Representing change: User interaction and data modelling of an identity paradox

Remo Grillo

Introduction

When does something become something else? Although this question is literally thousands of years old, it still requires new and adequate solutions in unexpected areas today. This riddle has been puzzling many cultures quite distant from each other, and in the history of western philosophy, it is sometimes known as the *Ship of Theseus* paradox. It can be formulated like so: "If every part of a specific ship gets replaced entirely, piece by piece, is it still the same ship?". More in general, this belongs to the family of problems about *identity over time*.[1] Surprisingly enough, after more than two millennia from its original formulation, in the Digital Humanities (DH) the same puzzle is still faced, but this time for very practical and technical reasons, rather than to satisfy the ambition of build a philosophical understanding of the world.

Digital Humanists, in fact, are sometimes called upon to formalise knowledge through the use of ontologies, but some historical objects and facts are so convoluted and faceted that it is hard to express them fully in natural sentences: much less formalise them with a rigorous ontology. As we know, in the world of semantic web, knowledge can be represented through structures called knowledge graphs, where every entity is connected to other entities through relations that describe characteristics, involvement into events, connections of any kind among themselves: everything is defined clearly. It is worth remembering that we do this, instead of using our ambiguous and blurry natural language to describe historical events, because in doing so computers can elaborate knowledge, thus allowing us to discover new information through automatic inference, in ways and means that are exponentially faster, universally accessible and more serendipitous than with traditional tools.[2]

But here is the challenge: ontologies – these fixed grids we use to grasp and pigeonhole knowledge – are discrete, while reality is a continuum. Ontologies work like Legos, but reality is like play-dough.

[1] According to early Buddhists, coeval to the ancient Greek philosophers, a brave "no" is the answer to the riddle and one of the pillars of buddhist philosophy: the concept of *impermanence*.

[2] Eero Hyvönen, 'Using the Semantic Web in digital humanities: Shift from data publishing to data-analysis and serendipitous knowledge discovery'. *Semantic Web*, by Pascal Hitzler and Krzysztof Janowicz, vol. 11, n. 1, January 2020, 187–93. *DOI.org (Crossref)*, https://doi.org/10.3233/SW-190386.

A gap between model and reality

Sometimes we do not notice the representational gap between model and reality, because the problematic nature of reality hides behind the apparent simplicity of the facts we want to describe: for instance, while it is easy to formalise in a graph the claim "Napoleon was born in Ajaccio in 1769, and he participated in the battle of Waterloo" (fig. 1), it is not always so easy to represent other kinds of facts which are less linear, and for which modelling solutions are not so obvious. During the

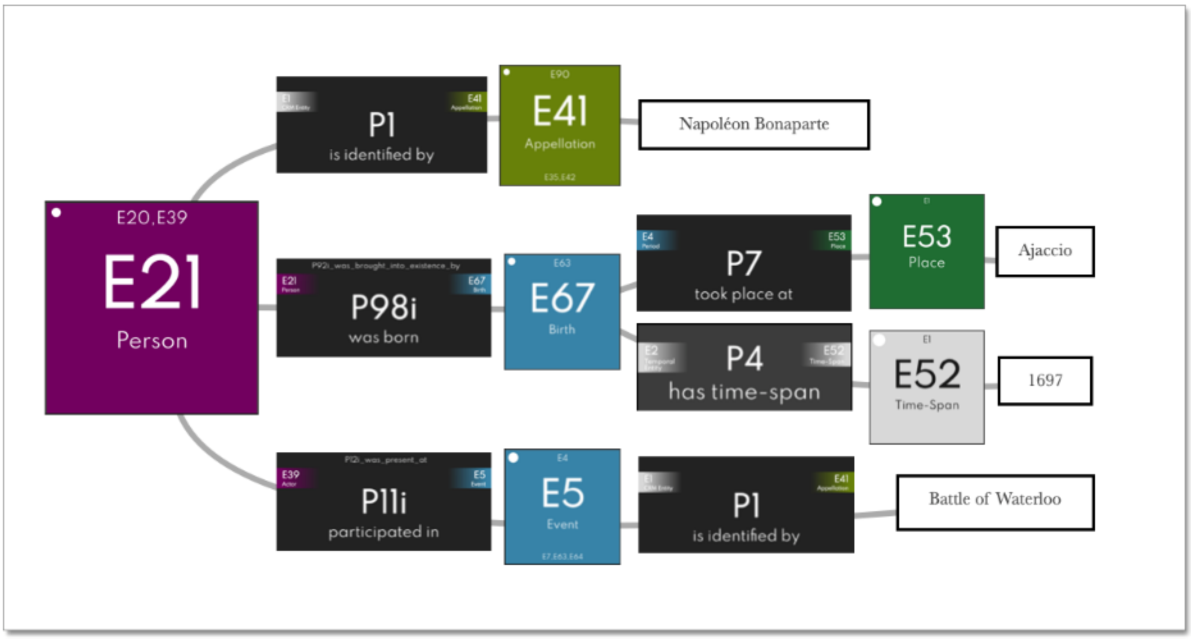

Fig. 1: A sample semantic formalisation of "Napoleon was born in Ajaccio in 1769 and participated in the battle of Waterloo" in Cidoc-CRM (simplified).

development of *Venice's Nissology (VeNiss)*,[3] we came across a surprising complexity trying to design a data model devoted to buildings in particular. In fact, VeNiss is a project with a peculiar focus on the history of architecture among other things and, by its very nature, it requires a good level of detail regarding the modification of buildings and their roles in the network of islands that constitute Venice's archipelago. Over the decades and the centuries buildings underwent all sorts of changes: they were extended, partially destroyed, restructured, their materials were modified or reused for other buildings, decorative elements or entire environments can be added, etc. In addition, built works that for centuries had important religious functions, such as churches or ecclesiastical complexes, in many cases were partially demolished during French and Austrian rule or incorporated in other buildings and then converted into strategic military zones. A continuous flux of change like this

[3] The project *Venice's Nissology. Reframing the Lagoon City as an Archipelago: A Model for Spatial and Temporal Urban Analysis (16th-21st centuries)* has been funded with a five-year grant by the European Research Council (ERC-2021-StG, n. 101040474).

can be surprisingly common for buildings, especially in a timespan of five centuries. Can we claim that the small church is *ontologically the same* building as the desecrated military facility rebuilt after an explosion centuries later? Perhaps not. Then it must have changed its identity at some point. When? Was it when the building changed its shape or when it changed its name and use? Was it when it was enlarged or only after it was bombed? If so: which characteristics of an entity are necessary to it so that, if removed, that thing is no longer itself but, rather, something else[4]? This is a necessary technical question required to explicitly model identities in a knowledge graph. If the small church *becomes* a military facility, this means that these two buildings are different entities in the graph, but they somehow have a connection in a chain of events.

Facing the paradox with a data model

We resolved to adopt a fairly strict methodology that makes a distinction about the typology of change based on whether this concerned the physical configuration or the function of a building. Physical changes involve its construction, modifications (additions, removals, substitutions, etc.), destruction. Functional changes are those involving what people make of it: a building can be named in one way or another, it can have a private or public *function*, it can have a specific *use* (religious, military, administrative, etc.) or a building of a certain *typology*: a convent, a shelter for boats, a wall, a house, a bell-tower, etc. This physical/functional dichotomy lead us to consider all possible combinations of types of changes of a building, and then find ad-hoc solutions for all combinations:

1) no changes
2) just functional changes
3) just physical changes
4) both physical and functional changes.

For each, Cidoc-CRM provides a good level of expressiveness that allowed us to conceptualise and formalise all these nuanced aspects (Fig. 2).

1. First of all, it is tautological that if a building does not change physically nor functionally then the entity corresponding to that built work inside our graph must remain unique, identical to itself in time.

2. In case a building remains substantially identical in physical terms, but its *use, name, typology or function* are changed, we materialise different usages and appellations for the same identical building, but referred to different time-spans.

[4] On necessity of identity see: Saul A. Kripke, *Naming and Necessity: Lectures Given to the Princeton University Philosophy Colloquium*. Cambridge, MA: Harvard University Press,1980.

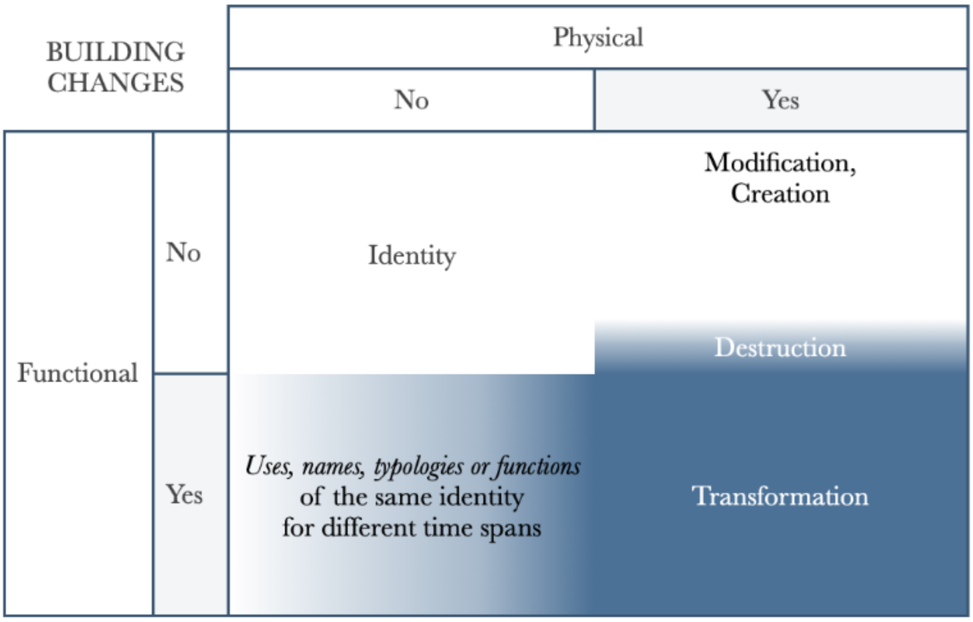

Fig. 2: Combination of possible varieties of changes for buildings with relative modelling solutions.

3. Solely physical changes give rise to *modification* events of the same building.[5] A church can be enlarged with a new bell-tower, restored, or partly destroyed and rebuilt, still keeping its identity over time. Although, complete material destruction makes that building cease to exist.

4. Changes of both function and shape happening simultaneously give rise to a new identity through an event of *transformation*,[6] which constitutes a causal/historical connection between two buildings:

- the one that ceases to exist;
- the new one, resulting from the transformation of the previous

Nonetheless, some extreme cases may require exceptions to this model, in accordance with the historian's understanding and the context-sensitive aspects of particular historical claims. This is particularly true for cases involving gradual changes of function and matter that do not happen simultaneously. For these situations it is important to remember that, when possible, the model should adapt itself to the scholar's request and vision, rather than the opposite.

[5] https://cidoc-crm.org/Entity/e11-modification/version-7.1.1

[6] Note that, in the Cidoc-CRM, *Transformation* is a subclass of both *End of existence* and *Begin of existence.*
E81(x) ⇒ E63(x), E81(x) ⇒ E64(x). Cfr. https://cidoc-crm.org/Entity/e81-transformation/version-7.1.1

Implying existence

Talking about the identity of something is not the same as talking about its existence. In a sense, we can say identity is a concept that relates an entity to other entities, as in questions about sameness or otherness: "can we claim that building A and B are the same?". Existence, on the other hand, relates an entity to time, as in questions of presence, such as: "did building A already exist in 1500?".

The existence of a building in a specified year is testified by historical documents, such as architectural plans, maps, paintings, land registries, etc. and is recognised by the scholar based on these proofs and related deductions. For instance, if the same building is both visible in a map from 1500 and in a document from 1700, we can assume that it existed in 1600 also. The opposite is also true: if a building is visible in a document from 1500 but not in one from 1600, we can assume it didn't exist anymore in 1700. Using this methodology, for a given space, it is possible to define a timeline (fig. 3) about the existence and co-existence of buildings, given that there can be also gaps and uncertainties – and therefore possible contradictions and overlaps in time and space – that need to be represented nonetheless.

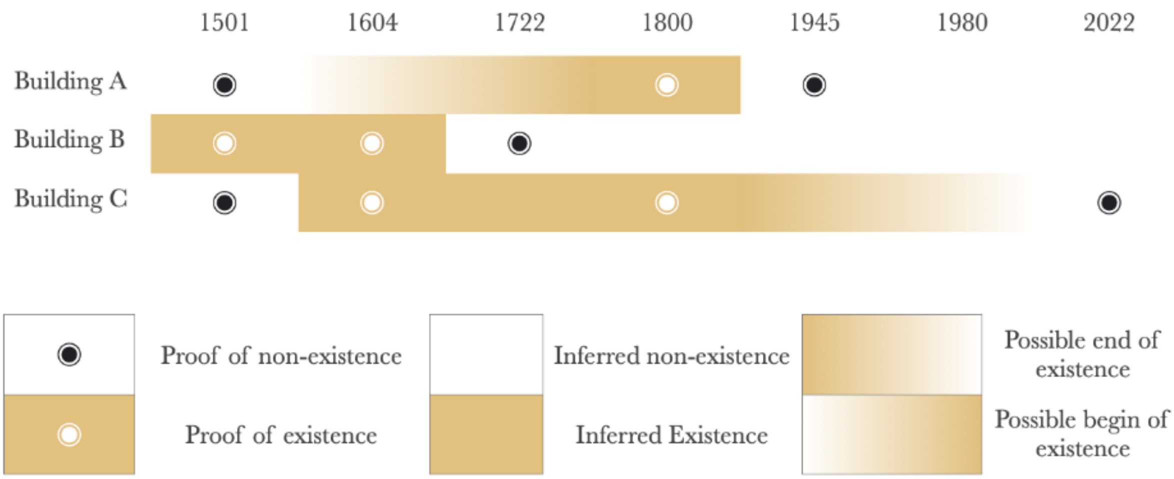

Fig.3 Three sample timelines, each rendering the possibilities of documented existence of buildings.

The interface

Providing the user with an intuitive result of this formalised process of knowledge is a specific task. The interface, in fact, should convey a temporal dimension, spatial and geometric data, qualitative information (names, functions, types, etc.): all at once and interactively.

As shown (fig. 4), georeferenced buildings are represented superimposed on different basemaps, whose opacity and order can be managed through stackable layers.

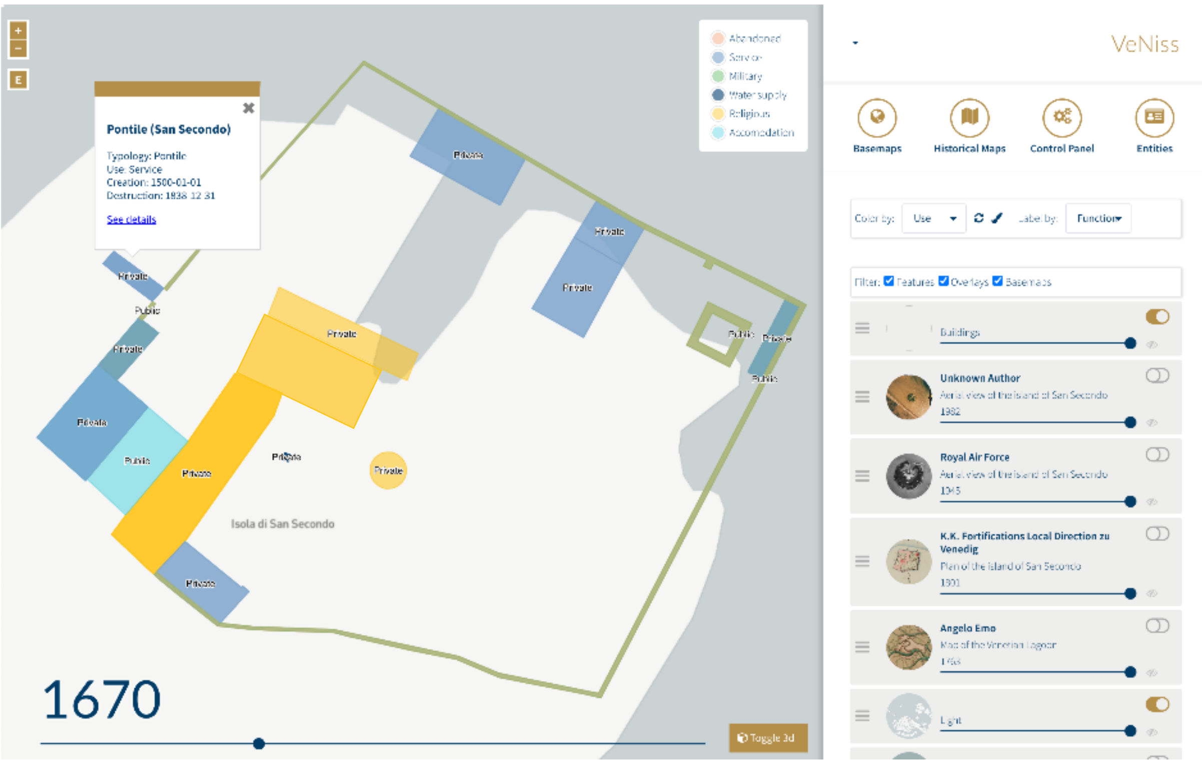

Fig. 4: The interface render as implemented in VeNiss at its current stage of development.

Qualitative information, mostly related to nominal and functional aspects of the buildings, are rendered using labels and colours. Colours can be managed through a map symbol component, of which it is possible to decide the taxonomy criteria used to discern the classes of buildings and manually modify the palettes.

The spacetime-volume of changes is filtered temporally with a year-by-year granularity, and a slider component can be dragged to move in time and appreciate changes on the map diachronically.

Conclusions

The Cidoc-CRM provides a valid expressiveness to represent change of buildings over time, and allows historian to produce formalised knowledge even about particularly extreme cases where the identity of a building is critically questioned. These changes can be rendered through a 4-dimensional GUI, using diachronical interaction where the time dimension can be manipulated, as can the spatial one. Causal chains tied to a specific building are not yet directly rendered in the interface, although that might be of a very specific interest for the project, and as such deserves future development.

Remo Grillo holds a Master's degree in Digital Humanities and Digital Knowledge with honours from the Alma Mater Studiorum - University of Bologna. His research focuses on knowledge representation and extraction through ontologies, as well as on natural language processing and algorithms applied to music, museum artifacts, and literature. He worked as a Full-stack Developer while attending the University of Naples Federico II and Ludwig Maximillian University of Munich, where he obtained a Bachelor's degree with a thesis on Logic, applied to a digital critical edition of the *Tractatus Logico-Philosophicus* by Ludwig Wittgenstein.

rgrillo@itatti.harvard.edu