HAL
open science

Roberto Busa (1913-2011), Pioneer of Computers for the Humanities

Corrado Bonfanti

To cite this version:

Corrado Bonfanti. Roberto Busa (1913-2011), Pioneer of Computers for the Humanities. Arthur Tatnall. Reflections on the History of Computing : Preserving Memories and Sharing Stories, AICT-387, Springer, pp.57-61, 2012, IFIP Advances in Information and Communication Technology (SURVEY), 10.1007/978-3-642-33899-1_4 . hal-01526801

HAL Id: hal-01526801

https://inria.hal.science/hal-01526801v1

Submitted on 23 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Roberto Busa (1913-2011), Pioneer of Computers for the Humanities

Corrado Bonfanti
Associazione Italiana per l'Informatica ed il Calcolo Automatico (AICA), Italy
corradobonfanti@hotmail.com

Foreword: Roberto Busa, *Societatis Iesu*, passed away on August 9, 2011.

I will commit a memory of such a remarkable and unusual personality to a concise – as well as arbitrary – excerpt from his own writings.

Before doing that, I ought to recall the last opportunity I had to meet Father Busa.

In the springtime of 2008, at the outset of a conference, I had the pleasure to drive him in my small car to the airport whence he had to depart. (Aged 95, he still travelled unperturbed on his own!) After so many years of formally respectful friendship, I suddenly felt the need to loyally disclose to him my personal distance from religions, included the one he so vividly professed. He listened quietly, without any comment. Then, at the moment of departure, he unexpectedly said "It's time we call us simply by name. Corrado – he smiled – you are free to call me father Roberto".

Keywords: Humanities, computing, informatician, life history.

1. A Priest and an Informatician

I am aware that the appearance of a priest[1] into the *ambulacra* of technology might appear rather bizarre: I feel I looked like a dromedary slipping into the stock exchange.

I ought to tell you: firstly, it was not my own choice but obedience to an assignment or mission; and this gives great peace and strength.

Secondly, also and precisely because I am a priest, I feel comfortable there: not only because of sympathy and admiration for informaticians; as a matter of fact, informatics is an interior and spiritual discipline.

Since man is child of God and technology is child of man, I think that God regards technology the way a grandfather regards his grandchild[2]. And for me personally it is satisfying to realize that I have taken seriously my service to linguistic research.

[1] Translated from [1], p.81. Last paragraph taken from [2], pp. 87-88.
[2] To be read 'information technology', according to context.

2. “It was not my own choice’’

I entered the Jesuit order[3] in 1933. I was then 20. Later my superior asked me: “*Would you like to become a professor?*” “*In no way!*” My wish was to be a missionary to take care of the poor. “*Good. You’ll do it, all the same.*”

By 1941, I had been assigned to work towards a PhD in Thomistic philosophy at the Pontificial Gregorian University in Rome. My research was aimed at exploring the concept of ‘presence’ according to Thomas Aquinas. While I was involved with this research, two major considerations became evident. I realized first that a philosophical and lexicographical inquiry into the verbal system of an author has to precede and prepare for a doctrinal interpretation of his work.

In 1946 as a result of these preliminary conclusions, I started to think of an *Index Thomisticus*, i.e. a concordance of all the words of Thomas Aquinas, including conjunctions, prepositions and pronouns, to serve other scholars for analogous studies. It was clear to me, however, that to process texts containing more than ten million words, I had to look for some type of machinery.

In 1949, I visited approximately 25 American Universities from coast to coast, asking about any gadget that might help in producing the type of concordance I had in mind. Jerom Wiesner of M.I.T., sent me to IBM in New York City, where someone was assigned to examine my project.

I knew, the day I was to meet Thomas J. Watson, Sr., that he had on his desk a report which said that IBM machines could never do what I wanted[4]. I had seen in the waiting room a small poster imprinted with the words: “*The difficult we do right away; the impossible takes a little longer*” (IBM always loved slogans). I took it with me into Mr. Watson’s office. Sitting in front of him and sensing the tremendous power of his mind, I was inspired to say: “*It is not right to say ‘no’ before you have tried.*” I took out the poster and showed him his own slogan. He agreed that IBM would cooperate with my project until it was completed ‘provided that you do not change IBM into International Busa Machines.’ I had already informed him that, because my superiors had given me time, encouragement, their blessings and much holy water, but unfortunately no money, I could recompense IBM in any way except financially. That was providential! [5]

Although some say that I am the pioneer of the computers in the humanities, such a title needs a good deal of nuancing. *A propos* of this, Mr. Lee Loevinger in the *Minnesota Law Review* (April 1949), in an article on jurimetrics said: “*Machines are now in existence which have so far imitated thought processes that they can solve differential equations. Why should not a machine be constructed to decide lawsuits?*” And on the stacks of the IBM library in New York City I had spotted a book (whose title I have forgotten), which was printed sometime between 1920 and 1940: in it someone mentioned that it was possible to make lists of names by means of punched cards. Maybe others too may claim that they have worked in this area prior to me.

---

[3] A combination of fragments taken from [2]

[4] Thomas J. Watson Sr. was the mighty IBM’s father-master. IBM machines, by that time, were punched card equipment.

[5] As a matter of fact, IBM generously supported Busa’s cultural enterprise until its completion.

Yet, isn't it true that all new ideas arise out of a *milieu* when ripe, rather than from any one individual? If I was not the one, then someone else would have dealt with this type of initiative sooner or later. To be the first one to have an idea is just a chance. If there is any merit, it is in cultivating the idea. During the following years I experienced the wisdom of another slogan, attributed to an American, Thomas Edison: "*Genius is 1% inspiration, 99% perspiration*."

In addition to the 10,600,000 words of *Index Thomisticus*, I processed five million more words, Italian, English, German, Russian, ancient Greek and Hebrew, Aramaic and Nabatean. The subjects ranged from nuclear physics and mathematics abstracts to Qumran [or Dead See] Scrolls and works of Dante, Kant, and Goethe. The processing of the Dead Sea Scrolls was publicized throughout the world on the front pages of newspapers in the spring of 1958. I was able to complete my *Index Thomisticus* 33 years after the conception of the project and 30 years after my first meeting with IBM. At that time it was the first undertaking in computer linguistics, the only published work of its kind with such dimensions and such characteristics.

3. Father Roberto in Numbers

Besides 245 congresses[6], I gave courses, lectures or conferences in 60 Italian sites, plus 40 in Europe, plus 17 in Northern America, plus 9 in Asia and Africa, plus 3 in Southern America. At this date [March, 2008], my written works are 421: 116 are books, amounting to 80,000 pages, 305 are articles and papers for more than 5,000 pages. Not counting those in Italian or Latin, 86 are written in or translated into Albanian, Hebrew, French, Georgian, English, Portuguese, Russian, Spanish, German.

> *Father Roberto avoided enumerating the awards he received along his life. I can testify his pride on being bestowed upon the highest honour of the Order of Merit of the Italian Republic; President Carlo Azeglio Ciampi created him a Knight Grand Cross in 2005.*
>
> *It is also worth mentioning that he was an invited speaker in a number of conferences held by AICA, the IFIP's federate Italian Computer Society. On the occasion of IFIP's WCC 2008 held in Milano, AICA issued an illustrated poster to recognize the Italians who markedly contributed to the progress of informatics and its applications; Roberto Busa is among them.*

4. Passages from his Life

Roberto Busa was born in 1913 in Contrada Busa di Lusiana[7] – near Vicenza, Northern Italy – whence the name of his family. He completed high school and the first two years of Theology at the episcopal seminary of Belluno, being there a

[6] Translated from [3].
[7] Translated, with minor adaptations, from [4].

schoolmate of Albino Luciani, later to become Pope Giovanni Paolo I. In 1933 he entered the Jesuit order, to be ordained priest in 1940. He gained the doctorate in philosophy (1937) as well as in theology (1941). In the years 1940-43 he served as a military auxiliary chaplain in the army and then with the partisan forces. At the Faculty of philosophy *Aloisianum* in Gallarate, near Milano, he was full professor of ontology, theodicy and scientific methodology; for several years he also served there as librarian.

In 1946 he conceived the project of an *Index Thomisticus*. In 1949 he started experimenting linguistic computerization at the IBM offices in New York and Milano. In order to manage and to exploit operations, he founded CAAL[8]. Operations were based at Gallarate until 1967; then for two years at Pisa and for two more years at Boulder, Colorado and finally, for nine years, in Venezia. Computerized photocomposition of the *Index Thomisticus* started there in 1974 and was concluded in 1980: 70,000 pages in 56 large size volumes.

During the same decades, the rhythm of international promotion for linguistic research and analysis was marked by the 143 congresses father Busa actively participated in, across three continents.

In 1967 prof. Antonio Zampolli – up to then an assistant of father Busa – founded in Pisa the Institute for computational linguistics which soon gained international reputation.

In 1983 a new association succeeded CAAL: it was named CAEL and housed at the *Aloisianum* of Gallarate[9]. Many Centres, from Israel to Czechoslovakia, Belgium, France and Germany have been inspired by Gallarate experience.

Later on, father Busa established the CIRCSE at the Sacred Heart's University of Milano, where he currently lectures Thomistic lexicography and lexicology and Computer linguistics[10]. He also gives courses of Text analysis by computer and Thomistic hermeneutics at the Gregorian Pontificial University in Roma.

> *This writing dates back to 1991 and could not capture the manifold enterprises to which father Roberto devoted the last and fruitful period of his life. Beside a series of lectures at the Polytechnic of Milano about philosophy and psychology in Artificial Intelligence and Robotics, let me at least mention the weighty project he named Disciplined Languages, aimed at automatic translation, a recurring concern of AI.*
>
> *In this project he likely resumed seminal ideas he conceived during early discussions with the cybernetics Eduardo Caianiello and Norbert Wiener and, more specifically, while participating in research, sponsored by EURATOM, for automatic translation of Russian scientific abstracts. The research was abandoned due to the unfair conclusions of the 1966 ALPAC report, after which the same US government discontinued funding Machine Translation[11].*

---

8 CAAL: Centro Automazione Analisi Linguistica.

9 CAEL: Computerizzazione delle Analisi Ermeneutiche e Lessicografiche.

10 CIRCSE: Centro Interdisciplinare per le Ricerche della Computerizzazione dei Segni dell'Espressione.

11 ALPAC: Automatic Language Processing Advisory Committee.

5. Sources

1 Busa R.: "Cinquant'anni a bitizzar parole"; in Convegno internazionale di storia e preistoria del calcolo automatico e dell'informatica – Siena, 10-11-12 settembre 1991; AICA, Milano, 1991, pp. 71-82.

2 Busa R.: "The Annals of Humanities Computing: the Index Thomisticus"; *Computers and the Humanities*, Vol. 14, n. 2, October 1980, pp. 83-90.

3 Busa R.: "*Report to my Superiors – March 3rd, 2008*"; not published, friendly copied to author.

4 Busa R.: *Author's Profile*; annexed to [1].

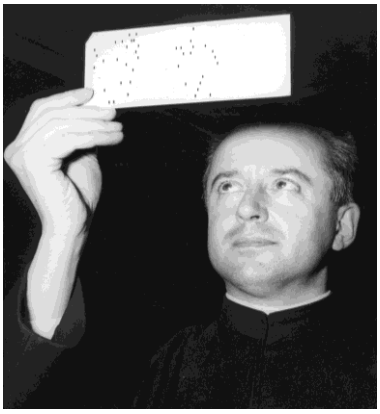

Figure 1 - Roberto Busa while "ostending" a punched card. (About 1965). Courtesy of CAEL

Figure 2 - Together with Norbert Wiener, at CAAL. (1958. Courtesy of CAEL)

Figure 3 - Roberto Busa decorated as Knight Gran Cross. (Roma, 2005. *CAEL Newsletter*, Dec. 2006)

Figure 4 - Father Roberto annually issued a CAEL Newsletter bringing, among other, a parody of himself sketched by artist Marina Molino. This one appeared *in memoriam*, featuring him finally reaching, in heaven, his beloved mentor Thomas Aquinas. (2011. Courtesy of CIRCSE)