

ANNALI DI STORIA DELLE UNIVERSITÀ ITALIANE

Comitato di direzione: Gian Paolo Brizzi, Antonello Mattone, Andrea Romano.

Comitato di redazione: Elena Brambilla (Università di Milano), Marco Cavina (Università di Bologna), Romano Paolo Coppini (Università di Pisa), Piero Del Negro (Università di Padova), Peter Denley (Queen Mary University, London), Mordechai Feingold (California Institute of Technology, Pasadena), Roberto Greci (Università di Parma), Paul F. Grendler (University of Toronto), Daniele Menozzi (Scuola Normale Superiore di Pisa), Mauro Moretti (Università per Stranieri di Siena), Paolo Nardi (Università di Siena), Luigi Pepe (Università di Ferrara), Mariano Peset (Universidad de Valencia), Maria Gigliola di Renzo Villata (Università di Milano), Hilde de Ridder Symoens (Universiteit Gent), Marina Roggero (Università di Torino), Roberto Sani (Università di Macerata), Elisa Signori (Università di Pavia), Andrea Silvestri (Politecnico di Milano), Maria Rosa di Simone (Università di Roma "Tor Vergata"), Gert Schubring (Universität Bielefeld), Jacques Verger (Université Paris Sorbonne-Paris IV).

Comitato dei consulenti editoriali: Girolamo Arnaldi (Emerito, Università di Roma "La Sapienza"), Patrizia Castelli (Università di Ferrara), Maria Luisa Chirico (Seconda Università di Napoli), Rosanna Cioffi (Seconda Università di Napoli), Ester De Fort (Università di Torino), Gianfranco Fioravanti (Università di Pisa), Giuseppina Fois (Università di Sassari), Gianfranco Liberati (Università di Bari), Angelo Massafra (Università di Bari), Aldo Mazzacane (Università di Napoli "Federico II"), Paolo Mazzarello (Università di Pavia), Simona Negruzzo (Università Cattolica del Sacro Cuore - Brescia), Maria Grazia Nico (Università di Perugia), Daniela Novarese (Università di Messina), Giuliano Pancaldi (Università di Bologna), Marco Paolino (Università della Tuscia - Viterbo), Lorenzo Paolini (Università di Bologna), Maurizio Ridolfi (Università della Tuscia - Viterbo), Achille Marzio Romani (Università Commerciale "Luigi Bocconi"), Maurizio Sangalli (Università per Stranieri di Siena), Ornella Selvafolta (Politecnico di Milano), Andrea Tabarroni (Università di Udine), Elio Tavilla (Università di Modena e Reggio Emilia), Andrea Tilatti (Università di Udine), Francesco Totaro (Università di Macerata), Francesco Traniello (Università di Torino), Ferdinando Treggiari (Università di Perugia), Gian Maria Varanini (Università di Verona).

Gli «Annali di storia delle università italiane» sono una pubblicazione periodica a cadenza annuale. Gli «Annali» si propongono come punto di incontro, di discussione e di informazione per quanti, pur nella diversità degli approcci storiografici e nella molteplicità dei settori disciplinari di appartenenza, si occupano di temi relativi alla storia delle università italiane.

La rivista è espressione del "Centro Interuniversitario per la Storia delle Università Italiane" (CISUI), cui aderiscono attualmente gli atenei di Bari, Bologna, Ferrara, Macerata, Messina, Milano "Luigi Bocconi", Milano Politecnico, Milano Statale, Modena e Reggio Emilia, Padova, Parma, Pavia, Perugia, Pisa, Roma "Tor Vergata", Sassari, Scuola Normale Superiore di Pisa, Siena "Università per Stranieri", Seconda Università di Napoli, Torino, Valle d'Aosta, Verona, della Tuscia (Viterbo).

Redazione: Ilaria Maggiulli, Maria Grazia Suriano

Direttore responsabile: Gian Paolo Brizzi

Autorizzazione del Tribunale Civile di Bologna n. 6815 del 5/6/98

I testi pubblicati sono preventivamente valutati dai curatori indicati, per ciascun numero, dal Comitato di redazione e dal Comitato dei consulenti editoriali. I testi sono altresì sottoposti al doppio giudizio in forma anonima di esperti interni ed esterni (double-blind peer review). Il modulo per la peer review è disponibile on-line all'indirizzo www.cisui.unibo.it/home.htm. Gli articoli pubblicati in questa rivista sono catalogati negli indici sotto elencati.

«Annali di storia delle università italiane» uses a double-blind peer review system, which means that manuscript author(s) do not know who the reviewers are, and the reviewers do not know the names of the author(s). It is covered by the following abstracting/indexing services:

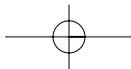
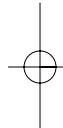
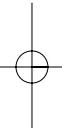
Acnp - Catalogo italiano dei periodici
Aida - Articoli italiani di periodici accademici
Bibliografia storica italiana
EBSCO Publishing - Historical Abstract
EIO - Editoria italiana online

Il CISUI ha la propria sede presso l'Università di Bologna:
Centro interuniversitario per la storia delle università italiane
Via Galliera 3
40121 Bologna
tel. +39+051224113; fax +39+0512088507
e-mail: cisui.redazione@unibo.it; indirizzo internet: www.cisui.unibo.it/
Corrispondenza redazionale: «Annali di storia delle università italiane», CP 82, 40134 Bologna 22
Abbonamenti e acquisti: CLUEB, via Marsala 31, 40126 Bologna

Copyright: tutti i diritti sono riservati. È vietata la riproduzione, anche parziale, con qualsiasi mezzo effettuata, compresa la fotocopia, anche ad uso interno o didattico, non espressamente autorizzata dalla Redazione della rivista.
© 2013 CLUEB, via Marsala 31, 40126 Bologna e CISUI, via Galliera 3, 40121 Bologna

Annali di storia delle università italiane





INDICE

9 STUDI

- 11 ANGELO MASSAFRA, *Mezzo secolo di storia dell'Università di Bari: note introduttive*

L'Università di Bari fra Otto e Novecento: politica, società e cultura

- 31 ELISA SIGNORI, *L'istituzione dell'Ateneo di Bari e la politica universitaria italiana del primo dopoguerra. Da Croce a Gentile*
- 45 TOMMASO DELL'ERA, *Strategie politiche ed esigenze scientifiche: il ruolo di Nicola Pende nell'istituzione e nell'organizzazione dell'Università di Bari*
- 69 LUIGI MASELLA, *L'Università, Bari e la Puglia: cultura, società e politica nel Novecento*
- 81 FRANCESCO PAOLO DE CEGLIA, *L'Università di Bari e le tradizioni scientifiche locali e regionali*
- 113 DORIANA DE TOMMASI-GIAMBATTISTA DE TOMMASI, *L'edilizia universitaria tra emergenza e progetto (anni '20-'70)*
- 135 MAURO DI GIANDOMENICO-LUCIA DE FRENZA, *Organizzare la ricerca nell'Università di Bari: uomini, strumenti, risorse*
- 147 BENEDETTA CAMPANILE, *Il materiale storico-scientifico dell'Università di Bari: i musei tra scienza, didattica e conservazione*
- 161 ANNA ORFINO, *L'Archivio generale di Ateneo dell'Università di Bari tra recupero, valorizzazione e ricerca storica*

Ambiti e protagonisti della didattica e della ricerca di un Ateneo in formazione: primi studi

- 175 LUIGI VOLPE, *Gli studi giuridici e la teoria del diritto: protagonisti e linee di sviluppo dalla fondazione della Facoltà di Giurisprudenza fino agli anni '60*
- 207 ORNELLA BIANCHI, *Dalla Scuola di studi corporativi alla Facoltà di Scienze Politiche*
- 229 EZIO RITROVATO, *Tra la Scuola Superiore di Commercio e la Facoltà di Economia: precedenti storici e sviluppi fino agli anni '70 del Novecento*
- 241 CLAUDIO ACCIANI-FRANCESCO ALTAMURA, *Le Scienze agrarie fra sviluppo, formazione e ricerca scientifica*
- 259 FRANCESCO ALTAMURA, *Vincenzo Ricchioni fra scienza, politica e governo dell'Università*
- 273 BRUNO GHIDINI, *Michelangelo Merlin e gli studi di Fisica nell'Università di Bari*
- 297 VITILIO MASIELLO-RAFFAELE RUGGIERO, *Mario Sansone e gli studi di Italianistica*
- 305 MARIO PANI, *Gli studi di Antichistica nella Facoltà di Lettere e Filosofia*

323 FONTI

- 325 STANISŁAW A. SROKA, Academic degrees obtained by Poles studying at Northern Italian universities in the second half of the 15th c. (Bologna, Padua, Ferrara)
- 333 SILVIA CONTI, La fisica sperimentale nell'istruzione sabauda del Settecento: ricerca dell'«uniformità» e rinnovamento dei saperi
- 355 GIAN LUIGI BRUZZONE, Francesco Selmi e Stanislao Cannizzaro
- 377 MARIA TERESA MARCIALIS, La “restituita” Facoltà di Lettere e Filosofia. La filosofia a Cagliari dal 1924 al secondo dopoguerra e oltre
- 399 GIOVANNI RITA, Le «Scienze sagre» nella Sapienza romana tra giansenismo e cattolicesimo reazionario. Itinerari bio-bibliografici
- 433 NATHALIE GOROCHOV, Les relations entre les *studia* de Paris et de Bologne et la naissance des premières universités d'Europe (XII^e siècle-début XIII^e siècle)
- 447 MARCO BARBIERI, Per una prosopografia dei laureati alla Facoltà teologica pavese nell'età delle riforme asburgiche. Stato dell'arte e prospettive di ricerca
- 455 MATTIA FLAMIGNI, Il processo epurativo all'Università di Bologna

475 ARCHIVI, BIBLIOTECHE, MUSEI

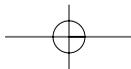
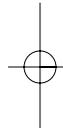
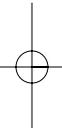
- 477 CARLO SARTI, The Capellini Museum (University of Bologna), the most ancient Italian Geo-Paleontological Museum
- 487 ALESSANDRA BARETTA-MARIA PIERA MILANI, Il *Fondo docenti* dell'Archivio storico dell'Università degli Studi di Pavia: i risultati di un progetto di recupero e valorizzazione
- 495 NICOLETTA TROTTA, Il Fondo Manoscritti dell'Università di Pavia, «scricino della memoria»

505 SCHEDE E BIBLIOGRAFIA

- 507 *Almum Studium Papiense. Storia dell'Università di Pavia*, I, 1, *Dalle origini all'età spagnola*, a cura di DARIO MANTOVANI, Milano, Cisalpino, 2012 (ROBERTO GRECI), p. 507; *Autographa*, I, 1 *Giuristi, giudici e notai (sec. XII-XVI med.)*, a cura di GIOVANNA MURANO, con la collaborazione di GIOVANNA MORELLI, Bologna, CLUEB (Studi, 16), 2012 (ILARIA MAGGIULLI), p. 508; ELISABETTA BARILE, *Per la biografia dell'umanista Giovanni Marcanova*, Treviso, Antilia, 2011 (MARIA TERESA GUERRINI), p. 509; MASSIMO BUCCIANINI-MICHELE CAMEROTA-FRANCO GIUDICE, *Il telescopio di Galileo. Una storia europea*, Torino, Einaudi, 2012 (PAOLO MAZZARELLO), p. 509; *Collegiate learning in the middle ages and beyond. 2. Coimbra group birthday seminar*, ed. by ANTONIO SAVINI, Milano, Cisalpino, 2012 (MAURIZIO PISERI), p. 510; *Formare alle professioni. Architetti, ingegneri, artisti (secoli XV-XIX)*, a cura di ALESSANDRA FERRARESI-MONICA VISIOLI, Milano, FrancoAngeli, 2012 (MARIA TERESA GUERRINI), p. 511; *La formazione del medico in età moderna (sec. XVI-XVIII). Atti della XXXVIII tornata degli studi storici dell'arte medica e della scienza, Fermo, 20-22 maggio 2010*, a cura di ROBERTO SANI-FABIOLA ZURLINI, Macerata, EUM, 2012 (ARIANE DRÖSCHER), p. 512; DANIELA FRANCHETTI, *La scuola di ostetrica pavese tra Otto e Novecento*, presentazione di PAOLO MAZZARELLO, Milano, Cisalpino, 2012 (CLAUDIA PANCINO), p. 514; *Galileo e la scuola galileiana nelle Università del Seicento*, a cura di LUIGI PEPE, Bologna, CLUEB (Studi, 14), 2011 (DIEGO DONNA), p. 515; *Laura Bassi. Emblema e primato nella scienza del Settecento*, a cura di LUISA CIFARELLI-RAFFAELLA SIMILI, Bologna, Editrice Compositori, 2012 (MARIA TERESA GUERRINI), p. 516; ANTONIO LOMBARDINI, *Diario universitario (1825-1835)*, a cura di SERGIO DI NOTO MARRELLA, Parma, Casa editrice Alessandro Farnese, 2013 (GIAN PAOLO BRIZZI), p. 517; *Mathematicians in Bologna 1861-1960*, ed. SALVATORE COEN, Basel, Birkhauser, 2012 (LUIGI PEPE), p. 518; CARLOS NIETO SÁNCHEZ, *San Clemente de Bolonia (1788-1889): el fin del Antiguo Régimen en el último colegio mayor español*, Madrid, publicaciones de la Universidad Carlos III, 2012 (CARLOS DEL CASTILLO RODRÍGUEZ), p. 519; *L'organizzazione dei saperi all'Università di Pisa. Dalle Facoltà ai nuovi Dipartimenti*, Pisa, Pisa University Press, 2012 (LUIGIAURELIO POMANTE), p. 521; GIUSEPPE PALMISCIANO, *L'Università di Napoli nell'età della Restaurazione. Tra amalgama, moti e repressione*, Bologna, Il Mulino, 2012 (ANNA MARIA RAO), p. 522; *Per la storia dello Studio di Siena. Documenti dal 1476 al 1500*, [a cura di] TIZIANA FERRERI, Milano, Monduzzi Editoriale, 2012 (GIAN PAOLO BRIZZI), p. 525; *Il Politecnico di Milano e la formazione delle classi dirigenti nazionali e locali*, Cinisello Balsamo, Silvana editoriale, 2013 (GIAN PAOLO BRIZZI), p. 525; *I Pronostici di Domenico Maria Novara*, a cura di FABRIZIO BÒNOLI ET AL., Firenze, Olschki, 2012 (ELIDE CASALI), p. 526; MIRELLA SPADAFORA, *Felicem peragrat Italiam. Viaggio di istruzione in Italia di Veit Königl giovane barone del Tirolo del Sud (1607-1609; 1609-1611). Libro delle spese di viaggio*, Bologna, CLUEB, 2012 (ARIANE DRÖSCHER), p. 528; *Gli studenti dell'Università di Padova caduti per l'Unità d'Italia. Documenti raccolti dalla Biblioteca del Consiglio Regionale del Veneto in occasione della cele-*

brazione solenne a Palazzo del Bo del 17 marzo 2011 Festa nazionale per il 150° dell'Unità d'Italia, Venezia, Consiglio Regionale del Veneto, 2011 (MARIA TERESA GUERRINI), p. 529; *L'Università di Macerata nell'Italia unita (1861-1966). Un secolo di storia dell'Ateneo maceratese attraverso le relazioni inaugurali dei rettori e altre fonti archivistiche e a stampa*, a cura di LUIGIAURELIO POMANTE, Macerata, Eum, 2012 (MARIA PIA CASALENA), p. 529; *Le Università e l'Unità d'Italia (1848-1870)*, a cura di ALESSANDRA FERRARESI-ELISA SIGNORI, Bologna, CLUEB (Studi, 17), 2012 (FRANCESCA SOFIA), p. 531; *Le Università erano vulcani... Studenti e professori di Pavia nel Risorgimento. Mostra documentaria, Università di Pavia, 9 febbraio-30 marzo 2011*, a cura di MARINA TESORO-ARIANNA ARISI ROTA, Pavia, 2011 (MARIA TERESA GUERRINI), p. 531; GEERT VANPAEMEL-MARK DEREZ-JO TOLLEBEEK, *Album van een wetenschappelijke wereld: de Leuvense universiteit omstreeks 1900*, Leuven, Lipsius, 2012 (ILARIA PORCIANI), p. 532.

- 535 Bibliografia corrente e retrospettiva
- 545 NOTIZIARIO
- 547 Convegni, seminari, incontri di studio
- 568 Attività e progetti
- 580 Tesi
- 584 Riviste e notiziari di storia delle università



Notiziario

I primi risultati di questo lavoro si sono espressi in particolare nella citata pubblicazione *Il Politecnico di Milano e la formazione della classe dirigente nazionale e locale*, dove, per evidenziare in modo esemplificativo l'apporto del Politecnico di Milano alla costruzione della classe dirigente nazionale si è fatto ricorso a cento percorsi individuali di varie personalità appartenute alla comunità del Politecnico di Milano, in quanto studenti e/o docenti. I cento profili biografici di cui si dà conto sono stati selezionati all'interno del database tra personalità non più viventi, la cui attività ha avuto ricadute significative in ambito professionale o si è accompagnata all'esercizio di ruoli politico-amministrativi. Tra questi cento personaggi, alcuni approfondimenti sono stati dedicati a figure di cui è sembrata esemplare la trasversalità sociale e istituzionale, ma che al contempo erano prive di una sufficiente illustrazione attraverso monografie dedicate. In altri termini si è ritenuto che proprio questi soggetti intermedi potessero evidenziare il carattere diffuso del contributo fornito dal Politecnico di Milano alla costruzione delle classi dirigenti nazionali.

Come seguito della ricerca, ci si propone di dare visibilità a questo lavoro tramite una pubblicazione on-line, implementabile e consultabile all'interno di un semplice pannello di navigazione, che possa evidenziare la fitta rete di relazioni che è in parte l'ordito su cui è stato costruito il nostro Paese.

STEFANO MOROSINI
ANDREA SILVESTRI
FABRIZIO TRISOGLIO

PROSO: prosopographic records. A model for the university students database ASFE

1. Introduction

The data needed to carry out prosopographic-based research is at the moment hard to use, collect and share.

These problems might be mitigated substantially by the use of modern technology, but these results can be obtained only if said technology is based on sound principles that are shared among both halves of the digital humanities sphere: humanists and information science specialists. Prosopographic data are inherently hard to collect. Most of the sources with the highest availability are usually written records stored on paper. The digitization of these records has a high cost, not just financially, but in time and effort required. To make thing worse, many sources or collections have been digitalized more than once, once for every project that needed to use them. Sometimes even activities of 're-digitization' are needed, because the collections have been published only on paper and not on electronic means. In many other cases, there exist an electronic copy of the data, but unfortunately it is often either in a format that cannot be understood or easily accessed, perhaps because it relies on technologies no longer available (e.g. old proprietary database applications), or it is modeled in a way fit to suit only a single project specific needs, thus making it very hard to plan information sharing and interoperability, for a number of reasons, including different assignment of meaning to unrestricted vocabularies, different choices in data modeling, and so on. Indeed, another justification often motivating re-digitization acts is the fact that a certain research group wants to describe the data using a conceptual model different from that used by the original authors.

The need to share data is sometimes felt as a non-issue by researchers because they see themselves as the only scholars that can be interested in a certain niche data (for example, how many research groups are working on scholars of noble origins that studied in universities of southern Italy in the 17th century?). The effort to make this data more easy to share and reuse might then be seen as a fruitless burden. However, although this concern may seem plausible while looking at a single project, when the scope rises to a more panoramic oversight of the targets of multiple projects

in the field of historical research, it is easy to realize that many of them involve overlapping or related data/information about several 'neighbouring' others. If the data coming from all these projects could be put together, it is a reasonable expectation that new insights could more easily emerge, similarly to what is happening with the many Linked Data repositories available on the web [10].

A first step towards a future where sharing prosopographic data will be easy and fruitful is the definition of a common format for such data. The word 'format' entails more than one thinks. Defining a common format means that, first, there must be a common data model agreed upon by the producers of the data and its users. The role of such a model is to define what are the entities that are discussed, what meaning and ideas are associated to them and how they are composed and linked together.

A second aspect of the format is the definition of a serialization format to use to write down the data from the abstract model in a file. This serialization format must define in precise terms how the data is written in terms of technical elements (for example XML elements, columns of a CSV file or RDF statements). In order for such a format to be successfully employed, it needs to address both the needs of scholarly researchers (in this case: historians) and of tool developers. Consequently, researchers need to ensure that the format is able express all the information they care about and that the way the information is expressed is sound, clear and not ambiguous. The tools developers must instead assert that said format is on par with the current best technologies, and is at the same time both easy to manage, scalable and does not require excessively complex tools to deal with.

We developed PROSO as our proposal of a format for the storage and exchange of prosopographic data. The design of PROSO has been driven by investigating the data that is already used in various projects [1, 2, 5, 6, 7, 9], as well as other models for prosopographic data [1, 8]. PROSO tries to address the needs of scholars as well

Notiziario

as those of developers; it balances easiness of use (a must for tool developers) with expressive power and the ability to extend the format to describe new or peculiar entities (characteristics needed by researchers in the field, to make sure that any data can be expressed in PROSO).

The rest of the paper is structured as follows. Section 2 reviews some of the existing prosopography database, the kind of data they deal with and their own data models and formats. The PROSO model is described then in Section 3, its XML format is briefly explained in Section 4 and its possible adaptation for use in the Linked Data world in Section 5. We conclude in Section 6 with a summary of advantages of PROSO and ideas for future developments.

2. Current projects and their data

Computer databases in prosopography have a long history of being appreciated as a fundamental research aide by scholars in the field. Indeed, the classification of knowledge extracted from heterogeneous sources towards a more structured format is a common goal. After all, prosopography «collects and exploits structured biographical data» [15]. In a sense, prosopographical records can also be said to have an inherent structure very much alike to the one most common in most relational databases, especially if we are to consider the importance that interactions between individual recorded in the prosopography. In some cases, even the primary sources might be characterized by a certain regularity of the records.

As a natural and logical consequence of the importance of investigating possible relationships between two or more different prosopographical records, as well as the need to quickly access, browse and query these knowledge archival systems, many projects have implemented web-based database systems, usually providing the public (or a selected group of researchers) with a web interface for querying the database and displaying the detailed information returned

by the system. However, given the massive scale of human history, many of these projects often focus only on a specific subset of prosopographical information, limiting their scope in time, space or with some other categories.

Nevertheless, a researcher might want to compare the outputs of different prosopographical databases, perhaps some sharing some common property (e.g. time period) but separated by another (e.g. geographical area). At the time being, even simple unified search services between similar web-based prosopographical databases are only far-sighted visions and other applications for complete interoperability even more chimerical. This is especially true considering the specific subset of prosopographical studies we had focused our activity on, namely the one of scholarly records from the late middle ages and early modern age.

Yet the well-known advantages of interoperability are not lost in the reference research community of historians: on the contrary, there has been a consistent strong push for the development of at least the basic foundations of it, in order to encourage cooperative efforts and enable a more effective and cost-efficient use of the results of scholarly efforts. This has been felt for over a decade. Indeed, since the early FASTI Workshop held in Amsterdam in 2001, «A consensus was reached about the necessity for a uniform and easy database structure for the study of university populations» [3]. In the successive FASTI workshop, held in Ghent in 2003, this inceptive request for interoperability retained its importance together with the need of widening the availability of used sources as well the perceived need for a common standard.

Right now, the leading cooperative effort in the field is Heloïse [4], a European network of historical databases for the people involved in ancient higher learning institutions. The project was established in May 2012 to promote scientific meetings and encourage the implementation of technical solutions for collaboration. The network coordinates also a mailing list, and the network participants and their

respective databases were the targets of our investigations on the interoperability and metadata needs and *desiderata* of this research community.

Aiming to research and propose a suitable exchange format to enable communication and eventual interoperability between these research projects, we did not confine our analysis of the state of the art to just the basic assessment of the available project, but we decided to contact the associate partners of the Heloïse network with a very short written questionnaire. The goal of this interview was to facilitate us to reach the aforementioned objective. We deem important that such a common schema of metadata should have the maximum possible consensus in the designated community of researchers. Thus the questionnaire aims to be a first step down that path, in order to highlight those elements and themes of shared relevance, confirming or refuting what we gathered from our investigations or the FASTI reports.

The 7 submitted questions aimed to assess the needs and the scope of each project, as well as integrating the information publicly available. The questions were worded as following:

1. Would you kindly summarize the scope and focus of your project? (study subjects, language and years of coverage)?
2. What are the main conceptual entities (e.g. people, roles, locations) and the relations between them (e.g. teacher-student relations, data about associations) that you manage? For example, some of the projects might be more interested in a certain class of people sharing common attributes (e.g. the career of all students from a certain area), while others more centered around other entities (for example, a specific location and all students passing through it).
3. Do you use controlled vocabularies, taxonomies, metadata schemas or ontologies in your projects? If so, which ones?
4. Is the data you manage normalized? Are any non-normalized form stored as well? Are the sources of the data (or of the normalization process) tracked?
5. Have you already considered the is-

Notiziario

sue of sharing your data with other projects in order to implement a common platform for search? If so, have you identified any critical priorities?

6. What kind of research results are you interested in publishing and sharing, either with the general public or to the other academic projects of the research community?

7. What type of searching capabilities do you suppose would be the most useful when envisioning a cooperative web environment (e.g. familiar relationship, geographical area of provenance, type of studies, ...)?

We collected answers from 7 of these projects: *Amore Scientiae Facti Sunt Exules* (ASFE) [1] of the University of Bologna; from the Universität Bern; *Repertorium Academicum Pictaviense* (RAP) [7] of the Université de Poitiers; SymoGIH [8] from the Laboratoire de Recherche Historique Rhône-Alpes (LARHRA); the 'Teachers of Arts and Medicine in Renaissance Italy' from the University of Warwick [9]; the *Onomasticon* of the University of Perugia [5]; and the *Catalogus Professorum Lipsiensium* (CPL) [2] of the Universität Leipzig.

We then proceeded to analyze the answers to assess if there existed any common ground between the respondents, and, if so, in what areas and about which concepts or information. From our analysis, the first thing that we could point out is that, at the current time, there is no available platform or method for any kind of interoperability or intercommunication between any of the projects, even if most respondents declared themselves interested, and some were already striving towards putting into practice some of the first steps for this.

We also discovered that there are just small intersections between the projects when considering their geographical or historical scope. There is almost no intersection of significance between all the seven of them, but the possible intersections grow when considering only pairs of project, especially from the point of view of the time period studied. This reinforced our idea of the field as a network aggregation of multiple 'neighbouring' scholarly efforts.

However, while at the moment there is little practical space and time concordance in the project's scope, there are very relevant affinities in the research subjects, both when considered as conceptual entities (e.g. people, especially scholars, universities, locations...) and on their related properties (e.g. simple biographical information about a person, or his/hers *curriculum studiorum*). Another shared need emphasized in the answers was the importance of source tracking and the possibility to express uncertain information (e.g. concerning dates) or variants of normalizations (e.g. names). Most importantly, all parties were interested in the possibility of sharing, querying and accessing relevant data to and from project partners (although some expressed concerns about keeping control of what to expose to others). We thus ascertained that existed a foundation for our effort in designing and developing a shared metadata model. This would need to suit the formal description of prosopographical records, satisfying the respondent's needs and their wishes to realize the idea of interoperability as a possibility to communicate information on topics of common interest, and to do so with a pre-established common vocabulary.

The way we decided to shape our proposal according to our findings will follow in Section 3, but before proceeding to it, we will briefly expound on the aims and the state of the art of some of the other projects participating in Heloise. We had the opportunity to work very closely with the scholars behind the ASFE database. This project consists in a web-based database offering three different thematic sets of prosopographic records, all covering the same time span (1500-1800). One is a prosopographical database of the students and scholars that have passed through the University of Bologna, regardless of their provenance or their final accomplishments, the second one is a prosopographical collection of all Italian graduates of the time, while the third one focus on students arriving in Italy from outside the Alps. The main emphasis of ASFE is on the people and their roles, and especially on their *curricula studiorum*,

that is the scholarly records of their accomplishments or positions. Little biographical information is usually given, aside from the name and geographical area of provenance.

RAG is one of the other big projects in this field: RAG's goal is to collect biographical and social data on those theologians, jurists, physicians and masters of arts who studied at a university in the Holy Roman Empire (HRE) between 1250 and 1550. Non-graduated noble visitors of universities are also taken into account. RAG aims in the end to be a «who is who of the scholars of the Old Empire», and strives to offer interdisciplinary perspectives from its vast collection of data, offering opportunity for both qualitative and quantitative research on the intellectual elite of the HRE, with special reference to social, cultural, and scientific history. Again, the main conceptual entity is a person, being a graduate in the HRE, described both by his/hers career, relations to other people, as well as more specific biographical data in comparison to ASFE (e.g. often including noble titles, career outside the university, information about birth and death). However, unlike ASFE, sources in RAG aren't tracked information per information, but on a whole per-record basis.

RAP is similar, although with a more narrow geographical scope. It aims to describe and record registered students of the Poitiers University, from the 1431 onwards. Their first goal is the digital transcription of grade records ('registres de grades') and registration records ('matricule'), with these tabular data stored in a simple database which can be accessed by the RAP web site. A second task is a prosopographical study of some of the more notable of these students, mainly during 17th century. People (students and, in some cases, professors), study subjects, grades and locations are the main entities involved in their research. Many of the primary sources are also attached to records as scanned images.

SymoGIH is a very interesting effort started in 2007 by Laboratoire de Recherche Historique Rhône-Alpes. It consists in a collaborative database enabling historians to store the informa-

Notiziario

tion produced by individual researchers or collective programs and to share it with others within this cumulative system. As it is, it is one of the projects, together with RAP, who is also geared towards interoperability: its data model is especially interesting for describing locations in historical time, as well as other entities (people, etc.), given the philosophy of separation between entity identifiers and the content associated with that identifier at a certain time.

Onomasticon's scope is to record news on professors and students attending the University of Perugia between 1300 and 1515. The web-based interface for the database allows to search for both biographical and bibliographical information. Aside from the usual data about people and study subjects, a good deal of emphasis is given to information of financial nature, such as payments and compensations earned by the professors and lecturers of the University, year by year.

CPL is an historical prosopographical database about professors of the University of Leipzig. It covers the ongoing lifetime of the University starting from 1409. Once again the main conceptual entities are people, as well as location and dates, as the catalog contains facts about historical individuals. All facts are mainly linked by the periods of life of the single professor. Periods of life are linked to bodies including organizations and institutions, as well to other people.

The University of Warwick has an ongoing project called 'Teachers of Arts and Medicine in Renaissance Italy' and it is a prosopographical database, not yet accessible to the public, about all the professors of Arts and Medicine in the Universities of Bologna, Pavia, Padova, Pisa and Rome for the period between 1350 and 1600. Again, aside from people, the main concepts handled are locations, teaching subjects, payment and locations.

Project Studium Parisiense, by the Université Paris-Sorbonne, is a database dedicated to members of the school and the University of Paris between 12th and 16th century. The records aim to ultimately store biographical and bibliographical information

of teachers, students and other people related to the University of Paris.

Summarizing what we could gather from the analysis of all these projects, they all deal with similar kinds of data. All of them describe people, their roles and contextualize most of their information with dates. The differences between all these projects do not lay in their models but in the data they care about; the data collected by these projects is often focused around small geographical areas or time periods. However, the information they deal with is very similar and the use of a shared reference model could foster the exchange of data between them.

3. The PROSO Model

The PROSO model is proposed as a common format to describe prosopographic data and metadata. It is structured around four main concepts:

Entities: the things that are described by the data, e.g. people, places;

Factoids: the assertions that contribute to the description of the entities, the carries of the data;

Collections: the set of factoids collected by a certain author about a certain entity;

Sources: the person, group or document from which a piece of data has been extracted.

Prosopographic collection based on PROSO are able to express all the data needed by the projects analyzed in section 2. In addition to just expressing the collected information, any data modeled using PROSO is assured to have appropriate properties making the data stored very rich. The regularity of PROSO also makes it easier to develop advanced tools. For example, PROSO data always has an explicit source (that may be unreliable or imprecise) and provenance data, features that make it possible to combine different datasets from different projects without polluting the results in case the data is of heterogeneous quality.

3.1 Entities

Entities, or conceptual entities, are main concepts described in the proso-

pographical database, for example a Person, a Studium or a Study subject. Entities are either the subject or the objects of the information provided by the data suppliers and they are described by the means of factoids. In the PROSO model every concept which can be at the center of some scholarly discourse is an Entity.

The PROSO model describes various kinds of entity: Person, Study subject, Place, etc. Entities of these types should follow the structure thought for them in the PROSO model. First of all, each entity must have an ID, in the form of an URI, so that it can be referred to by factoids in other entities and by other datasets. Also, each entity must be described only in terms of factoids. This requirement is fundamental in making sure that all the datasets are interoperable with each other and with various tools based on the factoids model.

In order to accommodate extensions and a progressive evolution of the model, PROSO allows for *ad-hoc* entities, i.e. entities whose type is not one of the well know types or that are structured in a different way. Such *ad-hoc* entities must still adhere to the general entity structure previously described: they must have an ID and be described only in terms of factoids. We think that allowing new *ad-hoc* entities to be introduced under the constraint of being described only by factoids strikes a good balance between extensibility, rigidity of the model and simplicity of implementation of the tools.

The IDs of entities require further discussion, as they play an important role in the dataset. In the PROSO model each entity has an URI so that it can be referenced by other entities. For example, a person entity may be linked to another person entity using its URI and a 'parenthood' factoid. In addition, forcing every entity to have its own ID greatly simplified the work of tool implementers, as they no longer need to take care of the case there is the need to link an entity that has no ID.

The decision of which URI should be used as ID is also very poignant. We suggest to give all entities an URI local to each project. In addition to this URI there should be a separate list of links

Notiziario

(usually called a linkbase) used to link these URIs to the URIs of similar entities in other projects or to perform HTTP redirects to canonical URIs (for example to DBPedia resources or other Linked Data sources).

3.2 Factoids

The use of factoids is inspired by Bradley and Short [11], the keystone of the PROSO model. The factoid technique uses statements such as «author ASFE states that source Alvisi affirms that Wolfgang Castner was a student from 1571 to 1573» instead of the usual «Wolfgang Castner was a student from 1571 to 1573». The difference is easy to spot, as factoids do not state any kind of assertion as true (or false), but convey the meaning that some source is making statements about a fact. In general, factoids are used to state a relation between two entities or between an entity and a value. This is reminiscent of RDF [17] object and data properties.

The factoid statement carries much more known knowledge than plain statements or RDF assertions. However, factoids do not provide more information about an entity than their plain counterpart. What factoids are able to provide, as a precious addition, is a structured context allowing for the same data to be used with more precision and versatility. For example, the fact that factoids provide source information allows the association of different levels of trust to different sources. Factoids also enable the introduction of contradictory or incomplete information in the dataset, without the risk of tainting any other information already in the dataset, empowering the user and the scholar alike. Factoids can be otherwise used to provide precise attribution to imported datasets, an important feature for legal and ethical reasons.

3.3 Collections

Collections are the way used to group together all the factoids about a certain entity and from a certain source. With collections it is easy to mix together factoids about the same entity coming from different sources. Collections also make it possible to state certain prop-

erties (e.g. the set of used sources) for all the factoid in that collection instead of repeating that information for all factoids.

3.4 Sources

PROSO deals with two kinds of sources: external sources and original sources (also referred to as authors). The distinction between these two kinds of sources is similar to that found in scholarship between primary and secondary sources. The similarity is, however, limited. In PROSO, external sources are all the sources that have been used by the author to extract the data stored in the factoids. This means that external sources could be original documents but also lists of facts compiled by other authors.

The idea behind the PROSO sources is that every author is required to explicitly state what source has been used to gather the data about a factoid. This allows for the construction of chains of sources and makes it able to understand what are the passages that the data have undergone between the original document from what has been taken to the current document where it is stated. A list of 'default' sources (fallback sources) can however be specified at record or collection level to allow for better retro-compatibility.

Sources are identified by a URI that acts as their ID. As with the entity IDs, these URIs as supposed univocally identify a source, but they are not meant to be unique, i.e. there could be different URIs that identify the same source. Just like for entities, we think each project should maintain its own set of URIs that it uses to identify sources and have a linkbase of matching between its sources URIs and equivalent URIs published by other projects.

4. PROSO in XML format

The most basic way to exchange PROSO data is through XML files created according to the PROSO schema. An example XML PROSO file is shown in the following excerpt.

```
<person xml:id="http://exchange.heloiseproject.eu/person/CastnerW01">
```

```
<info -collection
author="http://asfe.unibo.it"
xml:id="http://asfe.unibo.it/persona/
NG0306">
<fallbackSources>
<source
href="http://asfe.unibo.it/source/
Matschinegg-385"/>
<source
href="http://asfe.unibo.it/source/
ASFE"/>
</fallbackSources>

<name normalized="true">
<given-name>Wolfgang </given-name>
<surname>Castner</surname>
<source
href="http://asfe.unibo.it/source/
ASFE"/>
</name>
<name>
<given -name>Wolfgang </given -name>
<surname>Kastner</surname>
<source
href="http://asfe.unibo.it/source/
Sienal-123"/>
</name>
<name>
<value>Castner Wolfgang</value>
<source
href="http://asfe.unibo.it/source/
Belvisi-135v"/>
<source
href="http://asfe.unibo.it/source/
Belvisi-136r"/>
</name>

<location>
<place type="region">
<value>Tirrolensis Germanus</value>
</place>
<source
href="http://asfe.unibo.it/source/
Belvisi-136r"/>
</location >
<location >
<place
href="http://asfe.unibo.it/places/
Germanus"/>
<source
href="http://asfe.unibo.it/source/
Belvisi-135r"/>
</location>

<changeOfSocialRelation type=
"affiliation" class="nazione">
<studium>Padova</studium>
<uni>Iuristarum</uni>
<natio>Germanica</natio>
<moment>1571-04-18</moment>
<source useFallback="true"/>
</changeOfSocialRelation >

<factoid type="note">
<value>Probably received a scholarship
for 1 year.</value>
<source useFallback="true"/>
</factoid >
</info-collection >
</person >
```

The root of a PROSO XML file is the entity described by that document.

```
<person
xml:id="http://exchange.heloiseproject.eu/person/CastnerW01">
```

Notiziario

In this case the entity is a person whose ID is:

```
http://exchange.heloiseproject.eu/person/CastnerW01.
```

Then, all the factoids about this person are grouped in an *info-collection* element:

```
<info-collection author="http://asfe.unibo.it"
xml:id="http://asfe.unibo.it/persona/NG0306">
```

The *info-collection* element states the ID of the factoid collection and, more importantly, the ID of the author of this collection.

The first things present in the factoid collection is the list of *fallback sources* to be used when the sources used to describe a factoid cannot be stated with precision.

```
<fallbackSources >
<source
href="http://asfe.unibo.it/source/Matschinegg-385"/>
<source
href="http://asfe.unibo.it/source/ASFE"/>
</fallbackSources >
```

This mechanism has been included in PROSO in order to allow the conversion of prosopographic datasets where not all the data has been annotated with source is has been extracted from and only a coarse list of used sources is available.

Proceeding, the first factoids describing this person are *name* factoids.

```
<name normalized="true">
<given -name>Wolfgang </given -name>
<surname >Castner </surname >
<source href="http://asfe.unibo.it/source/ASFE"/>
</name>
<name>
<given -name>Wolfgangus </given -name>
<surname >Kastner </surname >
<source
href="http://asfe.unibo.it/source/Sienal-123"/>
</name>
<value>Castner Wolphgangus </value>
<source
href="http://asfe.unibo.it/source/Belvisi-135v"/>
<source
href="http://asfe.unibo.it/source/Belvisi-136r"/>
</name>
```

The first *name* factoid provides a normalized version of the person's name,

normalized by the author identified by the URI <http://asfe.unibo.it/source/ASFE>. The second *name* factoid tells us how the same person is referred to in the external source <http://asfe.unibo.it/source/Sienal-123>; the third how it is referred to in two other sources. Please note the difference between the content of the second and the third factoid. In the second factoid the information about the name is stored in a structured way: the first name and the surname have been identified and marked accordingly. Instead, in the third factoid, the author of the factoid collection provided only the raw value of that piece of information, probably because they believe that the source is not reliable enough to detect the structure of the information without forcing a precise interpretation. These name factoids shows that in PROSO it is possible to mix contrasting statement without interfering with existing data or polluting it.

The next factoids that describe this person are the *location* factoids.

```
<location>
<place type="region">
<value>Tirrolensis Germanus </value>
</place>
<source
href="http://asfe.unibo.it/source/Belvisi-136r"/>
</location>
<location>
<place
href="http://asfe.unibo.it/places/Germanus"/>
<source
href="http://asfe.unibo.it/source/Belvisi-135r"/>
</location >
```

The content of the first *location* factoid is not a structured string as in the case of name factoids, but another entity, an entity of type place, or more specifically, a type region. The PROSO model defines a standard way to deal with places, but in this case the authors of the collection decided to be more specific and used an *ad-hoc* type region. Tools that understand the region type will be able to understand this factoid better. At the same time all the other tools will still understand that this factoid refers to a place and process it accordingly, losing some information but not all.

The second *location* factoid shows another way to provide information about entities. In this case the factoid is stating that the place of this location is the entity that can be found at the URI <http://asfe.unibo.it/places/Germanus>. It is expected that that URI contain a machine-readable description of that resources, possibly expressed as a PROSO XML file.

The *changeOfSocialRelation* factoid uses an already seen feature (the use of the type attribute to describe an ad-hoc factoid) and shows a new feature, the concrete use of any fallback sources described at the beginning of the collection.

```
<changeOfSocialRelation type="affiliation" class="nazione">
<studium>Padova </studium>
<uni>Iuristarum </uni>
<natio>Germanica </natio>
<moment> 1571 -04 -18</moment>
<source useFallback="true"/>
</changeOfSocialRelation>
```

The *useFallback* attribute is used to state that the source of this information is not known with precision, thus it must come from any of the fallback sources. In this case the author is using the fallback mechanism because this information has been recorded in its database before the project started keeping track of the sources used in each factoid.

The last factoid is the *note* factoid, a factoid that is used to as a last resort to be able to specify any kind of data in a PROSO dataset.

```
<factoid type="note">
<value>Probably received a scholarship for 1 year.</value>
<source useFallback="true"/>
</factoid >
```

They are similar to the *name* factoids, but show other features of the PROSO model. The *note* factoid is intended for free text annotations and for all kinds of legacy data.

5. PROSO and linked data

The current best way to make data available on the Internet and part of the Semantic Web is to publish data follow-

Notiziario

ing the Linked Data principles [10]. Instead of using directly one of the many Semantic Web technologies (RDF, OWL) to publish PROSO data, we envision a set of transformation from PROSO XML files to RDF datasets. This design decision is grounded in the idea that current ontologies and vocabularies lack features that are fundamental to prosopographic data, first of all the ability to record contrasting statements on the same subject.

GRDDL [13] makes it possible to transform XML files in RDF datasets on the fly, without the need to store permanent RDF datasets. This allows the gradual evolution of these RDF datasets as the semantic technologies evolve, through the update of the GRDDL transformations, without the need to touch the original PROSO XML files. Our current transformation is based on various ontologies: FOAF for names and people entities [12], BIO for events [14] and Biography Light Ontology for other bibliographic data [16].

6. Conclusions

Right now most of the prosopographic data is only available in paper form and it is being slowly and costly digitalized. It is important to make sure that all these digitalization efforts, together with the new born-digital projects, are based on a solid and shared conceptual model, so that it will be easy to collect, use and share this new wealth of data. After getting in contact with many ongoing projects based on prosopographic data, we found that most of the research groups dealt with very similar data (at least from the model point of view) and faced similar problems.

In this paper we introduced PROSO, a conceptual model and XML format for prosopographic data. The conceptual model is aimed at scholars and gives them a set of modeling guidelines about how to structure their research data. The XML format is, instead, intended to appeal the more technically inclined members of research groups: it provides a small and extensible exchange format that can be easily implemented using current technologies. In addition to this, we also suggested ways to export the prosopographic data as Linked Data using some of the *de facto* standard ontologies of vocabularies. Looking forward, we hope that the PROSO model and format will be discussed and adopted by research groups around the world. We are confident that the current PROSO model is a solid base on top of which all the new entity types that may feel needed can be added.

References

- [1] ASFE. <http://asfe.unibo.it>.
- [2] Catalogus professorum Lipsiensis. <http://catalogus-professorum.org>.
- [3] Fasti. <http://www.fastionline.org>.
- [4] Heloise: European workshop on historical academic databases. <http://heloise.hypotheses.org/>.
- [5] Onomasticon. <http://old.unipg.it/Prosopografico/>.
- [6] Repertorium Academicum Germanicum. <http://www.rag-online.org>.
- [7] Repertorium Academicum Pictaviense. <http://repertorium.projets.univ-poitiers.fr/>.
- [8] Symogih. <http://symogih.org/>.
- [9] Teachers of arts and medicine in the Italian universities. <http://www2.warwick.ac.uk/fac/arts/italian/staff/lines/research/>.
- [10] TIM BERNERS-LEE, *Linked data*, 2006. <http://www.w3.org/DesignIssues/LinkedData>.
- [11] JOHN BRADLEY-HAROLD SHORT, *Texts into databases: The evolving field of new-style prosopography*, «Literary and Linguistic Computing: the journal of digital scholarship in the humanities», 19/5 (2005), p. 3-24.
- [12] DAN BRICKLEY-LIBBY MILLER, *FOAF Vocabulary Specification 0.98*, Tech. rep., Feb. 2010. <http://xmlns.com/foaf/spec/20100809.html>. Latest version available at <http://xmlns.com/foaf/spec/>.
- [13] DAN CONNOLLY, *Gleaning Resource Descriptions from Dialects of Languages (GRDDL)*. Recommendation, W3C, Sept. 2007. <http://www.w3.org/TR/2007/REC-grddl-20070911/>. Latest version available at <http://www.w3.org/TR/grddl/>.
- [14] IAN DAVIS-DAVID GALBRAITH, *BIO: A vocabulary for biographical information*, Tech. rep., 2011. <http://vocab.org/bio>.
- [15] KATHARINE KEATS-ROHAN, *Biography, identity and names: Understanding the pursuit of the individual in prosopography*, in *Prosopography Approaches and Applications. A Handbook*, Occasional Publication of the Unit for Prosopographical Research, Vol. 13, Oxford, 2007.
- [16] MICHELE R. RAMOS, *Biography light ontology: an open vocabulary for encoding biographic texts*, 2009. <http://metadata.berkeley.edu/BiographyLightOntology.pdf>.
- [17] RALPH R. SWICK-ORA LASSILA, *Resource Description Framework (RDF) Model and Syntax Specification*. Recommendation, W3C, Feb. 1999. <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222>. Latest version available at <http://www.w3.org/TR/REC-rdf-syntax>.

GIOELE BARABUCCI
(CRR-MM Università di Bologna)
barabucc@cs.unibo.it

JACOPO ZINGONI
(CRR-MM Università di Bologna)
jacopo.zingoni2@unibo.it