Did they know what (on Earth) Marco Polo was talking about?

Exploring toponymic variance through software prototyping

Marcel Schaeben (CCeH, University of Cologne)
Jun.-Prof’in Dr. Elisa Cugliana (CCeH and IDH, University of Cologne)

Colloquium in Digital Cultural Heritage, Cologne, November 22nd 2023
Agenda

● **Subject matter: Marco Polo**
  ○ The guy
  ○ The text
  ○ Our focus

● **Methodology: Software Prototyping**
  ○ Prototyping in engineering and software engineering
  ○ Prototyping in DH
  ○ Epistemics of prototyping in DH

● **Case study: Marco Polo prototype**
  ○ Features and goals / ideas
  ○ Collaboration and development process
  ○ Preliminary findings, discussion
Subject matter
Marco Polo

- Venetian traveller (1254-1324)
Our focus

- A lot of exotic place names
- Scribes in the Western world had troubles deciphering the names while copying Marco and Rustichello’s text
- Copying implies reading implies predicting what is written (e.g. “Do yuo fnid tihs smilpe to raed?...”)
- This trick doesn’t work if you don’t know the language
- It’s even worse if you’re trying to read foreign proper names written in ambiguous medieval handwriting (*cf.* a whole lot of linguistic and philological research on proper names)
- → scribes made tons of mistakes (unless they knew the place; e.g. Rome)
- The more variants (∼ mistakes) of a toponym we find, the less geographical knowledge we can attribute to the scribe, and probably the readers (Cruse 2017)
What to do about this?

- Mapping the variance of the toponyms: what areas of the world retold by Marco were known to the audience of the time?
  - Red (more variance) → less known place
  - Blue (less variance) → better known place

- Using medieval maps: how did readers actually imagine Marco’s description of the world?

- (Include *stemmata*)
Methodology
Let’s build a prototype!

but why? 😐

and how? 🤔
Software Prototyping – a DH research method?

prōtotupos – first impression / figure / type / example

“first, typical or preliminary model of something, especially a machine, from which other forms are developed or copied”

Oxford English Dictionary, *prototype*, 2020

“preliminary example, usually full size, of a machine such as a motor vehicle or aircraft used to evaluate design and performance”

Some Prototypes


**Prototype / Erlkönig Mercedes E-Klasse W213**, photo: Robert Körner (CC BY-NC-SA 2.0), source: https://www.flickr.com/photos/67611651@N03/17932922119

**Apple prototype**, photo: Shannon Lucas (CC BY 2.0 DEED), source: https://commons.wikimedia.org/wiki/File:Apple_Prototype_%28530373102%29.jpg

**Design Thinking: Wallet prototype**, photo: John Nash (CC BY-NC-SA 2.0), source: https://www.flickr.com/photos/illiac1/517234427
Prototypes in Software Engineering

“preliminary version of a software system in order to allow [...] aspects of that system to be investigated … additionally (or alternatively) a prototype can be used to investigate particular problem areas or certain implications of alternative design or implementation decisions”


- often **reduced complexity**, focussing on **specific parts** of the system (Sommerville 2016; Butterfield, Ngondi and Kerr 2016)
- can be **incomplete**, ** buggy**, or “**half-cooked**” (Sommerville 2016)
- often developed in **iterations** towards a final product (Schneider 1996)

- used for **demonstrating ideas** and to **elicit feedback** (Devadiga 2017, Sommerville 2016)
- assessing **usefulness and usability** (Houde and Hill 1997)
# Prototypes in Software Engineering

<table>
<thead>
<tr>
<th>Fully functional (F)</th>
<th>Computer-based (C)</th>
<th>Paper-based (P)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Main screen for messaging menus</th>
</tr>
</thead>
</table>

Mobile phone UI prototype in different fidelities. Source: Lim et al. 2006.
Our prototype – a software prototype?

- “classic”, engineering-style software prototyping is found in DH!
  - when developing tools and software systems oriented towards specific requirements: databases, (geo-)information systems, user interfaces for digital editions etc.

  but… what about the “Marco Polo prototype”?

- preliminary version of a software system or an aspect thereof
  - which software system?
- often reduced complexity
  - necessarily so?
- can be incomplete, buggy, or “half-cooked”
  - yes, definitely!
- often developed in iterations towards a final product
  - iterations, yes… but which final product?
- used for demonstrating ideas and to elicit feedback
  - sure, but about what?
- assessing usefulness and usability
  - again… of what?
Software Prototyping in DH

“The digital humanists on the team called the first version of the […] system a prototype, whereas the software engineering approach characterized the system as a product”

Randa El Khatib (2019) about working on a DH/CS interdisciplinary team
Software Prototyping in DH

- “experimental, exploratory software
- ... developed iteratively
- ... as an intertwined part of a research process
- ... without (!) the specific goal of an end-product software system” (El Khatib 2019)

- built as a way to engage with a research question (Ruecker 2015)

- has become an inherent part of DH research method and outputs (Kleymann 2023)

→ Software prototyping as research? How?
Epistemics of Software Prototyping

How can it contribute to knowledge and understanding in the (Digital) Humanities?
Prototyping as research method in DH

- Prototypes can embody **arguments** and can be **interpreted** (Galey and Ruecker 2010)
  - for example: a new kind of **visualization** based on a text
    - presupposes an inherent structure in the text that can be brought out this way
    - presupposes that bringing it out this way contributes to knowledge
  - such arguments can be **identified, discussed, and contested**

- Prototypes can be built to **test hypotheses** or to **generate new hypotheses** (Kleymann 2019)
Prototyping as research method in DH

- Prototypes can show some characteristics of theories (Ramsay and Rockwell 2012)
  - as “framework[s] for the interpretation and understanding of a phenomenon”
  - “hermeneutical instruments”
  - “telescopes for the mind” that show a phenomenon in a new light, leading to new insights
  - “[A theory] orients the user toward certain features in the phenomenon and away from others” (Rockwell and Sinclair 2016)

- Prototypes can embody or contain theories (Galey and Ruecker 2010, Ramsay and Rockwell 2012, El Khatib 2019)
  - Materializations or reifications of abstract theories
  - Operationalization of theories, make them “runnable”
  - Demonstration devices, rhetorical function

Prototyping as research method in DH

- Prototypes can make **implicit ideas** or **hidden biases** explicit (Kleymann 2020, Ramsay and Rockwell 2012)
  - *explicitly*: through affordances in their UI
  - *implicitly*: by not working as expected (productive failure, cf. McCarty 2005)

- Prototypes can have a transparent participative nature
  - open to manipulation and exploration (Rockwell 2001)
  - by fiddling with the UI or reading or changing the source code (El Khatib 2019)
Modeling vs Prototyping

Modeling is a core research practice in the Digital Humanities (McCarty 2005, Rockwell and Sinclair 2016, Ciula et al 2018 and tons more…)

A model is (Jannidis 2018):

- a representation of something
- for some purpose
- which concentrates on some aspects (features and relations between them) and disregards others

In DH specifically: computers require formal models expressed in some form of formal language

“[Modeling] is a complex iterative process of integration and exploration with repeated loops of testing, feedback and adjustment” (Ciula et al 2018)

Modeling can serve to (ibid.):

- test an hypothesis
- generate new explanations
- mediate between a theory and the physical world (materializing / reifying)
- operationalize an argument

…and can productively fail (McCarty 2015)
Modeling vs Prototyping

Modeling is a core research practice in the Digital Humanities (McCarty 2005, Rockwell and Sinclair 2016, Ciula et al 2018 and tons more…)

A model is (Jannidis 2018):

- a representation of something
- for some purpose
- which concentrates on some aspects (features and relations between them) and disregards others

In DH specifically: computers require formal models expressed in some form of formal language

“Modeling is a complex iterative process of integration and exploration with repeated loops of testing, feedback and adjustment” (Ciula et al 2018)

Modeling can serve to (ibid.):

- test an hypothesis
- generate new explanations
- mediate between a theory and the physical world (materializing / reifying)
- operationalize an argument

…and can productively fail (McCarty 2015)
Modeling vs Prototyping

- “[Modeling] is independent of the technology used to manifest it” (Fishwick 2018)
  - prototyping as a technology-specific manifestation of the resulting models?

- Prototype as an interface to model(s)?
  - a vehicle by which models can (more effectively) do what they are said to do?
A prototypical circle
Hermeneutic circle(s)


Source: Simon Fraser University Media Lab. Cycles of iteration development. https://www.sfu.ca/media-lab/cycle/presentation/design.htmlmld

Source: ChatGPT, personal communication, prompt: “Can you create a mermaid diagram of the hermeneutical circle”, November 18, 2023.
Prototype vs text

Creating a prototype can be compared to **writing a critical or interpretative text / essay** about a phenomenon

(e.g. Rockwell 2001, Galey and Ruecker 2010, El Khatib 2019)

... so why bother?
Prototype vs. text (more circles!)
Prototype vs. text

Prototyping modulates theorizing and modeling by:

- collection, modeling and interpretation of data
- algorithm design and development
- interface design
- choice of technology
- choice of visualization
- rapid feedback loop: code → test → code → test… (Ruecker 2015)
- …
Prototyping as hermeneutical theory?

- Hermeneutical theories explain how interpretation works
  - i.e. how to “make sense” or “derive meaning” from a phenomenon
- Hermeneutical theories provide a framework for “doing” interpretation

➢ Prototyping could be regarded as a “hermeneutical theory”, a framework with a specific sets of hypotheses and methods for of interpreting, understanding, making sense of a phenomenon

➢ It is characteristically different from writing a text as a form of hermeneutics
  - It brings a different set of practices to the research process
  - Therefore it could (or should?) lead to different kinds of knowledge / understanding of a phenomenon

On the hermeneutics of building prototypes and other digital things in DH see:
Ruecker (2011), Ramsay (2012), Kleymann (2019, 2020, and 2023) as referenced at the end, also:
https://doi.org/10.1002/9781118680605.ch23.
**Summary: Prototypes and Prototyping in DH**

**DH Prototypes (the artifacts):**
- Are **materialized artifacts** of an *iterative process of reasoning* about a phenomenon
  - Phenomenon is external to the prototype, not a “final version” of itself or a greater software system
- **Embody** and / or **demonstrate hypotheses, arguments, theories, models** of / about sth. external
- Invite / provide affordances to **explore, manipulate, test, collaborate, discuss**
- Can be subject to **interpretation** and can serve as **interpretative tools**

**DH Prototyping (the process):**
- Is the creation of prototypes in an *iterative process of reasoning* about a phenomenon
- Has **no clearly-defined end-goal**
- Is an **intertwined process** with other interpretative processes / research activities
  - Hypothesizing, arguing, theorizing, (different forms of) **modeling**
  - Presupposes and influences these processes
- Is a distinctive **hermeneutical theory** about how to derive meaning / understanding from a phenomenon

*Disclaimer:* This is no textbook definition, these are just ideas. Let’s discuss them!
Prototype goals, features and preliminary findings
Different medieval maps

Walspberger's map (in contrast to Fra Mauro's) has "no room" for Marco's geography, hinting at a different level of geographical knowledge of the German audience, and thus a possibly different reception of the text compared to the Italian audience.
Textuality

● Different versions (and witnesses) of the work
● DI (3 witnesses): starting point
  ○ copied from a Tuscan version
  ■ copied from a Latin version
  ● copied from a Venetian version

● Normalised or not normalised texts (integrated pipelines!)
● Interactive narrowing down the visualizations to a chosen set of manuscripts
Stemmata: theoretical basis

- Choice of underlying hypothesis
Stemmata: single variants

- VA, Lagiazza, Llagiazza
  - P, Glaza, Glazam, Glaçam, Glaze
  - TB, Chiazza, Chvązza, laGlaza

  - cgm 252, allagiazza, giazza, giassa
  - cgm 696, allagiazza, ciaczyma, giazza, giaza, giazza, lagiagazaar
  - inc 77, allaggiassa, giassa, lagiassa
From the variants to the text
From the text to the variants
Expectations and surprises

- **Expectation**
  Considering the European origin of the texts considered, we expected “distance” to play a role in the distribution of the colours (the more distant, the redder).

- **Surprise!**
  The regions of Armenia are almost as red as more distant places in Asia.
Deductions

- Time spent in a place is relevant! The more Marco dwelled on specific territories (such as Anatolia), the more he described little cities and villages.

- Important centres, such as Constantinople, have less variants than smaller places, like Iskenderun: readers probably identified the former (if found in other contemporary sources), and mentally placed the latter somewhere in the vicinity.

  → Route perceived as a sequence of better known centres surrounded by blurred areas?
What we realised in the process

- Places that occur only once in a source do not display variants. If they are kept in the visualisations, they will appear as blue markers, but in the context of our study this is misleading, as in this case “blue” does not necessarily mean “familiar”.

- Need for strategies to address the heterogeneity of the sources (some versions describe some places, other versions omit some of them…)

- Problem of representing place uncertainty (cf. Walsperger’s map)

- Different categories of variants?
Collaboration

- “Semi-Formalized” through user stories and interviews
- Preparation of data
- Editor function

Feature 4: Visualise the place name spelling variance for each place.

📚 As a philologist, I want to graphically represent the level of pre-existing knowledge for the contemporary scribes and readers, based on the inversely proportional relation that exists between the degree of variance of a place name and the knowledge of it on the part of the scribe.

💻 As a digital humanist, I want to translate long lists of words into another medium, able to “speak for itself”. Through visualization we can evince more insights into the object of study.

🎓 As a PhD researcher, I want to provide evidence that the proposed inversely proportional relationship between variance of place names and knowledge of the scribe really exists. I encounter all these variants while transcribing and I basically do the same job of the scribe so I really understand the difficulty of deciphering unknown names so it is clear to me that visualising variance means visualising knowledge and this is extremely interesting.
Data

- TEI-XML files by Elisa
  - Witness transcriptions with encoded place names and unique IDs
    - Normalized and non-normalized
  - List of places with geo-coordinates
- TEI-XML encoding of stemmata

Manuscript encoding with place name annotations

```xml
<lb n="5"/>
<placeName ref="#catai">Cathay</placeName> wegen die lateinischen hęn
```

Encoding of stemma hypotheses

```xml
<eTree type="hypothetical" corresp="#x"> ... </eTree>
<eTree type="hypothetical" corresp="#y"> ... </eTree>
<eLeaf type="extant" corresp="#cgm252"> ... </eLeaf>
```

Encoding of places on historical maps

```xml
<facsimile>
  <surface>
    <graphic url="/assets/fra-mauro/{z}/{x}/{y}.png" width="5037px" height="5032px"/>
    <zone ulx="24.957908099801484" uly="70.5314146661368" lrx="24.957908099801484" lry="70.5314146661368" corresp="#catai"/>
  </surface>
  ...
</facsimile>
```

Encoding of places on contemporary map

```xml
<place xml:id="catai">
  <placeName xml:lang="fr" ref="#catai">Catai</placeName>
  <location>
    <geo>38.998667, 98.661112</geo>
  </location>
  <note>Northern China</note>
</place>
```
Technical details

- Browser-only *TypeScript* app, no server component
- UI framework *Angular* + *Angular Material*
- Maps: *leaflet.js*
  - Modern Map: *OpenStreetMap* accessed externally
  - Historical Maps: JPEG Tiles served as part of the app
- Stemma visualizations: *d3.js* + *dagre-d3* (DAG library for d3)
- Storage of XML data (as is): *IndexedDb Browser API* + *dexie.js*
- Rendering of TEI transcriptions
  - Transformation of TEI elements into *HTML Custom Elements*
  - Adding them to the DOM
  - Style with CSS, add interactivity using *Angular*
- Data queries: *FontoXPath* (XPath 3.0 and XQuery 3.0 impl. in browser)
Developments and conclusions
For the future

- Improve data quality
  - normalization
  - coordinates
  - errors (eg. missing IDs for places in data)
- Homogenize types of sources
  - Editions vs. transcriptions
- Add more sources
"Software prototypes not only contribute to the evidence and plausibility of empirical knowledge or hypothetical assumptions. Moreover, they create a framework in which something is talked about."

Kleymann 2020, S. 15

Ideas for discussion:

- **Our prototype**: What do you think? Did it make sense building it? Did it inspire some thoughts or ideas? What could we do with it in the future?
- **Prototyping as a research method**: Scholarly value? Relationship to other scholarly activities in DH such as writing or modeling? Can it contribute to knowledge and understanding?
Marco Polo prototype links

● App:
  ○ https://mrcl.uber.space/marco-polo/

● Source code, XML data and technical docs:
  ○ https://github.com/olvidalo/marco-polo
Works cited (1)


Works cited (2)


Secret bonus slide: prototypes vs. tools

- [Tools are] software to make work easier for other scholars (Ramsay and Rockwell 2012)
- [They] don’t explain or argue but simply facilitate. (Ramsay and Rockwell 2012)
- Prototypes are "sandcastles": transient, unstable, interactive. Tools are pragmatic, functional and transferable (Hinrichs and Forlini 2017)
- prototypes can be disseminated early to others to engage in scholarly dialogue about them (Brown et al 2009)
- The success of [prototyping] projects is not pegged on completion, but measured in other ways (Brown et al 2009)
