Linking Manuscripts
from the Coptic, Ethiopian, and Syriac Domain:
Present and Future Synergy Strategies:

Preface to the Special Issue

Alessandro Bausi, Paola Buzi, Pietro Liuzzo,
and Eugenia Sokolinski

On 23 and 24 February 2018, a two-day workshop took place at Universität Hamburg, dedicated to 'Linking Manuscripts from the Coptic, Ethiopian, and Syriac Domain: Present and Future Synergy Strategies'.

Several projects based at the Hiob Ludolf Centre for Ethiopian Studies at Hamburg were behind the workshop. These included the Union of the German Academies-sponsored *Beta maṣāḥǝft: Manuscripts of Ethiopia and Eritrea* (see the project note, pp. 13–27) and the ERC-funded *TraCES*: *From Translation to Creation*: *Changes in Ethiopic Style and Lexicon from Late Antiquity to the Middle Ages* (see the project note, pp. 59–66), both chaired by Alessandro Bausi. The Landesforschungsförderung Hamburg provided significant funds for academic exchange between Ethiopianists in Hamburg and Copenhagen within the framework of the project *Transmission of Knowledge in the Red Sea Area*. The initiative was co-sponsored by the ERC-funded project *PAThs: Tracking Papyrus and Parchment Paths: An Archaeological Atlas of Coptic Literature. Literary Texts in their Geographical Context. Production, Copying, Usage, Dissemination and Storage* based at Sapienza Università di Roma and headed by Paola Buzi (see the project note by Paola Buzi, Julian Bogdani, and Francesco Berno, pp. 39–58). Cooperating projects included the projects *CMCL*: *Corpus dei Manoscritti Copti Letterari* (Rome / Hamburg), *Syriaca. org* (Vanderbilt University, Nashville), *IslHornAfr: Islam in the Horn of Africa: A Comparative Literary Approach* (Copenhagen), and *EMA: Ethiopian Manuscripts Archives* (Paris). Local organization was in the hands of Pietro Liuzzo and Eugenia Sokolinski.

The aim of the workshop was to have an informal exchange of practices accepted by the different research initiatives working with manuscripts from similar backgrounds and having a strong digital dimension. The participants had the chance to compare the challenges faced and the outcomes achieved or expected. The discussion focused on such points as (1) interactions between projects of digitization of catalogues of manuscripts from the Christian Orient, (2) alignment of authority lists for clavis identifiers, ancient places,

and ancient people, (3) standards for the reuse of primary canonical texts, (4) exploitation of common metadata standards for further outputs, and finally (5) future development perspectives for digital resources in the field.

The workshop was organized in three sessions, dedicated, respectively, to dealing with places, literary works, and manuscripts.

The first session (chaired by Paola Buzi) focused on the various strategies employed by the different projects in identifying and encoding information on (primarily historical) places in the respective research environments. Solomon Gebreyes Beyene illustrated how the *Beta maṣāḥǝft* project deals with encoding and annotating places on the example of the digital edition of an Ethiopic historical text, the *Chronicle of Galāwdewos* (see the contribution by Solomon Gebreyes Beyene and Pietro Liuzzo in this issue, pp. 121–141). Among other things, he illustrated how TEI-XML standards, and adaptation of existing ontologies, can help interoperability and visibility of data. In their paper on 'Texts and contexts: an effort to link Coptic literary texts to their archaeological context', Paola Buzi and Julian Bogdani showed, among other things, the approach adopted by the project *PAThs* to classifying and describing places relevant for the Coptic literary tradition (see the *PAThs* project note, pp. 39–58, §§1–2). The practices adopted by the *Syriaca.org* project for their Gazetteer (which, among other things, was also a source of inspiration for the *Beta maṣāḥǝft*) were presented (via Skype) by David Allen Michelson, who underlined the importance of assigning stable Universal Resource Identifiers (URI) to provide a digital structure for linking, data, aggregation and search functionalities. A special attention has been paid to the geographical fuzziness of historical places, a point the participants returned to during the discussion. The places (and other entities) recorded by the *Islam in the Horn of Africa* project were at the centre of the talk of Alessandro Gori (see the *IslHornAfr* project note, pp. 29–32).

The second session (chaired by Alessandro Bausi) focused on the Literary Works. The very first presentation, 'Identification of *one* work' by Tito Orlandi, formulated the main challenges scholars have to deal with when building up repertories, or claves, of works in a certain (in this case, Coptic) tradition (see the contribution by Tito Orlandi in this issue, pp. 107–114). Massimo Villa in his talk 'Encoding the Ethiopic literary heritage: issues and case studies' showed how this and other problems are being dealt with by the *Beta maṣāḥǝft* project (see the contribution by Massimo Villa in this issue, pp. 143–149). David Allen Michelson and Nathan P. Gibson (via Skype) spoke of the 'New Handbook of Syriac Literature (Syriaca.org/nhsl)' and of the 'Guide to Syriac Authors', other two core components of the *Syriaca.org* portal with data available in TEI and RDF which use stable URIs. They stressed the im-

portance of alignment to other authorities like Virtual International Authority File (VIAF). Paola Buzi and Francesco Berno returned to the Coptic tradition in their paper 'Coptic literature: authors, works, collections' (see the *PAThs* project note, pp. 39–58, §3).

The complexity of a relational database developed for the description of Islamic written production from the Horn of Africa in the *IslHornAfr* project was addressed in the presentation by Sara Fani 'Describing the Complex: the Multiple Dimensions of a Relational Database' (see this issue, pp. 89–96). In his talk 'Relation Labeling: The Case of Islamic Manuscripts', Michele Petrone focused on the solutions offered by the *IslHornAfr* project to managing the data elicited from endowment notes (*waqf*) on the example of an additional note in a manuscript from the collection of *šayḫ* Kamāl from Agaro (Gurage zone, Ethiopia). The *waqf* lists a number of entities assigned specific, fixed roles: founder, object of endowment, beneficiary, and controller. The approach chosen implies a full transcription of the *waqf*, related in the database to the specific codicological unit; the owners, founders, and beneficiaries (all being distinct entities in the database) are related to the paratext.

Finally, possibilities of linguistic study of digitally enoded texts were explored in the presentations on 'Multilayered digital annotation of Ethiopic texts' by Susanne Hummel, Vitagrazia Pisani, and Cristina Vertan (see their contribution in this issue, pp. 97–106) and on 'The digital Dillmann and a corpus-based lexicon' by Wolfgang Dickhut and Andreas Ellwardt (see this issue pp. 79–88).

The third session (chaired by Angela Bernardo) was dedicated to (digital) scholarship of manuscripts. Tito Orlandi opened the session by addressing the questions of how we define what is a manuscript, considering its changing and living nature, in his paper 'Identification of *one* manuscript' (see the contribution by Tito Orlandi in this issue, pp. 107–114). Dorothea Reule and Denis Nosnitsin offered some practical solutions to the various challenges posed by electronic cataloguing in the talk 'Encoding of Manuscripts in *Beta Maṣāḥəft*' (see the *Beta maṣāḥəft* project note, pp. 13–27). Anaïs Wion spoke of 'Corpus, manuscript, document: the basic XML-TEI architecture of Ethiopian Manuscript Archives (EMA) project and why manuscript matters' (see the *EMA* project note in this issue, pp. 33–38). Nathan Carlig in his paper 'Dealing with Coptic codices stratigraphy: two case studies' illustrated how the recently promoted 'archaeological' approaches to codicology has been applied to the description of manuscripts in the *PAThs* project database (see this issue, pp. 69–77). 'Coptic colophons and their relationship with manuscripts: typology, function, and structure' was the title of the paper of Agostino Soldati (published on pp. 115–119). Finally, Orhan Toy offered an insight into

the design of 'The *Islam in the Horn of Africa* Database'. The PostgreSQL database is a relational database management system (RDBMS) using Ruby on Rails as the general web application framework. The search is enabled by the ElasticSearch open-source engine.

Cooperation and interaction of the involved projects resulted in a number of immediate consequences. The direct contact allowed hands-on exploration of the possibilities of alignment of data architecture to a degree sufficient to allow for a maximum of data linking, sharing, and exchange. A significant degree of interoperability was achieved between the *Beta maṣāḥəft* web application and the *EMA* and *IslHornAfr* databases. Thus, the *Beta maṣāḥəft* schema was expanded to accommodate the typologies proposed by the *EMA* project, so that the *EMA* XML files could be directly imported and are now searchable and viewable on the *Beta maṣāḥəft* portal. The *IslHornAfr* data, encoded in a different format, could be converted and fully integrated with the manuscripts / works / places / persons entities of the *Beta maṣāḥəft* model. Members of the *Beta maṣāḥəft* and *PAThs* projects also extensively discussed the respective claves alignment as well as common strategies to dealing with ancient places and manuscript stratigraphy, resulting in set ups of interoperations among the active applications. *Syriaca.org* and *Beta maṣāḥəft* also joined efforts to produce accessible Linked Open Data aligning the ontologies used and setting up the possibility for cross-federated searches from the respective websites. The discussion also touched upon some future possibilities of joint academic initiatives, especially focused on production and exposition of Linked Open Data to allow queries across related datasets.

On a more general level, the discussion revolved around fundamental issues which have been central to the *COMSt* (*Comparative Oriental Manuscript Studies*) approach and community since the very inception of the work in 2009, and this makes it particularly reasonable that this collection of contributions appears as a monographic issue of this *COMSt Bulletin*: sharing and interoperability of data and resources, sustainability in time of generated data and later reutilization within a different framework, issues of common languages and formats (in cataloguing, describing, editing, annotating etc.), and in general, the search for common answers to common problems.

One should maybe also underline two further essential points:

(1) The projects involved in the workshop already gave important evidence and examples of concrete forms of mutual cooperation, as detailed above, but one should never forget that for most of them (*Beta maṣāḥəft*, *TraCES*, *PAThs*), their existence is deeply rooted in the work that was carried on for years, and even for decades, by former projects, the data of which it was possible to recover and re-utilize fully years later: this is the case, as far as

projects in Ethiopian studies are concerned, for the recovery of massive essential topographical, prosopographical, literary, codicological, and textual data from the *Encyclopaedia Aethiopica* and *Ethio-SPaRe: Cultural Heritage of Christian Ethiopia. Salvation, Preservation, Research* projects, and other and minor ones; and this is also the case of the *PAThs* project, with an even deeper chronological background, for the reutilization of codicological, literary and textual data of the *CMCL* (*Corpus dei manoscritti copti letterari*) project.

(2) If comparison on a larger scale is a fascinating and challenging task, the comparison on a smaller scale, as is possible for areas that are homogeneous from the point of view of their geographical distribution and cultural coherence and consistency, like the domain of Coptic, Ethiopic (including Ethiopian Islamic), and Syriac manuscripts, still offers a privileged vantage point of observation and experimentation, that cannot be replaced by comparison on a global and larger scale: in this small scale there is still a huge unexploited potential of cooperation that can contribute to optimize and make progress small fields at a fast pace by joining forces together, in a perfect '*COMSt* spirit', building up more solid premises for broader cooperations and large-scale comparisons.

*

The contributions in this Special Issue all originate in the papers presented during the 'Linking Manuscripts' workshop. The articles are grouped in two chapters, Project Notes, introducing the research initiatives involved in varying degree of detail, and Case Studies, illustrating individual solutions to particular problems. Within each chapter, the papers are arranged alphabetically, by project title in the Project Notes and by the author's name in the Case Studies.

An alphabetical index including proper names (persons, places, projects, institutions) and literary works completes the volume.