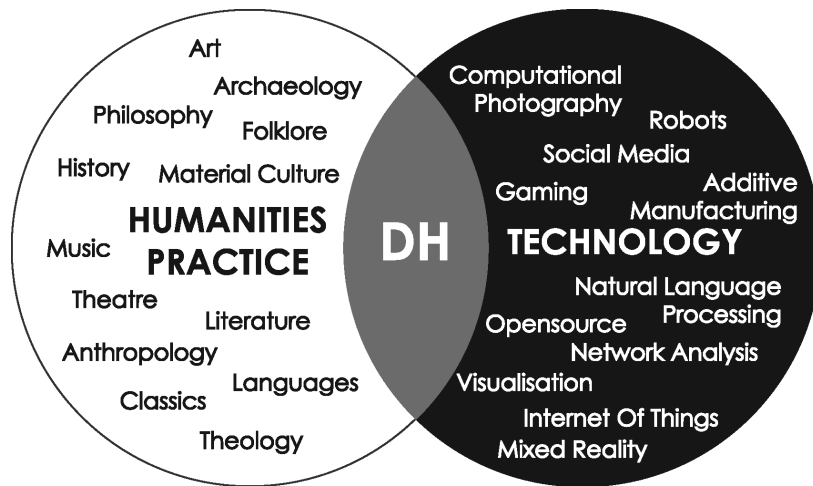# WhatIsDigitalHumanities.com

What Is Digital Humanities?

DH is the application of digital and computational methods, often data-intensive, to the traditional concerns of the humanities. These new methodologies and tools are not ends in themselves, but rather means of asking old questions in new ways, identifying new questions, and introducing new questions to the field that can only be asked through collaboration both among individual scholars and between scholars and software applications.

ANDREA HORBINSKI

NB: Refresh the page to get a new definition. Quotes were pulled from participants from the Day of DH between 2009-2014. As of January 2015, the database contains 817 rows and randomly selects a quote each time the page is loaded. If you want to do something cool with the data, I am providing a download for the CSV I compiled here.

Made by Jason Heppler. Problems? Questions? I'm @jaheppler on Twitter.

It's a gray area!?

Humanities Practice: Art, Archaeology, Philosophy, Folklore, History, Material Culture, Music, Theatre, Literature, Anthropology, Languages, Classics, Theology

DH

Technology: Computational Photography, Robots, Social Media, Gaming, Additive Manufacturing, Natural Language Processing, Opensource, Network Analysis, Visualisation, Internet Of Things, Mixed Reality

Credits: doi:10.1093/llc/fqy080

Arrogance is not helpful

Ian Bogost ✔ @ibogost   21 Jun
Replying to @berrydm
Maybe you can unbury the lede and tell me here what DH is.

David M. Berry   🐦 Follow
@berrydm

I don't think many people in DH care about definitions too much now. Thankfully the debates have moved on.

4:24 PM - Jun 21, 2017

💬 2   🔁   ♡ 13   ⓘ

DIGITAL HUMANITIES? JUST DO IT.

Identity crisis?

*Despite all this enthusiasm, the question of what the digital humanities is has yet to be given a satisfactory answer. Indeed, no one asks it more often than the digital humanists themselves. The recent proliferation of books on the subject–from sourcebooks and anthologies to critical manifestos–is a sign of **a field suffering an identity crisis**, trying to determine what, if anything, unites the disparate activities carried on under its banner. [Kirsch 2014]*

Ongoing identity crisis!?

At this moment, some humanities computing seems to be in a state of free-fall. It's tool-driven, a bad way to be. Better, be situation-driven.

Credits: Kenner [1989]
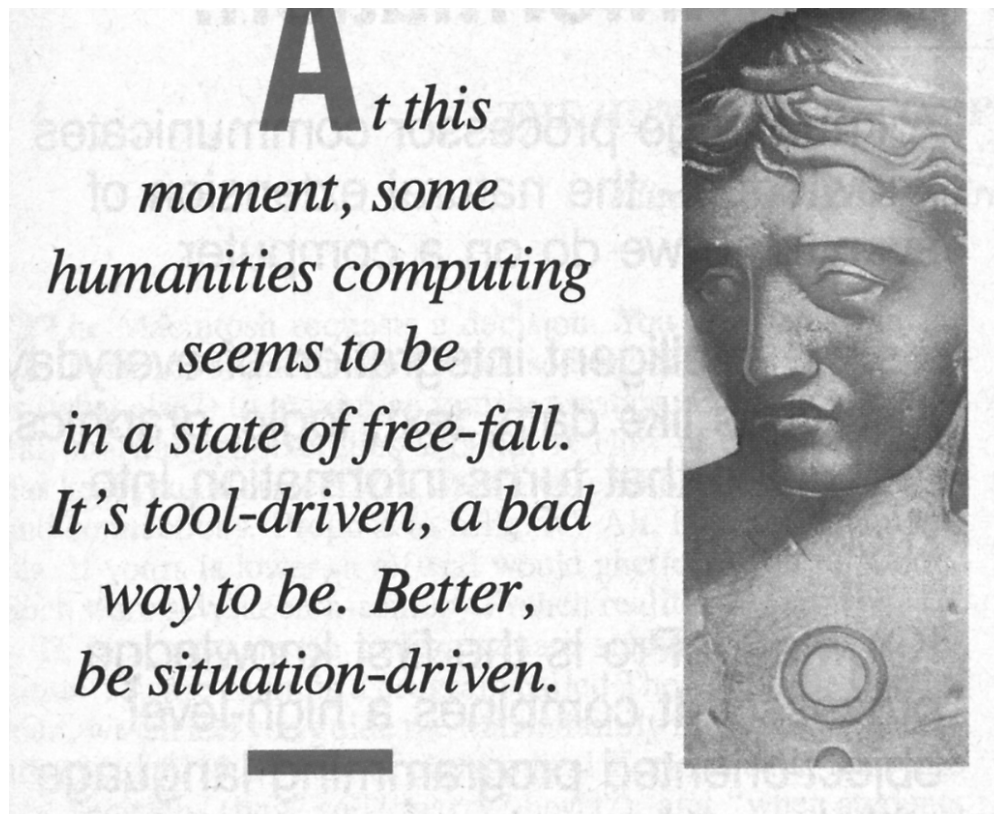
Computer science and the humanities

*Dès l'origine, **deux conceptions** se sont établies quant à la nature de ces rapports [entre l'informatique et les sciences de l'homme]. **La première, qui est encore la plus courante, les envisage surtout sous un angle opératoire et il est vrai que l'ordinateur est un instrument capable de modifier profondément les conditions matérielles dans lesquelles s'effectue la recherche dans les sciences de l'homme. L'autre point de vue lie cette dimension technique à son socle conceptuel (logico-mathématique) et voit dans le recours à l'informatique un facteur susceptible de faire évolue également les cadres théoriques et les référents épistémologiques des science de l'homme.** (Borillo 1985, pp. 5–6]*

Thinking machines

- ► Why this dualism?
- ► Unlike all previous machines, computers aren't intended to alleviate physical, but **mental** work.
- ► From the beginning, computer were often associated with thought ("electronic brains").
- ► Human thought is often considered the domain of the humanities.
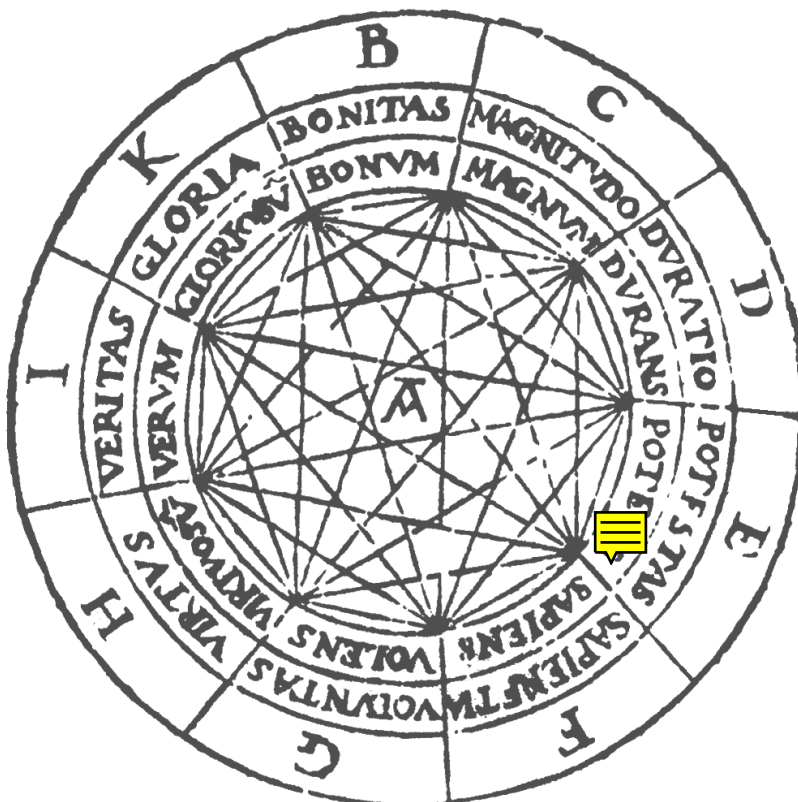
Thinking machines

Thinking machines

► Of course the title plays with the ambiguity–clearly it's not about "machines that think" but "machines for thinking."

► "Nous ne raisonnons que sur des modèles." (Paul Valéry?)

➔ **A "thinking machine" must be a "modeling machine."**

► In fact, computers were **created for the sake of modeling**: in the beginning primarily for mathematical models (in particular for ballistics), but already their use in cryptanalyis (Colossus vs. Lorenz code) anticipates a more general manipulation of symbols (↗ Llull, Leibniz).
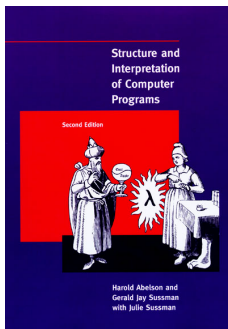
Ramon Llull: Ars magna

Credits: Ramon Llull, **Ars generalis ultima** (ca. 1305). Public Domain, Wikimedia
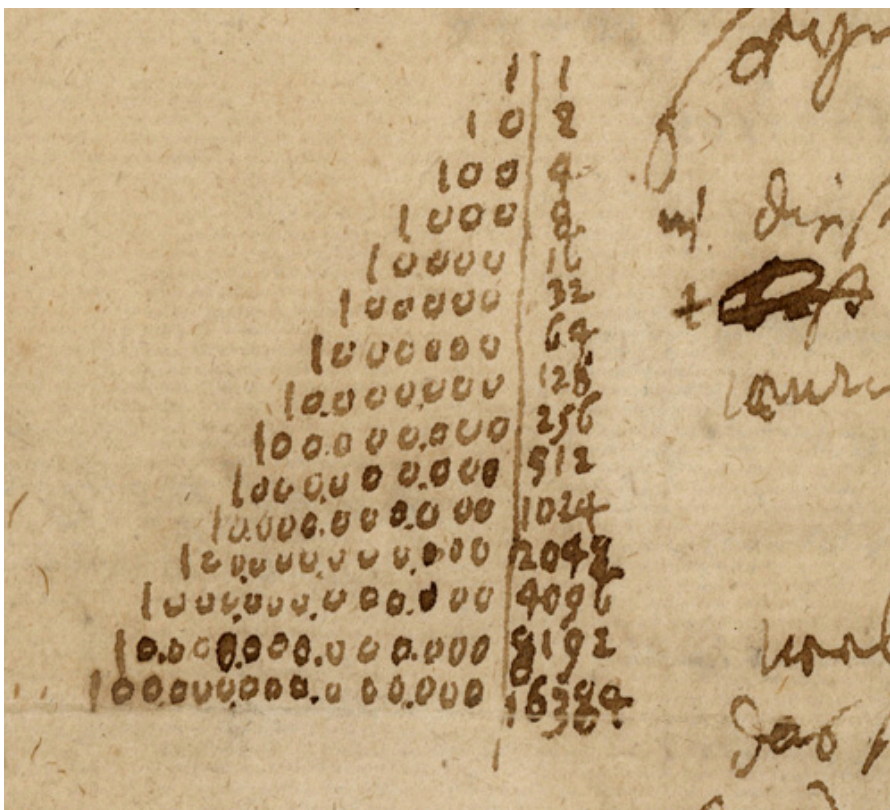
Thinking machines

Leibniz

► Leibniz's idea of a logic-based universal language (characteristica universalis) was directly inspired by Llull (Ars combinatoria, 1666).

► Idea: all concepts can be represented by symbols, by applying formal rules new concepts can be derived or generated mechanically.

► Leibniz went beyond Llull: Symbols → digits → numbers.

► Aimed to demonstrate that thinking can be formalized and thus mechanized—and that logical thinking and calculating are identical.

► In the future, theologians and philosophers would no longer have to fight (↗ Llull), but a disagreement between philosophers could be resolved like one between accountants:

quando orientur controversiae, non magis disputatione opus erit inter duos philosophos, quam inter duos computistas. Sufficiet enim calamos in manus sumere sedereque ad abacos, et sibi mutuo (accito si placet amico) dicere: calculemus

Leibniz

- ▶ These thoughts were the foundation for Leibniz's development of the binary system.

- ▶ He realized that digits could also represent other digits, and that 0…9 could be represented by combinations of 0 and 1.

- ➡ Decimal numbers, consisting of several digits, can be represented by a string of 0 and 1.

- ▶ The binary system is the basis for digital computers.

- ▶ "New Year's letter" to Count Rudolph August 1697: connection between mathematical structures and theological-philosophical interpretation of the world: "**omnibus ex nihilo ducendis sufficit unum**"–the world was created from nothing (0) and God's Word (1).

Leibniz


Credits: http://dokumente.leibnizcentral.de/index.php?id=54

Thinking machines

► After WWII, computers were quickly adopted by (more or less) civilian research, in particular in the natural and engineering sciences, which had already been using mathematical models.

► And the humanities? Contrary to popular belief, computers have long been used in humanities research:
  ↗ Computational linguistics
  ↗ Computational philology → Roberto Busa

► Busa–jesuit and philologist–may have been the first, but certainly not the only one who recognized the possiblities computers offer to the humanities (↗ e.g., Bowles 1967).

Roberto Busa
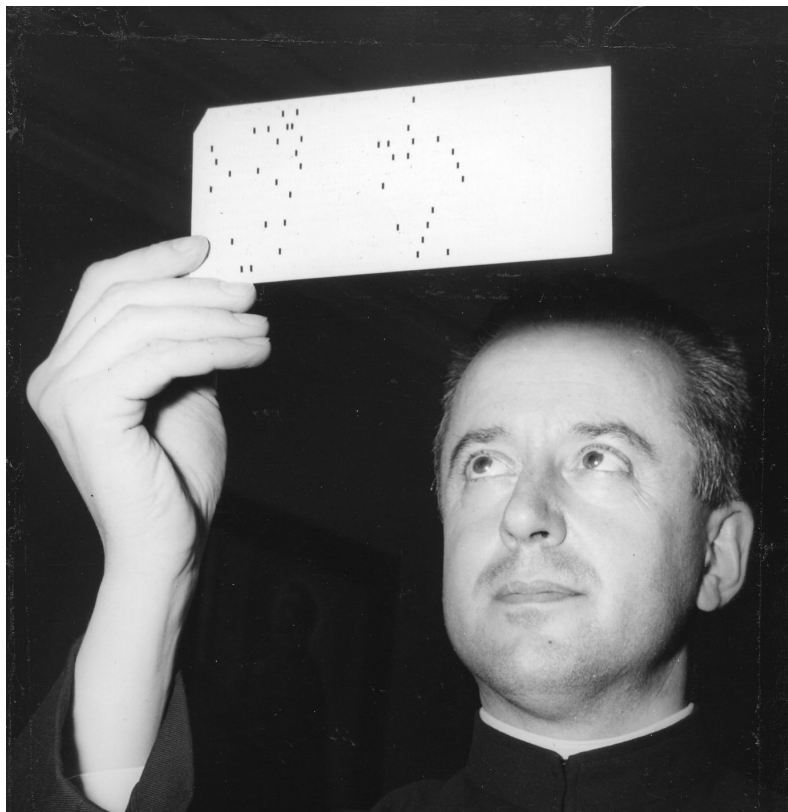
Roberto Busa (1913–2011)

Thinking machines

► The first humanities applications were primarily mechanical (e.g., sorting) and quantitative (e.g., frequency calculations).

► At some point, however, Busa realized that it's **not** just about making things easier:

> *In this field one should not use the computer primarily for speeding up the operation, nor for minimizing the work of the researchers.* **It would not be reasonable to use the computer just to obtain the same results as before, having the same qualities as before, but more rapidly and with less human effort.** *[...] To repeat: the use of computers in the humanities has as its principal aim the enhancement of the quality, depth and extension of research and not merely the lessening of human effort and time. (Busa 1980, p. 89)*

Traditional research process

| Sources | | Working Materials | | Narrative |
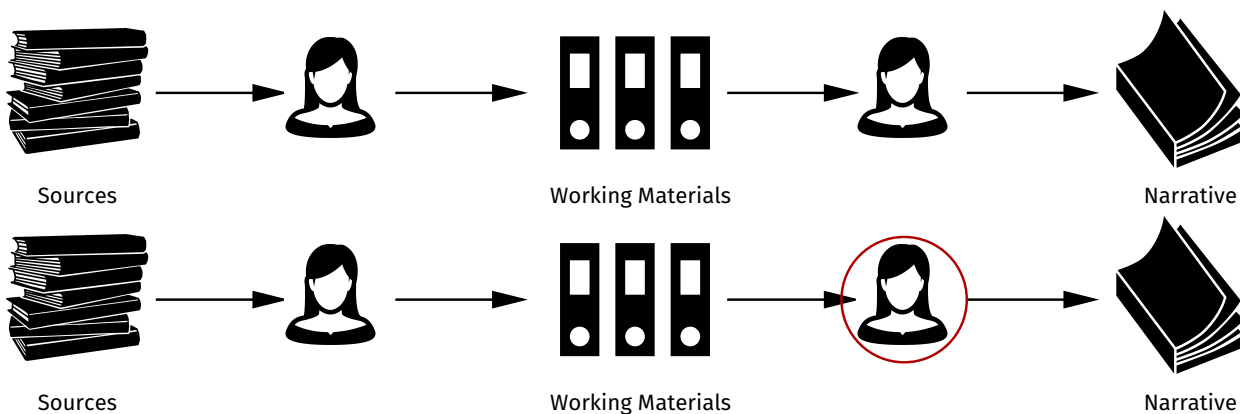
► Scholar reads and interprets primary and secondary **sources**

► Facts and insights are recorded as **working materials** in a variety of forms (on paper or electronically, as text, in spreadsheets, databases, etc.)

► Using the working materials, scholar constructs **mental model** to answer research question and describes the model in a **narrative**.

Models in the humanities?

► **All** scientific and scholarly research constructs models of their objects of research in order to understand phenomena that are too complex or not directly accessible to observation.

► In contrast to the natural sciences, models in the humanities are traditionally
  ► informal
  ► internal
  ► published as narratives

What do we mean by "model"?

We use the term in the sense of Stachowiak's **Allgemeiner Modelltheorie** ('General Model Theory,' Stachowiak 1973).
Three fundamental properties of models:

Mapping property Models are always **models of something** (no matter if natural, artificial, concrete, abstract, …)

Reduction property Models generally **do not capture all attributes** of the original they represent, but only those that seem relevant to the model creators and/or model users.

Pragmatic property Models are **not per se uniquely assigned** to their originals. They fulfill their replacement function
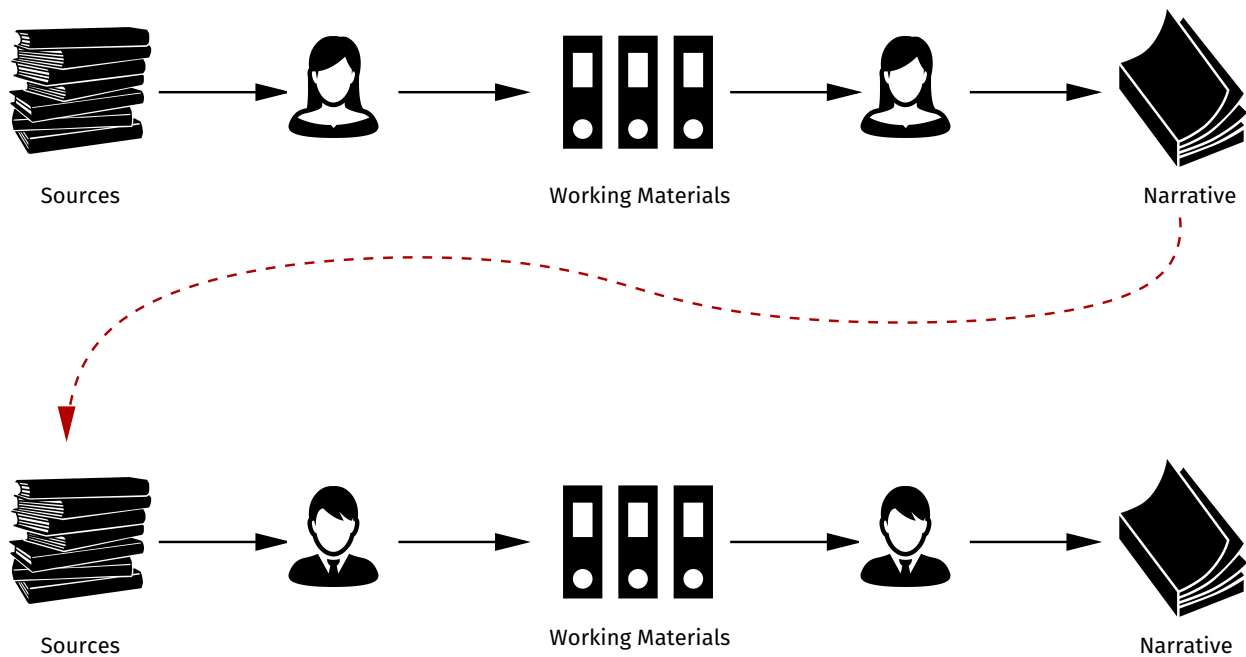 ⓐ for particular subjects,
  ⓑ within particular time intervals, and
  ⓒ restricted to particular mental or actual operations.

Original–model mapping

preterite attributes

abundant attributes

domain

attribute mapping

codomain

original
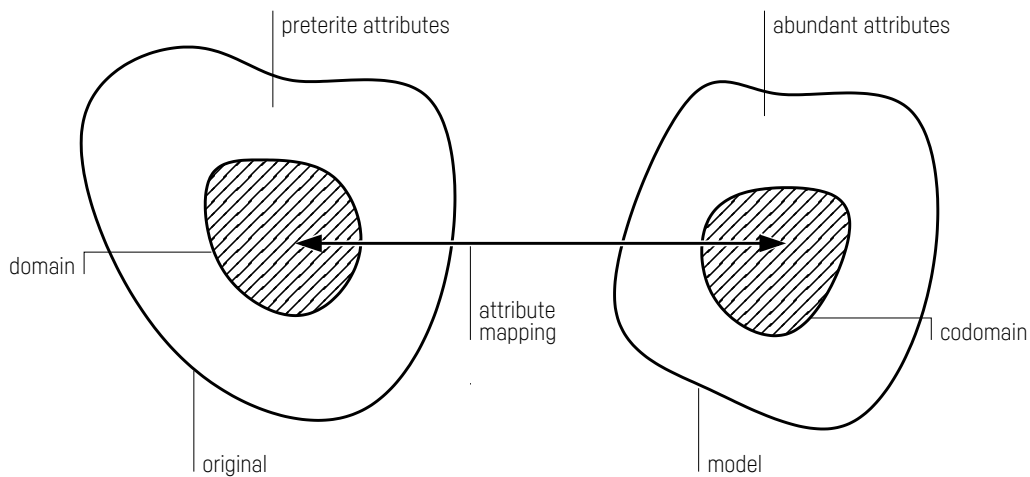
model

Building on the work of others (traditional process)

Sources

Working Materials

Narrative

Sources

Working Materials

Narrative

- ▶ Model behind others' narratives must be reconstructed hermeneutically
- ▶ **Problem:** limited intersubjectivity

Collaboration on a higher level through formal models

▶ Telepathy is unfortunately impossible!
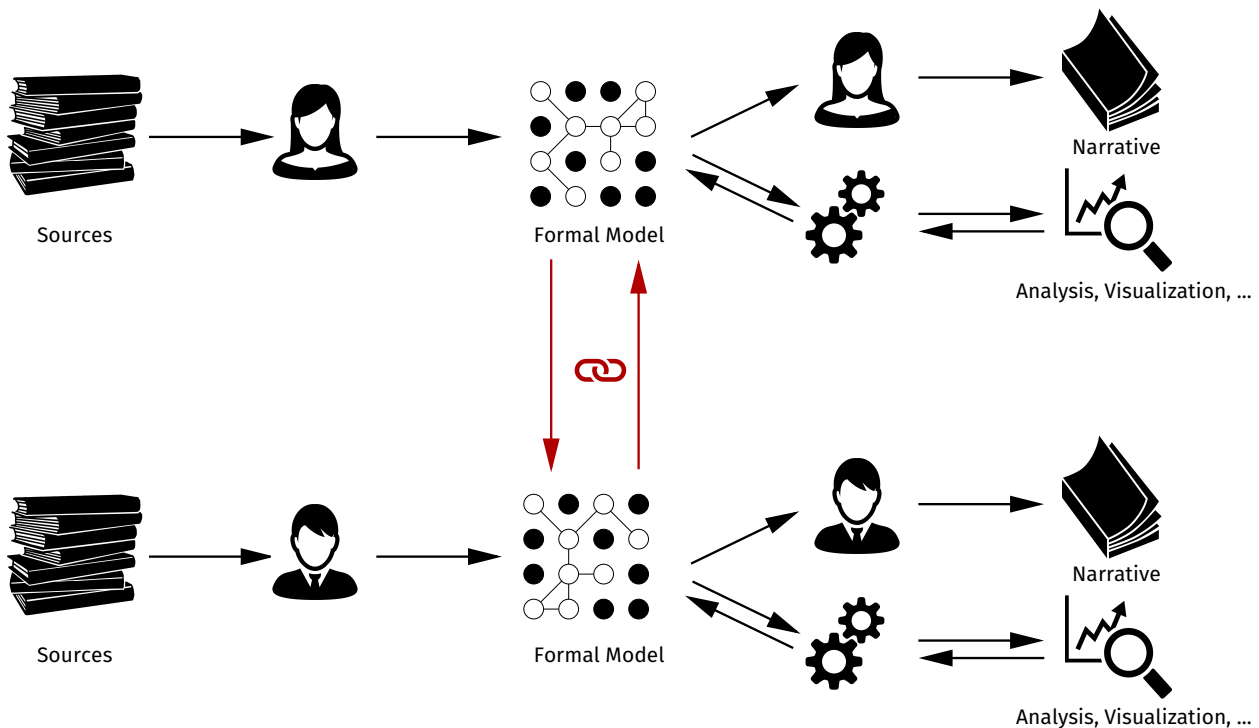
What do we mean by "formal"?

Слово «формальный» не означает ничего, кроме как «логически последовательный + однозначный + абсолютно явный». [Gladkij and Mel'čuk 1969, p. 9]

The word "formal" means nothing more than "logically coherent + unambiguous + explicit."

▶ A formal model is thus more than explicit.

▶ There are different degrees of formalization.

▶ In digital humanities, we are primarily interested in a degree of formalization that allows models to be **processed and manipulated by computers**.

➜ Computational models

Computational models and theory formation

Joseph Weizenbaum in **Computer Power and Human Reason** (1976):
*The connection between a model and a theory is that a model **satisfies** a theory; that is, a model obeys those laws of behavior that a corresponding theory explicitly states or which may be derived from it. […]*

Computers make possible an entirely new relationship between theories and models. *[…] Theories written in the form of computer programs are ordinary theories as seen from one point of view. […] **A theory written in the form of a computer program is thus both a theory and, when placed on a computer and run, a model to which the theory applies.*** [Weizenbaum 1984, pp. 143–145]*

Piotrowski 2016

Definition (Applied digital humanities)

Research in the humanities using formal models, and the methodology of constructing such models.

➜ Digital history, digital philology, digital musicology, etc.: **subfields of their disciplines**.

Definition (Theoretical digital humanities)

Research on the means and methods of constructing formal models in the humanities.

➜ Concerned with "construction materials" for formal models: **metascience**.

Can we formalize everything?

- ▶ **No!**
- ▶ But this does not exclude that formalization **can** yield new insights for a subset of phenomena.
- ➜ **Adequate modeling is decisive**

Adequate modeling is decisive

- ▶ Most phenomena of interest to the humanities are characterized by uncertainty and vagueness.
- ▶ The suitability of conventional computational approaches is limited.
- ▶ In order to gain new insights (≠ engineering) we need to **understand** our models—otherwise they do not fulfill their purpose.
- ▶ Modeling approaches (to address qualitative research questions) must primarily be looked for in the field of **knowledge representation**.
- ▶ Interdisciplinary challenge: finding formalizations that are adequate in two respects.
- ▶ **Interdisciplinarity manifests itself in the joint construction of models.**

Formal models in the humanities

Computers are valuable to the humanist exactly because what they produce compels him, or if you prefer, inspires him, to restudy and re-evaluate the artifact, the work of art, the human achievement which is the appropriate object of his dedication. (Kroeber 1967, p. 142)

In fact, the computer has even improved the quality of methods in philological analysis, because its brute physical rigidity demands **full accuracy, full completeness, full systematicity**. *[…] Using computers will therefore lead us to a more profound and systematic knowledge of human expression; in principle, it can help us to be more humanistic than before. (Busa 1980, p. 89)*

But only if we have good theoretical foundations!

Digitized, digitalized, or digitally transformed humanities?

► I believe "digital (or computational) humanities" have to be **digitally transformed humanities**.

► There's nothing wrong with any of them.

► But they're not one and the same!

► And it makes a difference beyond research in the stricter sense:
    ► Degree programs
    ► Funding
    ► Publications
    ► Careers
    ► Institutions
    ► …

References (1)

Borillo, Mario (1985). **Informatique pour les sciences de l'homme. Limites de la formalisation du raisonnement**. Bruxelles : Mardaga.

Bowles, Edmund A., ed. (1967). **Computers in Humanistic Research. Readings and Perspectives**. Englewood Cliffs, NJ, USA: Prentice-Hall.

Busa, Roberto (1980). "The Annals of Humanities Computing: The Index Thomisticus". In: **Computers and the Humanities** 14.2, pp. 83–90. URL: https://www.jstor.org/stable/30207304.

Gladkij, Aleksej Vsevolodovič and Igor Aleksandrovič Mel'čuk (1969). **Elementy matematičeskoj lingvistiki**. Moskva: Nauka.

Hackler, Ruben and Guido Kirsten (2016). "Distant Reading, Computational Criticism, and Social Critique: An Interview with Franco Moretti". In: **Le foucaldien** 2.1, pp. 7+. DOI: 10.16995/lefou.22.

Kenner, Hugh (Oct. 1989). "Curtains for Lady Buxley". In: **BYTE**, pp. 360–361.

References (2)

Kirsch, Adam (May 2014). "Technology Is Taking Over English Departments". In: **New Republic**. URL: https://newrepublic.com/article/117428/.

Kroeber, Karl (1967). "Computers and Research in Literary Analysis". In: **Computers in Humanistic Research. Readings and Perspectives**. Ed. by Edmund A. Bowles. Englewood Cliffs, NJ, USA: Prentice-Hall. Chap. 13, pp. 135–142.

Stachowiak, Herbert (1973). **Allgemeine Modelltheorie**. Wien, New York: Springer.

Weizenbaum, Joseph (1984). **Computer Power and Human Reason. From Judgement to Calculation**. Harmondsworth: Penguin.