Humanist Studies & the Digital Age, 5.1 (2017)
ISSN: 2158-3846 (online)
http://journals.uoregondigital.org/hsda/
DOI: 10.7264/N369720D

The Pico Project: Looking Ahead

Dino Buzzetti and Ernesto Priani

Abstract

In this essay we examine and assess text mining methods in order to find exact sources of Pico's theses ascribed to Medieval authors in his *Conclusiones nongentae.*

1. The digitisation of the *Conclusiones Nongentae Disputandae* at the Virtual Humanities Lab of Brown University opens up new research prospects for the Pico Project. The edited text is an accurate transcription of the *editio princeps* printed in Rome by Eucharius Silber on 7 December 1486. Pico's *Conclusiones*, or *Theses*, is a collection of short statements composed on the authority of former thinkers and schools, as well as according to his own opinion, on a wide range of philosophical and theological topics. Right at the outset of his "Preface," Pico describes his own work with these words:

> The following Nine Hundred Dialectical, Moral, Physical, Mathematical, Metaphysical, Theological, Magical, and Cabalistic opinions, including his own and those of the wise Chaldeans, Arabs, Hebrews, Greeks, Egyptians, and Latins, will be disputed publicly by Giovanni Pico of Mirandola, the Count of Concord. In reciting these opinions, he has not imitated the splendor of the Roman language, but the style of speaking of the most celebrated Parisian disputers, since this is used by almost all philosophers of our time. The doctrines to be debated are proposed separately by nations and their sect leaders, but in common in respect to the parts of philosophy—as though in a medley, everything mixed together. (Farmer and Pico 1998, 211)

Pico's theses deal with an astonishing amount of different subjects and refer to a large number of authors belonging to several philosophical and religious traditions. This means that their invention rested upon a prior knowledge of a vast and eclectic variety of doctrines and works. As Stephen Farmer observes, "no one can claim mastery over more than a small part of the traditions covered in Pico's text" (xiv), and so the problems of tracing back all of his sources still remain an open question.

Pico's *Theses* are organised in two main groups: the first 400 "historical theses" (32) according to the opinion of past philosophers, and the remaining 500 theses according to his own opinion. In turn, the philosophers' opinions are grouped by nations or sects—the Latins, the Arabs, the Greek Peripatetics and Platonists, the Pythagorean, Chaldean, Hermetic, and Cabalist wisemen—and for each nation by their sect leaders, such as Albert the Great, Thomas Aquinas, Averroes and so on. So apparently, we should know to whom the ideas conveyed by a given thesis is to be ascribed. However, Farmer warns us that "it is important to recognize that not all or even most of the nine hundred theses can be traced unambiguously to single sources", since many of them "express opinions assigned to authorities by common consent," and "others turn those opinions intentionally on their heads, apparently as challenges to rival philosophers." Moreover, "Pico drew some of his theses from epitomes, anthologies or florilegia, or even wholly from oral sources" and indeed others can be found that "combine materials from his sources in a highly idiosyncratic fashion, making it impossible again to point to one passage or another as his immediate source," since "he often compressed ideas spread out over dozens of pages […] into the exaggerated correlative forms characteristic of his own thought (192–93).

However, Farmer also admits that "in some cases, it is possible to follow Pico as he moves page by page through certain texts, drawing theses from various scholastic commentaries on Peter Lombard's Sentences or from favorite Greek sources" (192). Accordingly, we decided, for a start, to concentrate on medieval authors, considering that digitised corpora of their works are available for most of them. But also in these cases, we still have to face the difficulty of locating, for the majority of the theses, possibly exact references to the works of the author they are ascribed to.

2. The research activity of the ARTFL Encyclopédie Project at the University of Chicago[1] offers an instructive case of the application of digital-assisted approaches to text analysis in the successive stages of development of computational tools applied to textual studies. Scholars working at the project have all along been fully aware of the kind of opportunities offered to researchers by the diverse technologies available from time to time. The project avails itself of *PhiloLogic*, a tool developed specifically to work on the Treasury of the French Language, which provides full-text search, retrieval and analysis functionalities and has been continually upgraded to allow the collaborative development of higher-level, interoperable tools for Humanities Computing applications.[2]

ARTFL scholars discern three successive main stages of development of computational tools that enabled three distinct hermeneutical approaches, focussing respectively on "the interaction of digital methods with the text, the context, and the intertext" (Roe 2014, 89).[3] The first approach, practised already far before the digital turn, consists in the traditional provision of *concordances*, usually presented in digital form as Key Words In Context (KWICs). This method "is the oldest, simplest, and in many ways the most powerful tool of navigation and exploration of the texts" and it gives the opportunity "to compare different uses of a word in the context of the other words that surround it" (Morissey et al. 2016, 603). Accordingly, an interactive concordance programme affords "instant searching for words or combinations of words through large sets of texts" (604), so that "it gives an idea of their use and the lexical field in which they occur most frequently" (Roe 2014, 92). But together with those advantages, this approach shows also its limitations: "the specificity of the keyword and of its particular context drops gradually out of sight" (Roe 2014, 94), when dealing with increasingly larger corpora. Therefore, new analytical tools for search and retrieval are needed.

In the succeeding phase the attention was directed to implementing more efficient search and retrieval tools to explore word contexts in large digitised text collections. The study of context became the chief concern and *distant reading* was the "response to the growth of the digital contexts" (Roe 2014, 95). The resulting approach to text became overwhelmingly quantitative and induced analytical practices that were severely undermining textual integrity by

[1] *ARTFL Encyclopédie Project*. Accessed 4 March 2017. https://encyclopedie.uchicago.edu/.

[2] *PhiloLogic*. Accessed 5 March 2017. https://sites.google.com/site/philologic3/home.

[3] All translations of texts in French are ours.

"reducing the texts or parts of the texts to one or several 'bags of words' " (Morissey et al. 2016, 606). However, a capacity arose "to consider not only the raw frequency of words, but also their lexical field" (Roe 2014, 97). This led up to developing data mining tools, mostly based on machine learning. The resulting algorithms were then more "concerned about the frequency of words in relation to each other" and the "frequency distribution of words" (Morissey et al. 2016, 607). "Automatic classifiers" (Roe 2014, 99) were implemented to classify documents and to "establish ontologies that allow to navigate through massive textual collections," by building "a representation of the discursive content of each of these classes" (Morissey et al. 2016, 607). Moreover, by representing the frequencies of words as vectors in an *n*-dimensional space, algorithms were developed, that "seek to identify similar content by measuring lexical distance between items in a vector space" (611–12). So, "lexical similarity measures" (Roe 2014, 101) can be defined, since "the proximity of two vectors indicates a similarity in terms of the vocabularies and the frequency distribution of these vocabularies" (Morissey et al. 2016, 612).

Thus, since concordances hardly help in the exploration of large corpora, new methods have been introduced to inspect word contexts and vocabularies through the classification of texts, the development of ontologies, and the identification of lexical similarities. To support these new methods, new text technologies have been developed to "help the mobility between the macroanalysis or distant reading of massive collections and the microanalysis or close reading of the digital texts," in brief, to afford the interplay "between the text and the context, through the intertext" (Roe 2014, 103-104).

Finally, as regards the investigation of intertextual relations, the "heuristic value" (Morissey et al. 2016, 609) of such new methods is indisputable, but the "limitations" of these tools are equally obvious, since they tend to "remove the order of deployment of the words, so essential to the construction of sense" (613). Accordingly, in a later phase, sequence alignment techniques were used to identify similar passages. Glenn Roe quotes Roland Barthes — "every text is an intertext [...] every text is a new web of past quotations" (Barthes 1973, 1013) — to stress the hermeneutical significance of a thorough investigation of *text reuse*. Intertextuality, then, becomes the "theoretical foundation" of further digital implementations (Roe 2014, 104).

3. Finding the sources of Pico's theses in an author's corpus is a matter of context and intertextuality. So how can we proceed in our quest? The method of concordance could help, but it would necessarily require an exceedingly long human intervention and a painstaking close reading to identify the relevant contexts of a given key term. The method of alignment would bring optimal results, but very seldom a thesis is likely to turn out as an exact quotation, so that looking for text reuse by sequence-similarity searching would hardly produce good overall returns. Text mining methods would then seem the most productive. This approach comprises a wide range of algorithms and machine learning techniques and testing them by comparing their results seemed to be the preferable attitude, even though applying vector space models would look as the most promising procedure.

The vector space approach comprises, in turn, a great variety of algorithms, ranging from cosine similarity, count-based, predictive, random-vector, and so-called word-embedding models. Among them, the recently developed "popular software package" (Recchia 2016) word2vec (Mikolov et al. 2016) is widely employed for its "ease of use and state-of-the-art performance" (Recchia 2016). From a methodological point of view, it may be interesting to observe that word2vec implements a prediction-based model, which relies on an "intuition," namely that "words that share many contexts will be similar to each other." This intuitive step may indeed look quite "hand-wavy" (Goldberg and Levy 2014, 5) and it has been demonstrated (Levy and

Goldberg 2014) that "despite the apparent gulf between the inner workings of prediction-based and count-based models" what word2vec is doing, "mathematically speaking, is not much different than the sort of thing that some count-based models have been doing for awhile" (Recchia 2016). But in the face of these critical assessments, the same authors have nevertheless acknowledged that this "word embedding method" still "remains superior" in some respects (Levy and Goldberg 2014, 2177) to traditional natural language processing methods, such as distributional semantics solutions. The decisive intuitive step seems then to endow the word2vec approach of an indisputable heuristic force that makes it look preferable to other context analysis techniques.

4. Another approach worth exploring is that of The Concept Lab, a research project at the University of Cambridge led by Peter de Bolla. [4] This project aims at "tracking both the history of conceptual forms and their architectures by using data derived from digital archives" (de Bolla 2013, 2) and "is committed to the view that concepts are not equivalent to the meanings of the words which express them" (de Bolla 2015). Theories that treat concepts "as essentially in one-to-one correspondence to word senses" are therefore rejected (Recchia et al. 2017, 2) and the focus is set on conceptual structures and their internal and external manifestations, namely on their "form and function" as well as on "the conceptual networks within which concepts circulate" (de Bolla 2015).

The "distinction between word senses and concepts" is here deemed to be an important one, because "word senses change over time," but "a change in the frequency or lexical associations of a particular word does not necessarily entail a change in the concept" expressed in in the text (Recchia et al. 2017, 3). This circumstance is particularly relevant in relation to Pico's theses, since his vocabulary, depending as it does on the need to demonstrate the genuine concordance of all philosophical positions, is not necessarily the same as that of the authors they are ascribed to. As we have seen by quoting his "Preface," Pico is well aware that the expressive language he uses is a matter of choice. To cope with this problem, the Cambridge research project refers to the Concept Through Time (CTT) model that stems from the collaboration between computational linguists and cultural historians.

The CTT model is based on the assumption that "the vocabulary used in the debate on a concept may change over time" and intends to "focus on the stabilization and destabilization of relational vectors between words, i.e., the emergence and disappearance of words within a subset, as well as the shifting position of words in relation to one another" (Wevers et al. 2015). But does a change in the lexical associations of a given word "merit the addition" of a new concept, or "is this simply novel language for describing an old idea?" (Recchia et al. 2017, 3). At the present time, work towards finding a fitting algorithm to solve this problem is still in progress. The developers of CTT point out that it "a successful system for monitoring vocabulary shifts over time should strike a balance between an adaptive strategy that responds to changes in vocabulary, and a more conservative approach that keeps the vocabulary stable" (Kenter et al. 2015, 1191). In a similar vein, the researchers of the Cambridge project maintain that a model should be "flexible" enough (since "words whose meanings shift away from the conceptual core should drop out") and at the same time suitably "stable" (since "conceptual networks" should not "drift" significantly "away from the original conceptual core") (Recchia et al. 2017, 6–7).

[4] *The Concept Lab*. Accessed 15 March 2017. https://ccdkconceptlab.wordpress.com/.

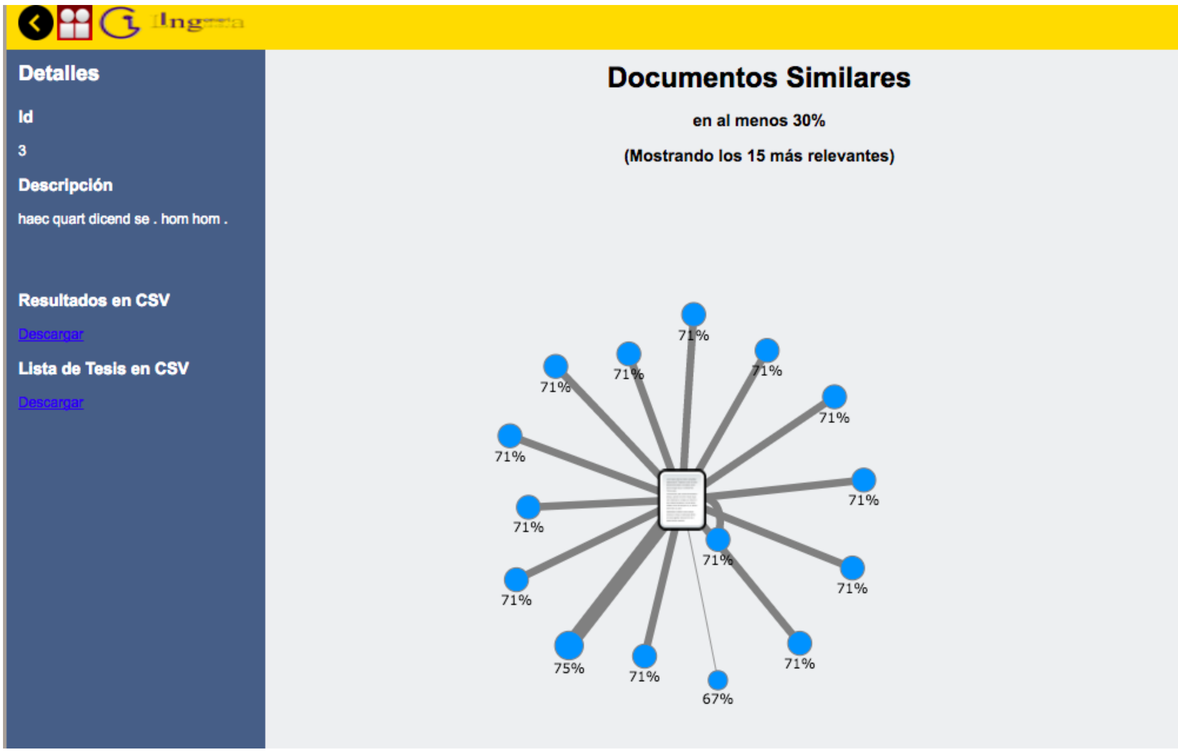
These being the options at hands, the decision was taken to test different methods, even though the vector space model—and word2vec in particular—would seem the most promising approach.

5. A first probing was conducted at the Universidad Nacional Autonoma de México (UNAM), in collaboration with the Grupo de Ingeniería Linguistica of the Instituto de Ingeniería, to locate the "sixteen conclusions" according to Albert the Great (Farmer and Pico 1998, 212–17) in his online corpus.[5] As an example, let us consider Pico's first thesis credited to Albert:

> Species intelligibiles non sunt necessariae, et eas ponere non est bonis peripateticis consentaneum. (Farmer and Pico 1998, 211)

Non-lexematic and categorematic words such as *non*, *sunt*, *et*, *eas*, *est* were excluded, in order to keep only those words that could help in identifying specific contexts and passages in Albert's works. The remaining words, *species*, *intelligibiles*, *necessariae*, *ponere*, *bonis*, *peripateticis*, and *consentaneum* were used as a 'bag of words' to search the *Alberti Magni e-corpus*. The search retrieves a collection of selected paragraphs of Albert's works that can be used to make a comparison with our thesis to identify relevant contexts. We chose the Cosine Similarity method to compare Pico's theses with passages retrieved from Albert's e-corpus, because it is the most straightforward procedure to detect text similarities. Moreover, it gives us the opportunity to compare the results we obtained, with those arrived at applying different methods. The results of our experiment are accessible online at the following address: http://www.corpus.unam.mx/tesis_latin/.

So, for instance, the 15 more relevant results concerning the third thesis on Albert the Great can be seen through the visual interface of the system, as shown here in Figure 1:



[5] *Alberti Magni e-corpus*. Accessed 17 March 2017. http://albertusmagnus.uwaterloo.ca.

— Fig. 1 —

This particular example is absolutely cogent to expose the difficulties depending on the use, as a term of comparison, of short sentences such as Pico's theses are, because the results we get are not decisive: in this case, the fact that 14 over the 15 highest results share the same score shows that the chosen bag of words is bound to yield a far too large number of similarity results. Consequently, for our purposes, the performance of the simple cosine similarity method turns out too low. To obtain significant results other and more efficient computational models have to be employed.

6. A further step in our testing may then consist in applying topic modelling and vector-space word-embedding techniques, such as the MALLET, a Latent Dirichlet Allocation, or LDA, based toolkit (McCallum 2002, Graham and Milligan 2012), and the word2vec packages.[6] This approach may not be decisive for taking into account vocabulary drifting and for locating possible sources of Pico's theses by retrieving exact passages referred to, but it would provide by all means reliable results in identifying recurring contexts and motifs, both in Pico's works and in his sources, that could be quite serviceable for hermeneutic purposes. The heuristic power of this method is of substantial support to any text analysis operation, such as producing linked open data, ontologies and annotation languages.

In more detail, a topic modelling analysis would allow a segmentation of the corpus under scrutiny into several sections relating to specific topics. A classification of the topics is then possible and can be structured hierarchically in a rigorous way as a formal ontology, that can provide a powerful means of navigation through the entire corpus. Topic modelling can then be employed as a classification tool, with results that, in the case of the French *Encyclopédie*, have been found "consistent with […] previous machine classification experiments" (Roe et al. 2014, 337) conducted on the same text (Horton et al. 2009). But the topic modelling results enable us to work also within the sections of the texts that deal with a specific topic. Distant and close reading can then meet and in defiance of the common opinion they can effectively show their complementarity.

According to the topic modelling approach, a topic is seen as bag of co-occurring words and we can deal with them in many useful ways. The full heuristic force of the topic modelling outcomes deserves to be thoroughly exploited. To give an example, we may come back again to the comparison of topic modelling and conventional techniques in computational linguistics. The latter can be conceived of as utterly quantitative, whereas the former is better thought of as an adaptive framework, relying on anticipatory assumptions. So, for instance, in purely computational linguistics terms, a Pointwise Mutual Information (PMI) measure—taken as a co-occurrence-based method to determine word relatedness—is seen as a function of two probabilistic events evaluating semantic similarity. This is apparently what word2vec also does, since it has been shown that its word-embedding prediction-based method "is implicitly factorizing a shifted PMI matrix—the well-known word-context PMI matrix from the word-similarity literature, shifted by a constant offset" (Levy and Goldberg 2014, 2177). But this actually means that the same data can be seen on the one hand as the representation of a pure matter of fact, and on the other as endowed of the anticipating power to disclose a full range of possible still unnoticed word relations. And it is precisely in exposing them that a close-reading approach comes of use.

[6] Moody (2016) has recently proposed a mixing of the two models called *lda2vec*. Similar models that aim at combining LDA and word2vec have been presented by Das et al. (2015) and Batmanghelich et al. (2016).

Co-occurrence detection can thus be exploited to define a topic, or to produce RDF statements and populate a triple store repository, to build and increase graph databases in order to store and manage linked data, or again to introduce possible and new ontology classes. These activities can be carried out either manually, or alternatively in a semi-automatic or fully automated way. Thus, a controlled annotation language, obtained out of topic modelling and co-occurrence results, would possibly enable an intermediary course of action between manual and automated procedures, consisting for instance in a formalised annotation practice.

7. A promising attempt to process annotation data comes from machine learning and artificial intelligence (AI), through the application of probabilistic soft logic (PSL).[7] Probabilistic soft logic is a recent development in the "field of statistical relational learning (SRL)" (Getoor and Taskar 2007), a successful approach to treating "noisy, or uncertain, multi-relational data" produced by manual or automated information-extraction procedures. In our case, topic modelling results could be mapped onto RDF linked data graphs and processed through PSL programmes, such as have been profitably applied to "mining" annotation graphs. Their outcome produces consistent "graph summaries" quite helpful to "handle uncertainty" present within this sort of "sources of evidence about concepts." For graph summaries can actually "group entities and relations based on similarity" and create "a graph at a higher level of abstraction" (Memory et al. 2012, 75), well suited to "filter out noise" and help with "the identification of structure and meaning" in the ingested data (Liu et al. 2017, A:1–2).

This example can be generalised, since the output of manual annotation practices and automated information-extraction systems is usually "spurious" and "often hampered by noise." On that account, the "statistical relational learning approach" (SRL) is specifically aimed at removing undesirable "noise." From an SRL point of view, to solve that problem, is to solve "the problem of knowledge graph identification," a process that "infers a knowledge graph" (Pujara et al. 2015, 65–66) from "the noisy output of an information-extraction system" (Pujara et al. 2013, 556). SRL techniques "incorporate statistical information and logical dependencies" (Pujara et al. 2015, 65–66) and formalise collective probabilistic reasoning in relational domains by building knowledge graphs, or "probabilistic graphical models" (Bach et al. 2015, 2) of "relational data, i.e., data composed of entities and relationships connecting them" (47). The end result is a more consistent representation of the "logical dependencies" (17) between a set of entities (concepts, in our case), their attributes, and their relations, that can be handled by formal ontologies enabling further reasoning to derive conclusions not explicitly expressed in the resulting knowledge-representation data structures.

Probabilistic soft logic (PSL) is also directly employed to develop a "programming framework for designing custom topic models," called "latent topic networks" (LTNs) that can be employed to explore "influence between scientific articles" (Foulds et al. 2015, 778), or for that matter literary and philosophical texts in general. A crucial step in the development of LTNs and other PSL applications, consists in the use of potential functions, or semantic constraints, to handle knowledge graph identification, a process that "incorporates semantics, in the form of an ontology," and employs "ontological constraints as weighted rules" (Pujara et al. 2015, 65–66). As a matter of fact, the "semantic dependencies" between nodes in a knowledge graph (66) are expressed "in the form of a potential function or constraint, along with any necessary parameters, such as the weight of the potential" (Bach et al. 2015, 13). This is done because, "if a modeler does not know how the domain behaves, the potentials should capture how it might behave, so that a learning algorithm can find weights that lead to accurate predictions" (4). Learning

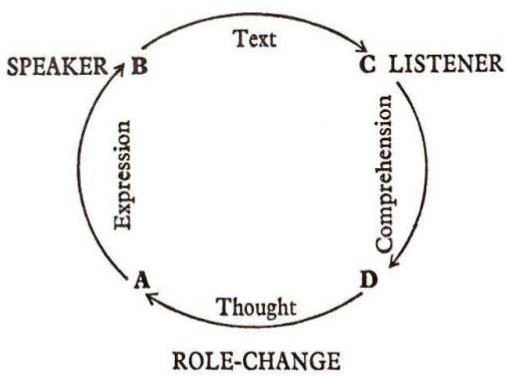
[7] *Probabilistic soft logic*. Accessed 18 March 2017. http://psl.linqs.org/.

algorithms are designed to enable inferences that comply with the assigned potential semantic constraints, so as to identify abstract probabilistic dependencies. Semantic constraints are then treated as "learned prediction functions to assign a confidence score" to uncertain dependencies between concepts: virtually, they are used "as 'hints' to find" the "correct" dependencies, in order to establish reliable knowledge graphs (Pujara et al. 2015, 66).

The whole graph identification process "requires combining two disparate elements: the statistics output by an information extraction and ontological constraints derived from the semantics of the knowledge graph" (68). Accordingly, it qualifies as a new development of the "mechanism known as Bayesian updating," a process in which "information is used to adjust the initial prediction (=prior belief) to the reality of the environment, resulting in a new adapted belief about the world (posterior belief)" (De Ridder et al. 2013, 1). In our case, 'the reality of the environment' consists of textual data, but again—as we have seen in regard to the Concept Through Time (CTT) model—by examining lexical dependencies are we dealing with concepts and ontology classes, or are we dealing with words and phrases in our texts? The point is crucial and deserves, for our purposes, a little more attention.

As is well known, Bayesian procedures admit of quite divergent interpretations, ranging from a strictly objectivist to a purely subjectivist view. Are we trying to find what is actually there, or are we just measuring the risk of our bet? Our surmise is that the two opposite points of view may even be compatible, depending on one's intentions and on where one's interests lie. Is the face you spot in a cloud actually there, or is it just a fiction of your imagination? The latter may indeed be the case, but you would not see anything if the cloud had not been there. If the cloud is made of words, the game is open. Out of metaphor, data mining applications using artificial adaptive systems (Tastle 2013), such as PSL models are, move from a principle derived from an analogy with natural language: "Exactly as through writing a natural language can create cultural objects... in the same way the Artificial Sciences can create through the computer automatic models," whose dynamics develops somewhat autonomously and independently from the observed data: "models... generate rules dynamically" in a way "similar to the Kantian transcendental rules" (Buscema 2011, 17-20). Models can then be construed either as cultural objects, in certain respects independent from what they represent, or as objective representations of what is being described. There is both an active, so to speak, and a passive aspect in their construction.

The same duality is present in the practice of textual interpretation. According to the "conversational cycle" (Fig. 2) proposed by Frederick Parker-Rhodes in his *Inferential Semantics* (1978, 16), the "thought that the speaker had intended to convey... will be processed in the listener mind, and possibly result in the elaboration of a new thought" (17).



— Fig. 2 —

This means that the comprehension process, or for that matter the interpretation of a text, produces a model of the text not totally dependent from the actual process of its expression. It is this kind of partial autonomy from the effective state of affairs that endows the model of its significant heuristic force. The epistemological rationale behind our last observations consists in admitting "the possibility... of presenting the difference of system and environment within the system" (Luhmann 1990, 12). We think it proper to the textual condition that both the model and the textual data be represented, and interpreted, within one and the same system.

Back to the application of PSL to text analysis and natural language processing, we can now show that both close and distant reading can play similar roles in respect of quantitative as opposed to qualitative, or factual as opposed to conceptual approaches, and that they can both be regarded as complementary practices. In PSL applications, input data are always assumed as strictly quantitative and purely observed ones, but they may be provided either by text mining and distant reading results, or by the outcomes of annotation and close reading practices. Accordingly, both close reading and distant reading provide, in this case, factual information. On the other hand, a conceptual and interpretational step may be essential in producing the outcome of specific distant- or close-reading practices. This is patent in a close-reading approach, where a reflective aspect is always present, but it turns out to be so also in distant-reading data-processing approaches, as we have seen in the case of word2vec. Likewise, in the application of PSL to text analysis and natural language processing, an anticipatory step can be singled out in the choice of semantic constraints that identify links of potential dependence between words and phrases in a text—or, in more technical terms, between the logical atoms and predicates that constitute the nodes and edges of a probabilistic knowledge graph.

8. To sum up, what we have proposed so far is a tentative survey of possible models for a digital approach to the study of Pico's works, but whatever the methods we choose, the focus of computational analysis, be it cast on the hermeneutics of the text, of its context, or its intertextual connections, must preserve the distinctive character of that kind of critical reflection, as much conceptual as factual, which is proper of the humanities frame of mind.

Works cited

Bach, Stephen H., Matthias Broecheler, Bert Huang, and Lise Getoor. 2015. "Hinge-Loss Markov Random Fields and Probabilistic Soft Logic." arXiv:1505.04406 [cs.LG]

Barthes, Roland. 1973. "Texte, théorie du." In *Encyclopaedia Universalis*, vol. 15, 1013–17. Paris: Encyclopaedia Universalis France.

Batmanghelich, Kayhan N., Ardavan Saeedi, Karthik Narasimhan, and Samuel Gershman. 2016. "Nonparametric Spherical Topic Modeling with Word Embeddings." In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Berlin, August 7-12, 2016*, vol. 2, *Short*

Papers, edited by Katrin Erk and Noah A. Smith, 537–542. Stroudsburg PA: The Association for Computational Linguistics. Accessed 22 March 2017. http://anthology.aclweb.org/P16-2087.

Buscema, Massimo. 2011. "Artificial Adaptive Systems: Philosophy, Mathematics and Applications." In *Advanced Networks, Algorithms and Modeling for Earthquake Prediction*, edited by Massimo Buscema and Marina Ruggieri, 17–37. Aalborg, Denmark: River Publishers.

Das, Rajarshi, Manzil Zaheer, and Chris Dyer. 2015. "Gaussian LDA for Topic Models with Word Embeddings." In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, ACL 2015, July 26-31, 2015, Beijing, China*, vol. 1, *Long Papers*, edited by Chengqing Zong and Michael Strube, 795–804. Stroudsburg PA: The Association for Computational Linguistics, and Red Hook, NY: Curran Associates, Inc. Accessed 21 March 2017. http://www.aclweb.org/anthology/P15-1077.

De Bolla, Peter. 2013. *The Architecture of Concepts: The historical formation of human rights*. New York: Fordham University Press.

–––. 2015. "About the Concept Lab." Cambridge: Centre for Research in the Arts, Social Sciences and Humanities. Accessed 15 March 2017. http://www.crassh.cam.ac.uk/publications/post/about-the-concept-lab.

Farmer, Stephen Alan, and Giovanni Pico della Mirandola. 1998. *Syncretism in the West: Pico's 900 theses (1486): The evolution of traditional, religious, and philosophical systems*. Tempe (AZ): Medieval & Renaissance Texts & Studies.

Foulds, James, Shachi Kumar, and Lise Getoor. 2015. "Latent Topic Networks: A Versatile Probabilistic Programming Framework for Topic Models". In *Proceedings of the 32nd International Conference on Machine Learning, ICML'15, Lille, July 6-11, 2015*, edited by Francis Bach and David Blei, Proceedings of Machine Learning Research, vol. 37, 777–86. Brookline, MA: Microtome Publishing on behalf of JLMR (The Journal of Machine Learning Research). Accessed 2 April 2017. http://proceedings.mlr.press/v37/foulds15.pdf.

Getoor, Lise, and Ben Taskar. 2007. *Introduction to statistical relational learning.* Cambridge (MA): MIT press.

Goldberg, Yoav, and Omer Levy. 2014. "word2vec Explained: Deriving Mikolov et al.'s Negative-Sampling Word-Embedding Method." arXiv:1402.3722 [cs.CL].

Graham, Shawn, and Ian Milligan. 2012. "Review of MALLET, produced by Andrew Kachites McCallum." *Journal of Digital Humanities* 2 (1). Accessed 21 March 2017. http://journalofdigitalhumanities.org/2-1/review-mallet-by-ian-milligan-and-shawn-graham/.

Horton, Russell, Robert Morrissey, Mark Olsen, Glenn Roe, and Robert Voyer. 2009. "Mining Eighteenth Century Ontologies: Machine Learning and Knowledge Classification in the *Encyclopédie*." *Digital Humanities Quarterly* 3 (2). Accessed 21 March 2017. http://www.digitalhumanities.org/dhq/vol/3/2/000044/000044.html.

Kenter, Tom, Melvin Wevers, Pim Huijnen, and Maarten de Rijke. "Ad Hoc Monitoring of Vocabulary Shifts over Time." In *Proceedings of the 24th ACM International Conference on Information and Knowledge Management, CIKM 2015*, 1191–1200. New York: ACM. doi: 10.1145/2806416.2806474.

Kimmig, Angelika, Stephen H. Bach, Matthias Broecheler, Bert Huang, and Lise Getoor. 2012. "A short introduction to probabilistic soft logic." In *Proceedings of the NIPS Workshop on Probabilistic Programming: Foundations and Applications*, edited by Vikash K. Mansinghka, Daniel M. Roy, Noah D. Goodman, 1–4. Accessed 6 April 2017. http://stephenbach.net/files/kimmig-probprog12.pdf.

Levy, Omer, and Yoav Goldberg. 2014. "Neural Word Embedding as Implicit Matrix Factorization." In *Advances in Neural Information Processing Systems 27 (NIPS 2014)*, edited by Zoubin Ghahramani, Max Welling, Corinna Cortes, Neil D. Lawrence, and Kilian Q. Weinberger, 2177–2185. Red Hook, NY: Curran Associates, Inc. Accessed 18 March 2017. http://papers.nips.cc/paper/5477-neural-word-embedding-as-implicit-matrix-factorization.pdf.

Liu, Yike, Abhilash Dighe, Tara Safavi, and Danai Koutra. 2017. "Graph Summarization: A Survey." arXiv:1612.04883v2 [cs.IR.].

Luhmann, Niklas. 1990. *Essays on self-reference*. New York: Columbia University Press.

McCallum, Andrew Kachites. 2002. MALLET: "A Machine Learning for Language Toolkit." *UMass Amherst*. Accessed 21 March 2017. http://mallet.cs.umass.edu/.

Memory, Alex, Angelika Kimmig, Stephen H. Bach, Louiqa Raschid, and Lise Getoor. 2012. "Graph Summarization in Annotated Data Using Probabilistic Soft Logic." In *Proceedings of the 8th International Workshop on Uncertainty Reasoning for the Semantic Web, Boston, USA, November 11, 2012*, edited by Fernando Bobillo, Rommel Carvalho, Paulo C. G. da Costa, Nicola Fanizzi, Kathryn B. Laskey, Kenneth J. Laskey, Thomas Lukasiewicz, Trevor Martin, Matthias Nickles, Michael Pool, CEUR Workshop Proceedings, vol. 900, 75–86. Aachen: CEUR-WS.org. Accessed 5 April 2017. http://ceur-ws.org/Vol-900/paper7.pdf.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Efficient Estimation of Word Representations in Vector Space." arXiv:1301.3781v3 [cs.CL].

Moody, Christopher E. 2016. "Mixing Dirichlet Topic Models and Word Embeddings to Make lda2vec." arXiv:1605.02019 [cs.CL].

Morrissey, Robert, Glenn Roe, and Clovis Gladstone. 2016. "La littérature à l'âge des algorithms." *Revue d'histoire littéraire de la France* 116 (3): 595–618.

Parker-Rhodes, Arthur Frederick. 1978. *Inferential semantics*. Hassocks, Sussex: Harvester Press.

Pujara, Jay, Hui Miao, Lise Getoor, and William W. Cohen. 2013. "Knowledge Graph Identification." In *Semantic Web – ISWC 2013: 12th International Semantic Web Conference, Sydney, NSW, Australia, October 21-25, 2013: Proceedings*, Part I, edited by Alani, Harith, Lalana Kagal, Achille Fokoue, Paul Groth, Chris Biemann, Josiane Xavier Parreira, Lora Aroyo, Natasha Noy, Chris Welty, Krzysztof Janowicz, Lecture Notes in Computer Science, vol. 8218, 542–557. Berlin: Springer.

———. 2015. "Using Semantics and Statistics to Turn Data into Knowledge." *AI Magazine* 36 (1): 65–74.

Recchia, Gabriel□ Ewan Jones. 2016. "Numberless Degrees of Similitude: A Response to Ryan Heuser's 'Word Vectors in the Eighteenth Century, Part 1'." Blog entry. Accessed 10 March

2017. http://www.twonewthings.com/gabrielrecchia/2016/06/11/numberless-degrees-of-similitude-word-vectors/.

−−−, Paul Nulty, John Regan, and Peter de Bolla. 2017. "Tracing Shifting Conceptual Vocabularies Through Time." In *Proceedings of Drift-a-LOD 2016: The first workshop on Detection, Representation and Management of Concept Drift in Linked Open Data, Bologna, 20 November 2016*, edited by Laura Hollink, Sándor Darányi, Albert Meroño Peñuela, and Efstratios Kontopoulos, CEUR-WS, CEUR Workshop Proceedings, vol. 1799, 2–9. Aachen: CEUR-WS.org. Accessed 4 March 2017. http://ceur-ws.org/Vol-1799/Drift-a-LOD2016_paper_1.pdf.

Roe, Glenn. 2014. "L'étude littéraire à l'ère du numérique: du texte à l'intertexte dans les 'digital humanities'." In *Literaturwissenschaft im digitalen Medienwandel*, edited by Christof Schöch and Lars Schneider, [85–111]. PhiN Beiheft, 7. Accessed 4 March 2017. http://web.fu-berlin.de/phin/beiheft7/b7t06.pdf.

Roe, Glen, Clovis Gladstone, and Robert Morrissey. 2014. "Discourses and Disciplines in the Enlightenment: Topic Modeling the French Encyclopédie." In *Digital Humanities 2014: Conference Abstracts*, 336–38. Lausanne: EPFL – UNIL.

Tastle, William J. 2013. *Data Mining Applications Using Artificial Adaptive Systems*. New York, NY: Springer.

Wevers, Melvin, Tom Kenter, and Pim Huijnen. 2015. "Concepts Through Time: Tracing Concepts in Dutch Newspaper Discourse (1890−1990) Using Word Embeddings." In *DH2015: Global Digital Humanities*. Sydney: University of Western Sydney. Accessed 15 March 2017. http://dh2015.org/abstracts/xml/WEVERS_Melvin_Concepts_Through_Time__Tracing_Conc/WEVERS_Melvin_Concepts_Through_Time__Tracing_Concepts_I.html.