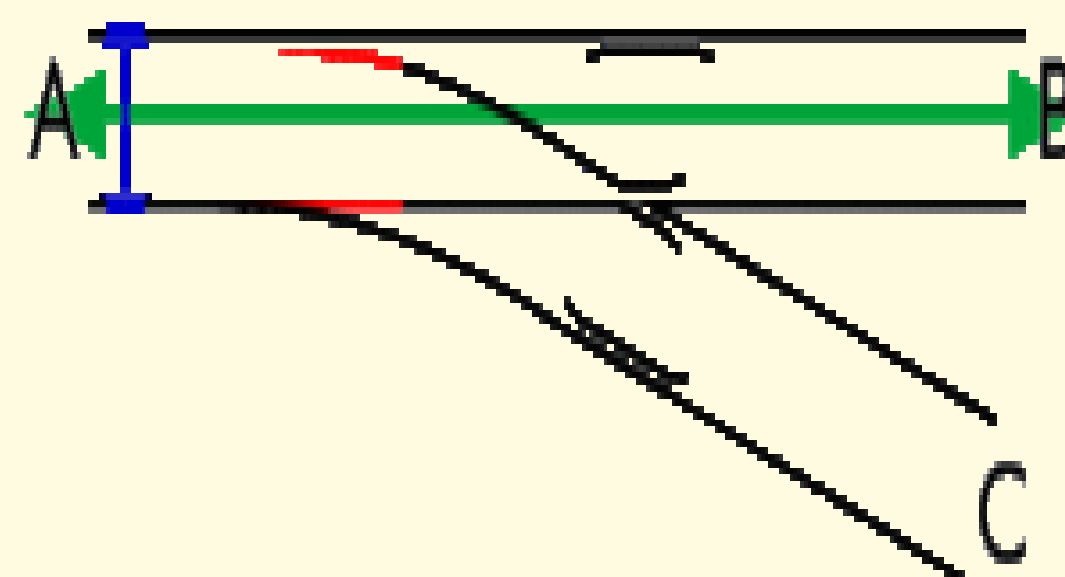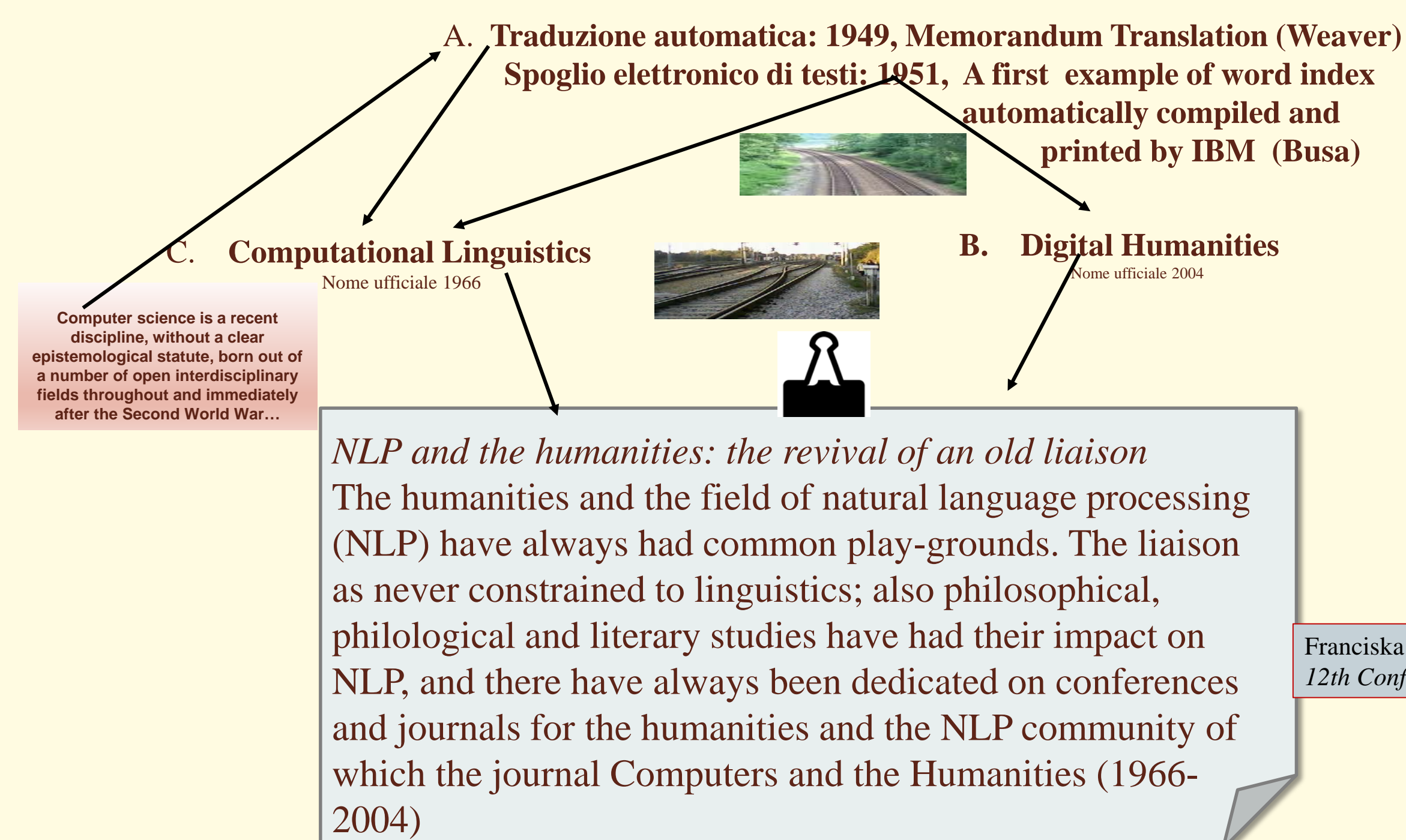# Strolling around the dawn of Digital Humanities

Gabriella Pardelli*, Sara Goggi*, Federico Boschetti*
Istituto di Linguistica Computazionale, "A. Zampolli", CNR-ILC, Italy *

Linguistica Computazionale e Digital Humanities: parole dai testi

A. **Traduzione automatica: 1949, Memorandum Translation (Weaver)**
Spoglio elettronico di testi: 1951, A first example of word index
automatically compiled and
printed by IBM (Busa)

C. **Computational Linguistics**
Nome ufficiale 1966

B. **Digital Humanities**
Nome ufficiale 2004

Computer science is a recent discipline, without a clear epistemological statute, born out of a number of open interdisciplinary fields throughout and immediately after the Second World War…

NLP and the humanities: the revival of an old liaison
The humanities and the field of natural language processing (NLP) have always had common play-grounds. The liaison as never constrained to linguistics; also philosophical, philological and literary studies have had their impact on NLP, and there have always been dedicated on conferences and journals for the humanities and the NLP community of which the journal Computers and the Humanities (1966-2004)

Franciska de Jong (2009). NLP and the humanities: the revival of an old liaison. *Proceedings of the 12th Conference of the European Chapter of the ACL* (EACL 2009), 10-15.

A junction, in the context of rail transport, is a place at which two or more rail routes converge or diverge. This implies a physical connection between the tracks of the two routes
…(https://en.wikipedia.org/wiki/Junction_(rail))

LE PAROLE DAI TESTI: IL PERIODO PIONIERISTICO
Nel decennio pionieristico non esisteva un termine generale per identificare l'elaborazione elettronica di dati linguistici per lo spoglio di testi e tecniche di analisi testuale nelle discipline umanistiche (fine anni '40 – inizio anno '60)
Non sono sicuramente e ampiamente note al grande pubblico le prime e fondamentali conferenze internazionali organizzate negli anni '60

Des travaux de type artisanal, nous passon brutalement aux techniques industrielles. L'utilisation des machines a introduit dans notre ordre de recherches un bouleversement total.
(Quemada, 1964) BESANÇON

Fasi preliminari delle ricerche linguistiche:
a) La transcription de textes naturels, à partir du livre imprimé, sur bande magnétique en vue d'élaborations électroniques…
b) …la compilation d'indices et de concordances de mots, de morphèmes, de graphèmes, de syntagmes, de fréquencies, etc…
(Busa, 1966) **PRAGA**

LES MACHINES DANS LA LINGUISTIQUE

… some humanists who have adopted computers are not yet fully adapted to them.
(Bessinger, 1964) **New York**

The Literary Data Processing Conference

The Literary Data Processing Conference brought together scholars in the humanities so that they could exchange knowledge and ideas in the field of literary data processing.

The Conference was organized by the Department of Scientific and Technical Information, Technology and Engineering Staff, IBM Corporation. It was sponsored jointly by that Department, IBM Data Processing Division - Industry Programs, and the Office of the IBM Director of Systems Research

LE PAROLE DAI TESTI: LA SECONDA DECADE
Il periodo che segue quello euristico e pionieristico sembra aver raggiunto un equilibrio in parte somministrato da una maggiore specializzazione degli umanisti in campo informatico

The second period lasted through the Seventies until the early Eighties. Humanities scholars learned now to write their own programs, as programming languages became easier to master
(Kenny, 1992).

Nasce nel 1966 la rivista *Computers in the Humanities*
- *Froger entusiasticamente commenterà: .. une revue specialisée, Computers in the Humanities …*
- Melic nel primo numero della rivista scriverà:
We define humanities as broadly as possible. Our interests include literature of all times and countries, music, the visual arts, folklore, the non-mathematical aspects of linguistics, and all phases of the social sciences that stress the human. When, for example, the archeologist is concerned with fine arts of the past, when the sociologist studies the non-material facets of culture, when the linguist analyzes poetry, we may define their intentions as humanistic; if they employ computers, we wish to encourage them and to learn from them… With the inauguration of Computers and the Humanities, the time has perhaps arrived for a more serious look at the position of the humanistic scholar in the world of data processing (Melic, 1966)

Terza generazione dei computer (1965-1980), uso di circuiti integrati. PASCAL e PL1 i nuovi linguaggi di programmazione:

There is no doubt that the third generation computers will eventually measure up to much that was promised. They are extraordinarily fast and can reach enormous amounts of data in immediately available storage; performance of more than one program at a time is feasible; the diversity of special devices such as voice response and graphic mechanisms that can be incorporated into a system are most impressive

(Wishy, 1967)

LE PAROLE DAI TESTI: LA TERZA DECADE
La terza decade, contrassegnata dall'avvento del Personal Computer e a seguire dallo sviluppo di Internet e dalla proliferazione di risorse digitali e di infrastrutture

Nasce nel 1986 la rivista *Literary and Linguistic Computing*

LITERARY & LINGUISTIC COMPUTING
Journal of the Association for Literary and Linguistic Computing

Nel 2004 pubblicazione del libro
A Companion to Digital Humanities
S. Schreibman, R. Siemens, and J. Unsworth (eds.). 2004. A Companion to Digital Humanities. Blackwell

Dal 2015 diventa
Digital Scholarship in the Humanities
DSH di ADHO e EADH

LLC

Language Resources and Evaluation

The term Digital Humanities has been commonly used to describe the application of computational methods in the arts and humanities for 10 years, since the publication, in 2004, of the Companion to Digital Humanities. "Digital Humanities" was quickly picked up by the academic community as a catch-all, big tent name for a range of activities in computing, the arts, and culture. [Terras, 2016]

When I talk about digital research products obviously I'm not referring to the classic article or monograph (although in digital form). Anyone who has really been involved in a research project in **Digital Humanities** knows that most of the research work consists in the modeling and creation of the data sets, computational analysis methods, tools and frameworks. If the goal of a research project is the creation of an archive of documents, a digital scholarly edition or a visualization tools to display historical data, those are the real products of research, not the articles describing them (often tedious and report style). The assessment of Digital Humanities research must apply to these objects. The creation of a platform of this kind requires: an ontology of the roles and functions performed by the researchers involved in a project; a set of minimum quality requirements (standards compliance, interoperability, scientific control, coverage of the domain, etc.) to assign a digital resource the status of research product; a metric based on these requirements to enable (Giotti, 2015)

LC&DH : IL LESSICO

<digital form>, <computational analysis methods, tools and frameworks>,
<digital resource the status of research product>, <visualization tools>,
<archive of documents>, <platform>, <ontology>,
<standards compliance>, <interoperability>

Dal 2005 diventa
Language Resources and Evaluation

Per le citazioni si veda la bibliografia nel Book of Abstracts AIUCD2019

Contacts: * gabriella.pardelli@ilc.cnr.it – federico.boschetti@ilc.cnr.it