ENEA Contributions to Safeguarding Material and Immaterial Knowledge from Ancient Mesopotamia

Thanks to the "Duplicazione e Rinascita" (Duplication and Rebirth) Project, funded by the Ministry of Foreign Affairs and aimed at supporting the Baghdad Museum, since the beginning of the year 2000 studies have begun in ENEA, with the purpose of protecting material goods and the great deal of knowledge of the ancient civilizations flourished in the ancient Mesopotamia, now Iraq.
Although the studies, projects and activities carried out share the use of data processing in the broadest sense, they focus on three different lines of research, which have faced and are still facing the problem of preservation of ancient artifacts or – equally important – maintenance and dissemination of ancient knowledge from different standpoints. The three technologies at issue are, respectively: the Reverse Engineering and Rapid Prototyping; the integration of Language Technologies (Multilingual Text Mining) and GRID Technologies in the ENEA-GRID infrastructure, for the analysis of ancient texts in Assyriology; the use of ICT technologies for the collection, use and dissemination of cultural resources

Paola Negri, Sergio Petronilli, Giovanni Ponti, Daniela Alderuccio, Giorgio Mencuccini, Alessio Rocchi, Flavio Fontana

Introduction

Among the many archaeological objects, a special place is occupied by clay tablets written in cuneiform writing and coming from the Ancient Near East (the so-called Fertile Crescent). Beyond the purely assyriological problems, there are many issues crucial to the conservation and use of these ancient documents. In the last decade, three classes of problems have been identified, for which solutions could be found thanks to the technologies employed within ENEA in a framework of interdisciplinary cooperation.

Millions of texts were written in ancient times, and though only hundreds of thousands of those tablets are preserved, it is nevertheless an extremely important amount, often difficult to be mastered by scholars.

Whilst clay is one of the strongest writing materials employed by the man, it is not free from damage, that may result from chemical factors, such as the crystallization of the salt content in the clay, or by mechanical factors, such as fragmentation and breakage. Moreover, in ancient times tablets were not always baked – an expensive procedure reserved

Paola Negri
ENEA, Unità Tecnica Tecnologie dei Materiali
Sergio Petronilli
ENEA, Unità Trasferimento Tecnologico
Daniela Alderuccio, Giorgio Mencuccini, Giovanni Ponti, Alessio Rocchi
ENEA, Unità Tecnica Sviluppo Sistemi per l'Informatica e l'ICT
Flavio Fontana
ENEA, Unità Tecnica Modellistica Energetica Ambientale

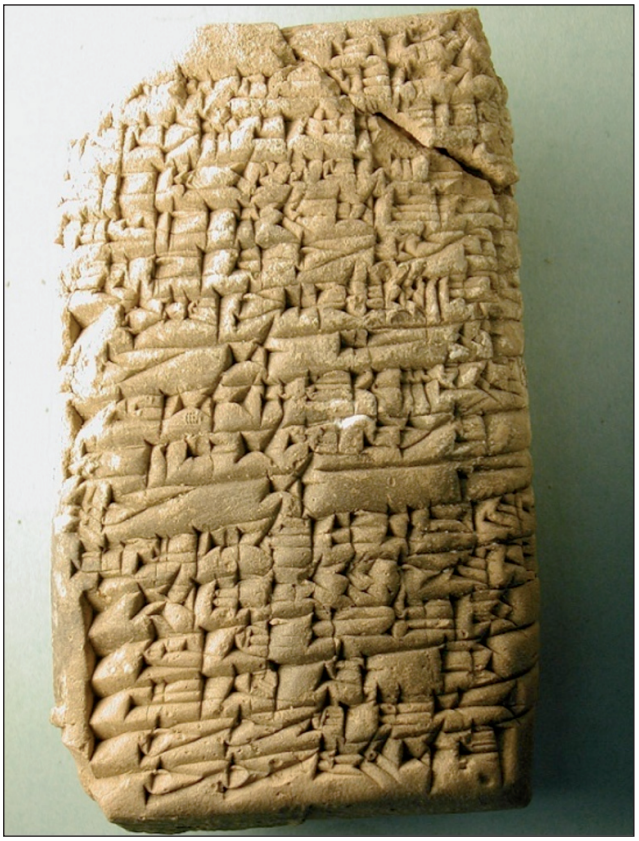

FIGURE 1 Oj 32, an Old-Babylonian cuneiform tablet (about 1800 B.C.) from Isin related to a sale of real estate
Source: ENEA / Centro Studi Diyala

for the most valuable documents destined to last – whereas the other ones, more ephemeral, were only sun dried and therefore are more subject to damage. Moreover, scholars necessarily handle the tablets to study the texts and this can damage them, although conservation treatments and methods have been developed since many years.
In addition, for study, teaching or exhibition purposes, often in modern times copies are made from casts, with procedures that may damage the tablet, because plastic substances come into direct contact with the tablet itself. So often fragments of the surface material are removed by contact with disastrous results for the conservation of the tablet itself. To solve the problem, the proposed technology is the Reverse Engineering and Rapid Prototyping, which is well developed here in ENEA.
Furthermore, in order to enhance access to text knowledge extracted from the cuneiform tablets, the integration of Language Technologies (Multilingual Text Mining) and GRID Technologies in the ENEA-GRID infrastructure offers a new perspective in the analysis of ancient texts in Assyriology, supporting scholars in mastering a large number of texts from the linguistic point of view and in managing the complexity of grammar.
Still a third class of problems arises from the number of the texts, so that for years in the field of Assyriology many different types of databases have been prepared separately by a variety of institutions. This type of information does not allow fast access, consistent and complete documentation. ICT applied to this particular type of cultural heritage can be an important response to this problem.
So, applying existing ENEA resources to an extremely complex field can both refine the techniques in our possession and provide high-value tools for a better use of cultural heritage.

Reverse Engineering and Rapid Prototyping

The term RAPID PROTOTYPING actually indicates a group of related technologies that are in use to make objects directly from CAD data sources. This methodology builds objects adding and bonding materials in layers. It is a fast and suitable solution for solving fabrication problems of objects with any geometric complexity or intricacy, that are difficult to make by a material subtraction manufacturing.
For this reason it is widely employed in several fields; in ENEA, the PROTOCENTER Laboratory employs this complex methodology in support of SMEs and handicraft, like fashion houses and jeweller's art, offering not only technological support, but also knowledge dissemination to enterprises.
In the Nineties it was observed that works of art and archaeological objects are also reproduced with this technology, and therefore in ENEA some experiments were made: as a result of these experiments some

objects have been replicated, the most relevant of which is the *Situla della Certosa*, in the Museum of Bologna.
In 2000, after that experience, a new type of archeological objects was taken into consideration, for the first time in the world: cuneiform tablets, and a preliminary attempt of applying RP to them was carried on in ENEA by a multidisciplinary team including S. Petronilli e P. Negri. Subject of the test was the cast of a XV century B.C. tablet. Testing that technology on a cast and not on a real tablet ensured that the tablet did not run any risk.
Main problems in this experiment were: 1) to obtain a readable cuneiform text; 2) to match questions concerning the form of the tablets, that are usually convex, with the text covering not only obverse and reverse, but also the four edges.
The preliminary model was built by scanning the cast of the tablet by a laser triangulation scanner. The resulting points were used as a reference for assembling the resulting "clouds of points" obtained by a high-resolution piezoelectric system. Therefore, the flexibility of a laser in reassembling "clouds of points" and the very high resolution of a piezoelectric system were mixed together.
In this preliminary experience six laser scanning and two piezoelectric scanning were performed in order to acquire all the surfaces of the tablet.
The resulting file fed an RP machine working with a multi jet modelling technology, and a model of the cast was made by adding and bonding a thermoplastic material in layers. The results encouraged the team to go on with experiment, in order to improve the use of the technology. An RP process for replicating cuneiform tablets was then established, which is similar to that usually employed for other materials, but takes into consideration the specificities of the tablets.
Based on that previous experience, in ENEA Laboratories another Project started, the *kima labirišu* Project – which means "like its original Project" – aimed at co-operating with Institutions and Museums to safeguard the cuneiform clay tablets. That Project was developed on the occasion of a special Project in support of the Baghdad Museum, the "Duplicazione e Rinascita" Project, funded by the Ministry for Foreign Affairs (MAE).
On that occasion 89, badly damaged tablets belonging to a private art dealer were scanned and replied. The *kima labirišu* Project aims not only at making copies or casts of tablets without touching them and reducing their handling to the minimum, but also at experimenting the creation of 3D images so precise that they could be used by students to read, analyze, study and publish texts.
In the three-dimensional scanning of the tablets, two different systems have been employed to obtain a high resolution acquisition: a laser triangulation system and a structured light system. For each tablet about 30 scannings have been processed, to a 400 p./mm2 resolution. The aim is to acquire a complete geometry of the cuneiform signs. This means to obtain and manage a file of 300Mb.
Starting from the 3D models, replicas of the tablets have been created through the technology of the 3DP (Three Dimensional Printing) Rapid Proto-typing. By applying the texture of the original tablet it is possible to obtain a coloured copy of the tablet, sending the machinery a virtual model of the tablet on which it has been applied.
The case of these 89 tablets has been emblematic: 1) thanks to either virtual models and real copies, very fragile exemplars are at the disposal of the scholars who can handle, study, analyse them without any problem for their safeguard and conservation; 2) although they belong to a private collection, and are therefore not freely available to the scientific community, thanks to the copies and 3D images obtained with permission by the collection owner they can be studied and the information drawn from them can be used into the wider system of knowledge. 3) Moreover, as a result of the *kima labirišu* Project, data concerning the 89 tablets have become available to the Iraq Museum, together with the results of the "Duplicazione e Ricerca" Project.
This Project has not come into being without a successor: as a further activity the group of the scholars of the "Duplicazione e Ricerca" Project, of which P. Negri is also part, is publishing the texts just starting from the 3D images. From a preliminary examination, it results that these tablets offer some new information on the environment of the Southern Mesopotamia in the II

millennium B.C. and add new knowledge on forms of contracts, chronology (new dates of kings) and scribal system.
For its part, the PROTOCENTER Laboratory is going on with activities concerning cultural heritage, as the Horemheb Project with the Museum of Cairo demonstrates. The aim of the Horemheb Project is to make copies of some stone reliefs so as to bring the tomb of the pharaoh Horemheb, located in Egypt in the necropolis of Saqqara and currently incomplete, to its original condition. The copies of the reliefs will be made starting from 3D models performed by the ENEA PROTOCENTER Laboratory by techniques of high resolution laser scanning, so as to avoid any contact with the originals and then the copies in stone material similar to the original will be produced through a system of 3-axis milling.

ENEA-GRID and e-Assyriology

With its fast and secure access to software and hardware resources, ENEA-GRID allows to share computing resources across collaborative projects, attracting, engaging and supporting a wide range of users and researchers from science and industry communities. For these reasons, ENEA-GRID facilities can be used in several contexts and fields of interest, such as weather forecasting, fluid dynamics, 3D applications, etc.
In this contribution, we want to show our activities and research on the integration of language technologies (e.g., text mining) in the ENEA-GRID infrastructure to provide a novel approach for studying and analyzing cuneiform corpora. In the domain of Assyriology, ENEA-GRID offers a digital collaborative environment, to share knowledge and digital resources with the integration of Multilingual Text Mining Software, Lexical Resources and Data Visualization tools for Network Analysis. The GRID enables researchers to perform quantitative and comparative studies on transliterated cuneiform texts, providing access to computational resources for the storage and processing of large textual corpora.
For this purpose, the TIGRIS Project (**T**oward **I**ntegration of e-tools in **GR**id infrastructure for e-a**S**syriology) started in 2008, and preliminary studies by P. Negri and D. Alderuccio [Negri-Alderuccio, 2009 and 2011] provided the basis for the activity for integrating and analyzing assyriological texts.
A first experimental application of new technologies to a corpora of Babylonian texts has been carried out using Multilingual Text Mining in ancient languages, in the prospective of an integration into the digital environment in ENEA-GRID. The main goal here consists in offering an innovative approach for the study of cuneiform corpora, from which software developments and collaborative working could coexist to match the specific needs of scholar communities.
The first case-study is referred to the texts of the town of Nuzi [Negri-Alderuccio et al., 2011], belonging to the small Kingdom of Arrapkha and located East of the Tigris River (XV-XIV centuries B.C.). Among the whole corpus of Nuzi texts, the best preserved texts belonging to the scribal family of Šeršiia (with his son Hupita and his grandson Muš-teššup) has been analyzed, in order to highlight innovative and conservative stylistic elements in the redaction of texts (administrative texts and contracts) through three generations. Documents have been tagged in order to put graphic and graphemic, grammatical, and prosopographical data into light. The small dimension of the corpus allows for an exhaustive control of the results in the application of the software. Preliminary results invite to continue with the application of these methodologies also to the whole corpus of the Nuzi texts.
Recently, we started from the past experience on assyriological texts and proposed a new strategy to analyze such data. In particular, our latest study consists in exploiting data mining algorithms tuned on assyriological transliterated corpora from cuneiform tablets. Such an approach has been developed by G. Ponti (UTICT) [Ponti et al., 2012] and supports document clustering of assyriological e-texts.
Clustering analysis is an unsupervised technique that allows correlating data and grouping them to highlight hidden relations and relevant patterns. Such information is particularly precious for the works of assyriologists, who may exploit such a data reorganization to discover new interesting relations and considerations for their studies.
In our study, we used as a case study a corpus of 50 letters from old-Babylonian Kingdom of Eshnunna, flourished

(XXI-XVIII centuries B.C) along the Diyala River, an affluent of the Tigris River. In particular, in an ambitious attempt of reconstructing the grammar of that local form of old-Babylonian language (XVIII century B.C), we resorted to letters, instead of administrative texts or contracts, because their language is more varied, thanks to the different, widely ranging arguments, whilst standard, repetitive formulas are prevailing in the contracts; for these reasons, letters are particularly suitable for text analysis. Cuneiform texts have been preprocessed (as we will describe in the following paragraph), and then clustering algorithms have been executed. We employed the well-known K-means clustering algorithm, which partitions the corpus into groups (i.e., clusters) according to the number chosen by the analyst. K-means typically assumes that texts are represented by means of the Vector Space Model, where a document is seen as a vector of its terms and each term relevance depends on the statistical/correlation measure employed. A two-stage analysis has been performed: a quantitative analysis accessed the quality of the clustering task exploiting quality-based indexes known in literature, whereas a qualitative analysis aimed at describing data relations and affinities discovered by the clustering algorithm involving assyriologists.

Both cases of study require a preprocessing phase, connected with the graphemic and linguistic characteristics of the documents. Text Mining and Lexical Analysis Software require adaptation in order to deal with problems related to cuneiform texts, such as graphemic ambiguities and inconsistencies, the use of more than a language in the same text, questions connected with printing fonts and the use of particular characters, and other peculiarities like the lack of punctuation and/or space and carriage return (the only separators of graphic forms). The preprocessing task on Eshnunna corpus consisted in two main steps, that are the transliteration phase, in which cuneiform texts have been transcribed in Latin alphabet, and the lemmatization phase, where nouns, adjectives, and verbs have been brought back to the base standard form.

The research is still in progress, but some aspects and preliminary results can be underlined: in particular, the analysis of the extracted forms aims at best evaluating either the completeness of the extracted data and their typology. At the moment attention is focused on the declension of the nouns and the conjugation of the triradical stark verbs, with a particular care to the D forms, in view of a future phase in which semantic aspects will be taken into account.

Data collected until now in both case-studies are very interesting and encourage ENEA researchers to carry on. Next steps will be to extend the analyses to the whole corpora and to open an international co-operation, thanks to the HPC systems of the GRID and the opportunities offered by the GRID Virtual Lab TIGRIS (www.afs.enea.it/project/tigris). ENEA-TIGRIS is a Virtual Lab in e-Humanities, based on the integration of e-tools in GRID infrastructure for e-Assyriology.

TIGRIS Virtual Lab is the digital environment where the Assyriologists and research teams can meet and collaboratively work together, exchanging research results, best practices, fonts, software, e-text collection, lexical resources (dictionaries, grammars, ontologies, genealogies and name dictionaries, etc), images, etc.

At present, integrating e-tools in the GRID virtual environment enables ENEA-GRID for e-Humanities.

ICT and H-DATA (Historical Documents on Ancient Technologies and Artifacts)

From many years in ENEA, attention is devoted to a sector of the cultural heritage: the documents on ancient technologies and on the way in which artifacts and works of art were created in ancient times. The reconstruction of ancient technologies can concern academic purposes, like history of the science and history of art, or conservative purposes, in particular restoration, and also other fields, like social and economic history. In the last decades, interest in these documents long considered not particularly significant has been shown, as many activities in the world widely demonstrate: one for all, the recent experiences made by the Fraunhofer Institut in the German National Vitruvius Project, where analyses of ancient technologies are linked with the Latin text and with the experimental reconstruction of roman decorative wall painting techniques.

This branch of knowledge is part of the immense cultural and scientific legacy of thought and knowledge passed on from ancient times to the present time Europe, to ensure the conservation of which great attention is paid. Safeguarding either the immaterial aspects, such ideas and knowledge, and material supports, such as books and written documents in general, is presently one of the scholars' tasks. The great amount of these cultural resources makes it often difficult to ensure their conservation and accessibility.

The study of ancient materials extends in large span of time incompatible with the short project times, and scientists are facing and defining more and more complex data. In the past this kind of work was carried on exclusively in a restricted way, without any cooperative network approach and integration. The applications, for which medium-sized personal databases are used, generally have limited functionality. In fact, two solutions to these problems were possible: 1) a unique standard, static, relational database, limited in data domains, 2) several different, heterogeneous databases, more complex to manage. Moreover, multiple standard technologies, sometimes not specifically dedicated to their requirements, have been often employed. They require an ICT expert support because of their complexity in being managed. Moreover, if any modification to the databases is necessary, realizing it is very difficult and requires a co-presence of the technical staff, with high costs in terms of time and money. The existence of strictly thematic data bases, which interact with great difficulty, makes it often impossible to match the data and information loaded in different supports without a particularly complex intervention of technology experts. So, for example, a text record cannot be linked with an archaeological record containing relevant information, unless this link has been projected in advance; a later planning should cost more and more time and money. Therefore, a project facilitating the access to different areas of cultural heritage through a wide range of technical formats, which can enable the efficient use of digitized cultural resources, seems particularly important, as it allows promoting active research and new knowledge, easier access to the documentation, included that out of reach and from less known sectors, and also the development of new products in creative industries.

In consideration of the presence and importance of ICTs in everyday life, which can offer innovative solutions to a wide range of problems, included those concerning cultural heritage, P. Negri and F. Fontana are engaged from some time in an ICT project that can connect written data on ancient technologies with archaeological and archaeometric data.

Preceded by moments of experimentation and preparation particularly useful to fine-tune a few factors, the project strategy regards the implementation of innovative intelligent systems finalized to collect different types of documents and historical materials, with autonomous software agents and standard procedures, to catalogue them with innovative ways that reduce the specific experts' manual work. To perform this activity, experts could define the selection criteria to the platform in a previous task, simply using their knowledge to define the characteristics and parameters for the classification and selection phases to the system. Then, the system will select the information and knowledge of H-DATA interest autonomously.

In the H-DATA project, the design and implementation of a portal (H-DATA Meta Portal - MP) are planned, together with an advanced communication system (H-DATA Meta Multi Platform - MMP), that can interface with other present and operational e-learning platforms and databases. The MP is finalized to the functional aspects of the H-DATA Project, and manages and retrieves information and multimedia data. The platform will also be used for the following activities:

1) Information about the textual, historical documents will be refined and integrated with audio, video and textual contributes;
2) The data structure could be dynamically modified in relation to the type of documents and their specific characteristics;
3) This platform will be dynamic and include applications suitable for reducing divergence of technical approaches, in order to improve the usability of the system;
4) The platform will be integrated with an innovative meta search engine, to offer new perspectives and possibilities to end users;

5) With a parallel and independent development, the platform can integrate also a meta e-learning platform based on different resources and 3D objects;
6) Links and connections with other existing Databases and systems can allow the integration among other scattered materials.

This way, the H-DATA platform can offer an answer to particular requirements in the cultural heritage and the possibility for final users of accessing freely and easily to interdisciplinary materials and data. There is still another aspect of great importance: loading data is often a very complex burden, because of the number and complexity of data and the many possibilities of error. A system that can reduce problems in loading data, without reducing quality, is certainly welcome. The H-DATA platform offers a revolutionary solution to all these problems and also opens to further unexpected developments, because it goes beyond the system of data base, with innovative devices based on a new approach to problems.

At present, as the first case-study, texts taken into account are those related to building techniques, glass and metals, in Greek, Latin and also in Assyro-Babylonian (languages of ancient Mesopotamia): the latter are very interesting, difficult to be found by non-specialists, and considered the antecedents of the classic cultures and important comparison elements. Mesopotamian receipts for making glasses offer data on the beginning of a technology, while documents spanning from the Mesopotamian handicraftsmen to the Roman *vitrarii* offer information on the development into a workers' category. Tablets on brick manufacturing (costs for materials and workers, working days, etc.) can be paralleled with Greek epigraphs on the construction costs and times of the Parthenon. Apart is the sector of metallurgy, and of documents on metals, leagues and handicraftsmen. And examples could be innumerous.

In any case, thanks to the integrated knowledge that the platform could allow, further data and information can be retrieved in order to offer information to a very large audience and to different types of users.

Conclusions

In this framework, it is interesting to take also into consideration what S. Ross observed already in his 2004 paper, where the importance of the cultural and scientific heritage of Europe as a source of raw materials for economic, intellectual, social, and cultural development in the 21st century is put into light; in particular not only does Ross stress that the value of these raw materials multiplies when they are available in digital form, but also that these materials contribute to improvements in the quality of life, enabling education, supporting life-long learning, underpinning the development of new products by creative industries.

References

[1] S. Ross, 2004, "Progress from National Initiatives towards European Strategies for Digitisation" in Towards a Continuum of Digital Heritage: Strategies for a European Area of Digital Cultural Resources, European Conference, (Den Haag: Dutch Ministry of Education, Culture and Science), pp 88-98.

[2] F. Fontana, M. Moscarini (2004), An Advanced e-Learning Platform to Manage Virtual Courses and Multimedia: MATRIX 3, ENEA, University of Rome, Italy, ED-MEDIA 2004, Lugano, Switzerland.

[3] F. Fontana, E. Cosimi, G. Ponzo, ENEA Usability Lab, Department of Computer Science, An Advanced Web 3D Mobile e-Learning System to Manage Virtual Courses and Multimedia, University of Rome, Italy, E-LEARN 2010, Orlando, Florida, USA.

[4] P. Negri Scafa, D. Alderuccio, (2009), ENEA, "A new experimental approach to text computer analysis applied to the Nuzi texts", 55th RAI 2009, "Family in the Ancient Near East: Realities, Symbolisms, and Images", Paris, France, July 6th-9th 2009.

[5] P. Negri Scafa, D. Alderuccio (2011), Giovanni Bracco, Silvio Migliori, ENEA "A preliminary analysis of a Nuzi scribal family in view of an application in the ENEA-GRID", 57th RAI 2011; "Tradition and Innovation in the Ancient Near East", Rome, Italy, July 4th-8th 2011.

[6] G. Ponti, P. Negri Scafa, D. Alderuccio, G. Mencuccini, A. Rocchi, G. Bracco, S. Migliori, (2012) ENEA, "Toward the integration of informatic tools and GRID Infrastructure for Assyriology Text Analysis" 58th RAI 2012; "Private and State (Privat und Staat) in the Ancient Near East", Leiden, The Netherlands, July 16th-20th 2012.

[7] ENEA PROTOCENTER: www.protocenter.enea.it

[8] ENEA GRID: www.eneagrid.enea.it

[9] GRID Virtual Lab TIGRIS: www.afs.enea.it/project/tigris

[10] ICT_H-DATA: odl.casaccia.enea.it

[11] www.desire-net.enea.it