Digital Humanities, Latour and Networks

Moses A. Boudourides

University of Patras, Greece

Moses.Boudourides@gmail.com

Digital Humanities in Estonia Conference

Tartu, Estonian Literary Museum, October 12-14, 2016

Foundational Standpoints on Digitality and Relationality

- Two fundamental ontological (set–theoretic) notions:
  - ***One*** or **Count–as–One**: ***Ex*clusive disjunction** (*distinction*)
  - ***Two*** or **Count–as–Pair**: **Cartesian product** (*association*)
- Two types of observational units:
  - ***Monads***: **Individual** entities (*Being–in–Itself*)
  - ***Dyads***: **Transindividual relations** (*Being–With–Others*)
- Two representations of entities:
  - ***Analogy***: *Transcending* the **as–structure** (*Philosophy*)
  - ***Univocity***: *Immanentizing* the **in–structure** (*Science*)
- Two ways to transform entities:
  - ***Analog***: Mapping **as–structures** but **not in–structures**
  - ***"Digital"***: Mapping **in–structures** but **not as–structures**
- Two foci of vision:
  - ***Pattern Recognition***: (emergence of) **Regularities**
  - ***Regularity*** or ***Anomaly Detection***: **Singularities**
- Two pervasive risks:
  - ***Apophenia***: Seeing the non–existent (William Gibson)
  - ***Control Society***: Algorithmic politics (Foucault, Deleuze & Guattari)

The Digital "Folding" of a Literary Text (Berry 2011)

- **Translating a literary text into digital code:**

  From second order *univocity* to a second order *in–structure*.

  - Inscription (and storing) into a socio–technical device through the process of a *discrete* (binary) encoding.
  - Stabilization of the digital representation through "cleaning" (mangling), "datafication" and division into discrete subprocesses for particular user tasks.
  - Internal event–driven computational processing and algorithmic manipulations.
  - Description (according to technology prescriptions), displaying, interpreting, negotiating findings of analysis.

Periods ("Waves") of Development of Digital Humanities (Schnapp & Pressner 2009)

First wave: Database search and retrieval, text analysis and automating corpus linguistics.

Second wave: Hybrid quantitative–qualitative methodologies of editing and database processing, hypertext and hypermedia production and sharing and new publication models (digital libraries, open access peer review journals).

Third wave: From search and hypertextualize to detect and recognize.

- Media changes effectuating literary experimentation, innovation and aesthetic regeneration ("*change tools to change epistemologies*").
- How code (and big data) is inflitrating the academia and society? (Berry 2011)
- What is culture after it has been "softwarized"? (Manovich 2008)

The Computational Turn in Digital Humanities

- **The algorithmic eye (Pasquinelli 2015) in literal data analysis aims at**:
  - Extracting the digital surplus–value (the *network effect*) of literary production (similarly to what Google, Facebook etc. do to the Internet).
  - Forming a repository of accummulated *big data* as a huge historical–archival lab that can be accessed easily practically by eveybody.
  - Using techniques of *Machine Learning* in order to test, validate, monitor and predict social tendencies based on cultural, artistic and literary evidences.
  - Corraborating and improving the design and efficiency of algorithms themselves (by using case studies from the humanities).
  - Analyzing complex patterns of high–level abstraction (not necessariry associated with experience or perception), it becomes a vehicle for exploration beyond the limit of common experience (Terzidis 2003).

The Three Formal Abstractions of Algorithmic Digital Humanities

1. **Quantitative Analytics**
2. **Network Analytics**
3. **Data Visualizations**

From Gabriel Tarde...

"The public journals will become socially what our sense organs are vitally. Every printing office will become a mere central station for different bureaus of statistics just as the ear-drum is a bundle of acoustic nerves, or as the retina is a bundle of special nerves each of which registers its characteristic impression on the brain. At present Statistics is a kind of embryonic eye, like that of the lower animals which see just enough to recognise the approach of foe or prey."

— Gabriel Tarde, *The Laws of Imitation* (1890, p. 136)

... To Franco Moretti

"*A canon of two hundred novels, for instance, sounds very large for nineteenth–century Britain (and is much larger than the current one), but is still less than one per cent of the novels that were actually published: twenty thousand, thirty, more, no one really knows – and close reading won't help here, a novel a day every day of the year would take a century or so... And it's not even a matter of time, but of method: a field this large cannot be understood by stitching together separate bits of knowledge about individual cases, because it isn't a sum of individual cases: it's a collective system, that should be grasped as such, as a whole.*"

— Franco Moretti, *Graphs, Maps, Trees: Abstract Models for a Literary History* (2007, pp. 3–4)

Bruno Latour, the "Prince of Networks" (Harman, 2009)

- The Three Stages of Latour's Work on Networks:

  1. **Actor-Network Theory** (**ANT**) and the sociotechnical premise.
  2. **Reassembling the Social** through heterogeneous networks.
  3. Testing **Gabriel Tarde**'s theories: *The Whole is Always Smaller than Its Parts* (TWIASTIP).

- Here we are going to restrict our attention on Latour's third stage.

Gabriel Tarde and the Quantifiability of the Social: A Brief Discussion

- **Gabriel Tarde** (1843-1904), until the 1890's was the major French sociologist.
- After he was succeeded by **Émile Durkheim**, Tarde was forgotten almost completely.
- The interest to Gabriel Tarde was revitalized after the 1980's by **Gilles Deleuze** and **Bruno Latour**.
- Tarde's main sociological contribution refers to a number of strong positions he was raising on the dualisms of:
  - **micro–macro**,
  - **atoms–structure**,
  - **parts–whole**,
  - **elements–aggregates** and
  - **individuals–society**.

- Tarde was resolutely rejecting any form of atomism. For him, "**everything is a society**," in nature, history, culture, economy, technology etc., absolutely everywhere.
- For a Tarde, such a *panspermic society* was not, either an emergent, or supervenient holistic form of existence of beings, which was organized at a higher lever than its parts and was constituting something ontologically and epistemologically different but equivalent to its parts. On the contrary, for Tarde, parts of a totality were striving to organize both the substantatiation of their own identities and the terms of their collective existence.
- For Tarde, the elementary ingredients and constituents of any social form were called **monads**. Apparently, this notion was borrowed by **Leibnitz**, although instead of God, who according to Leibnitz was guaranteeing the structural cohesiveness of monads, Tarde was attributing their coherence in their interactional faculties of "**imitation**." Thus, the "imitative rays" that monads emit were the realization that monads share attributes modified by each sharing, the result of which is a list made up of the "same" item repeated with difference (Deleuze).

- For Tarde, the basis of any motives for action and interaction that monads possess results as a combination of three psychological stances:
  1. **senses**,
  2. **beliefs** and
  3. **desires**,
- Among the three:
  - Senses are always **qualitative**, because they lack variation and nuances.
  - The other two, beliefs and desires, are always **quantitative**, because they are oppositional, contested, polarized along certain axes over which differences can be measured, compared and assessed.
- That was why Tarde believed in the ultimate quantifiability of sociology through statistics, though which, a time will come that "***public broadsheets will be to the social world what the sensory organs are to the organic world***."

The Three Ages of Quantitative Sociology (Boullier 2014)

- The following periodization (Boullier 2014) captures the whole range of the development of the *quantitative social* from the time of Tarde till the present:

  Age 1: **Statistics and the idea of society**

  Age 2: **Polls and the idea of opinion**

  Age 3: **Digital traces and the idea of vibrations**

Gabriel Tarde's idea of **vibration**

"*If Statistics continues to progress as it has done for several years, if the information which it gives us continues to gain in accuracy, in dispatch, in bulk, and in regularity, a time may come when upon the accomplishment of every social event a figure will at once issue forth automatically, so to speak, to take its place on the statistical registers that will be continuously communicatedto the public and spread abroad pictorially by the daily press. Then, at every step, at every glance cast upon poster or newspaper, we shall be assailed, asit were, with statistical facts, with precise and condensed knowledge of all the peculiarities of actual social conditions, of commercial gains or losses, of the rise or falling off of certain political parties, of the progress or decay of a certain doctrine, etc., in exactly the same way as we are assailed when we open our eyes by the vibrations of the ether which tell us of the approach or withdrawal of such and such a so–called body and of many other things of a similar nature.*"

— Gabriel Tarde, *The Laws of Imitation* (1890)

'Latour's Deconstruction of

'Network is a Structure Composed of Actors and Their Interactions"

- Latour's working hypotheses in TWIASTIP:
  - "**There is no individual actor. Actors don't interact; there is no whole superior to the parts**."
    - Bruno Latour: *Network actor's are Tarde's monads*.
  - **There are no interactions among actors**.
    - Gabriel Tarde: *Agents cannot be said, strictly speaking, to 'interact' with one another: they are one another, or, better, they own one another to begin with, since every item listed to define one entity might also be an item in the list defining another agent*.
    - Bruno Latour: *A monad is not a part of a whole, but a point of view on all the other entities taken severally and not as a totality*.
  - **There is no society** (as a macro–entity) but ''**everything is a society**."
    - Gabriel Tarde: *Monads conspire with one another without ever producing a structure*.

- **Bruno Latour: ''Actors' profiles modify them and aggregate relations."**
  - John Dewey: ***Association*** *is not what happens after individuals have been defined with few properties, but what characterize entities in the first place* .
  - Bruno Latour: "*The '**wholes**' are nothing more than several other ways of handling the **interlocking of profiles***."
- **Methodological consequences of Latour's monadological network approach**:
  - "*Actors' **profiles** are actors' **attributes***."
  - "*Actors organize their **interactions' profiles***."

What About Time?

- In what concerns **temporal graphs** or **time–dependent networks** (or **longitudinal networks**), in principle, there are *two ways to follow the **flow** of the relational variability in time*:
  - A **Eulerian perspective**: Follow the network as a **time–varying relational totality** (e.g., a flock of flying birds) and, from such an *absolute perspective*, observe the patterns of how actors' transient profiles keep on modifying, rearranging and reorganizing the topology of their network.
  - A **Lagrangian perspective**: Follow the **trajectories** of particular actors (e.g., a singular **bird's–eye view**) and, from such a *relative perspective*, observe how the rest of the actors are positioned relationally (pair–wise) with regards to the focus of observation as the time flow of the network evolves.
- **Two different visualizations!**

The Complexities of Networks of Literary Text

- Given a corpora of literary texts, for instance, literature novels or any other books (including the case of a single book), how can a digital humanist study it as a network of some sort?
- One has to make a number of decisions:
  - If the aim is to understand the context of the studied corpora inside the **ecology** of a more inclusive stream of literal or socio–cultural phenomena, one has to expand the coropora with other pertinent bibliographic sources (for instance, a **bibliometric analysis** might indicate a wider range of relevant textual material) and also with any available **social or historical (archival) data** that frame more general issues around the examined case.
  - In any case (independently of scale):
    - What entities may considered as **actors/monads/units** of a single overarching network?
    - What can be the **interactions/relations** among these actors/monads/units?
    - What would be a functional notion of **time** inside these books, which would produce appropriate time–dependent networks?

A Few (Personal) Answers

- Actors/monads/units in text corpora may be:
  - The **characters/heroes/branded–names** of the text (*difficult to be extracted automatically without being inscripted by "human" annotattors*).
  - Particular "**words**" appearing in the text, like **noun phrases**, **toponyms** etc. (*rather easily to be extracted automatically*).
  - General **n–grams** from the text . In computational linguistics, n–grams are contiguous sequences of n items, like phonemes, syllables, letters, words or base pairs according to the application (*easy to be extracted automatically*).

- Interactions/relations among actors/monads/units in the texts may be:
  - The **real interactions** (including their conversations) of the characters/heroes/branded–names in the text corpora (*extremely difficult to be extracted automatically without being inscripted by "human" annotattors*).
  - n–gram **similarity** and **distance** in the text (*easy to be extracted automatically but substantively not useful*).
  - **Co–occurrences** of characters/heroes/branded–names/"words"/n–grams inside well–defined parts of the text, like **sentences**, **paragraphs**, **sections**, **chapters**, **acts** etc. (*rather easily to be extracted automatically*).
  - **Correlations** among the previous co–occurrences (*easy to implement computationally*).