Knowledge Representation: From Observation to Data

Emilio M. Sanfilippo, PhD
ISTC-CNR Laboratory for Applied Ontology, Trento, Italy
ISTC-CNR Catania, Italy
emilio.sanfilippo@cnr.it

Statements*

In different application contexts, need for explicit representation of both:

- **Domain entities**
- ***Statements** (observations, claims, perspectives, etc) on domain entities

For instance:

- Domain entity: Guernica (Paris, 1937), a painting by Picasso, technique Oil on canvas
- Various statements about its aesthetic features, symbolic meaning, political meaning, etc.

Picasso's Guernica

An accurate depiction of a cruel, dramatic situation, *Guernica* was created to be part of the Spanish Pavilion at the International Exposition in Paris in 1937. Pablo Picasso's motivation for painting the scene in this great work was the news of the German aerial bombing of the Basque town whose name the piece bears, which the artist had seen in the dramatic photographs published in various periodicals, including the French newspaper *L'Humanité*. Despite that, neither the studies nor the finished picture contain a single allusion to a specific event, constituting instead a generic plea against the barbarity and terror of war. The huge picture is conceived as a giant poster, testimony to the horror that the Spanish Civil War was causing and a forewarning of what was to come in the Second World War. The muted colours, the intensity of each and every one of the motifs and the way they are articulated are all essential to the extreme tragedy of the scene, which would become the emblem for all the devastating tragedies of modern society.

Guernica has attracted a number of controversial interpretations, doubtless due in part to the deliberate use in the painting of only greyish tones. Analysing the iconography in the painting, one *Guernica* scholar, Anthony Blunt, divides the protagonists of the pyramidal composition into two groups, the first of which is made up of three animals; the bull, the wounded horse and the winged bird that can just be made out in the background on the left. The second group is made up of the human beings, consisting of a dead soldier and a number of women: the one on the upper right, holding a lamp and leaning through a window, the mother on the left, wailing as she holds her dead child, the one rushing in from the right and finally the one who is crying out to the heavens, her arms raised as a house burns down behind her.

At this point it should be remembered that two years earlier, in 1935, Picasso had done the etching *Minotauromaquia*, a synthetic work condensing into a single image all the symbols of his cycle dedicated to the mythological creature, which stands as *Guernica*'s most direct relative.

Incidents in Picasso's private life and the political events afflicting Europe between the wars fused together in the motifs the painter was using at the time, resulting both in *Guernica* itself and all the studies and 'postscripts', regarded as among the most representative works of art of the 20th century.

Paloma Esteban Leal

Source: Museo Reina Sofia, Madrid, Spain

Statements (not exhaustive)

Different kinds of statements based on:

- Speculative theories, e.g., Hermeneutics, Marxisms, Psychoanalysis;
- Mathematical methods, e.g., statistics;
- Formal, computational methods, e.g., formal reasoning, pattern recognition, machine learning;
- Empirical research, e.g., historical research, physical measurements, carbon-14 analysis, etc
- ….

All possible combinations

In a nutshell

By expressing statements:

- We express a **point of view**, a **perspective** on a certain phenomenon
- That is not necessarily **true** with respect to reality
- That is not necessarily **compatible** with other perspectives
- That is **based on** multiple factors (speculative theories, empirical research, formal methods, etc.)

From the expression and representation of Facts to

Models for Observational knowledge and Data

Statements in the Humanities

In many disciplines of the humanities, truths are hard to come by and facts are rare. In most cases, we use words such as facts and truths just to mean "*statements for which there is an acceptable trail of supporting sources*", or "*statements that are more or less accepted by the majority of the relevant scholars*" or "*statements that so far have not been disproven*".

We [...] believe that we should explicitly aim at representing competing points of view and opinions, and make sure that we fully document their existence, their strengths and the ideas behind them so that our audiences can finally perceive representations that are truer and more interesting than the sterilized and boring renditions forced by so-called objectivity.

Barabucci, G., Tomasi, F., & Vitali, F. (2021). *Supporting complexity and conjectures in cultural heritage descriptions*, COLCO 2020.

Statements in Natural Sciences and Engineering

Very similar scenarios

Example:

- The temperature of this room is 16 degrees C

  → the temperature of this room *has been measured* with a sensor

  → the sensor can measure temperatures within a range with a certain level of precision, etc.

Some (recent) literature on the topic

Models, theories, data, observations … classic topic in philosophy (of science), e.g.,

- Leonelli, S. (2019). What distinguishes data from models?. European journal for philosophy of science, 9(2), 22.
- Morgan, M. S., Morrison, M., & Skinner, Q. (Eds.). (1999). Models as mediators: Perspectives on natural and social science (No. 52). Cambridge University Press.
- Suppes, P., & Krantz, D. H. (2007). Foundations of measurement: Geometrical, threshold, and probabilistic representations (Vol. 2). Courier Corporation.

Some (recent) literature on the topic

- Barabucci, G., Tomasi, F., & Vitali, F. (2021). Supporting complexity and conjectures in cultural heritage descriptions.
- Doerr, M., Kritsotaki, A., & Boutsika, K. (2011). Factual argumentation - A core model for assertions making. Journal on Computing and Cultural Heritage (JOCCH), 3(3), 1-34
- Freedman, R., et al. https://crimproject.org/about/home/
- Gangemi, A., & Presutti, V. (2022). Formal Representation and Extraction of Perspectives. Creating a More Transparent Internet: The Perspective Web, 208.
- Gardin, J. C. (1997). Le questionnement logiciste et les conflits d'interprétation. Enquête. Archives de la revue Enquête, (5), 35-54.
- Janowicz, K., Haller, A., Cox, S. J., Le Phuoc, D., & Lefrançois, M. (2019). SOSA: A lightweight ontology for sensors, observations, samples, and actuators. Journal of Web Semantics, 56, 1-10
- Masolo, C., Botti Benevides, A., & Porello, D. (2018). The interplay between models and observations. Applied Ontology, 13(1), 41-71.
- Niccolucci, F., & Hermon, S. (2017). Expressing reliability with CIDOC CRM. International Journal on Digital Libraries, 18, 281-287.

Our purpose today: WHAT and HOW

To brainstorm on observations with an **interdisciplinary methodology** at the intersection between:

- **Specific disciplines** (e.g., archeology, literary studies, musicology): domain-specific knowledge, data, desiderata
- **Philosophy**: centuries of critical thinking, theories, conceptual tools
- **Knowledge representation methods**, logics included: rigorous formal means to formally represent (portions of) observations

Our purpose today: WHY

Among others:

- To escape from narrow, scholarly caves, hence
- To look at observations from a wider perspective
- To come up with a **conceptual framework** that can support scholars in documenting (perhaps only partially) their claims
- To support the definition of **formal models** and **digital applications** to (partial) document, compare, analyze, share etc. observations
  (For those interested in research at the intersection with computer science)

Some open questions

To brainstorm about observations:

- What is a (scholarly) observation? What are its features?
- What kinds of observations are expressed in a certain community?
- By means of what methods are observations expressed?

From a knowledge representation perspective:

- How can knowledge representation approaches deal with the modeling of observations?
- Which formal languages shall we rely on?
- What strategy shall we adopt to deal with incompatible observations?
- Since scholars may adopt multiple and not necessarily compatible approaches, is it possible the creation of a general model for the expression of observations?