REGNET

Cultural Heritage in REGional NETworks

Deliverable D4

Available Content and Products

Project acronym	REGNET	**Contract nr.**	IST-2000-26336
Type and Number	D4 – REGNET – Available Content and Products		
Work package	WP2 - Implementation of the System and Preparation of Services and Product Generation		
Task	T 2.1 Preparation of content and products		
Date of delivery	2002-04-19		
Code name	RN_D4v01	**Version** 01 draft ☐ final ☑	
Objective	Report		
Distribution Type	Restricted		
Authors (Partner)	IMAC, AIT, ONB, SUL, LMG, NRM, KVA, TARX, MECH, MUS, SPAC, ALI, CC, IAT, GRAN, ICCS		
Abstract	The report focuses on the work and main results of task 2.1 belonging to WP 2. According to the plans and guidelines elaborated in Task 1.2 digital content available at the content providers side was produced and prepared for the integration in the REGNET system. The report describes the collections available for the demonstration phase, the process of preparation and conversion together with the available tools. It gives examples for the presentation of data e. g. in topic maps.		
Keywords List	Content provider, content preparation, digitizing activities, data entry tools, content examples.		

ist information society technologies

Table of Contents

Executive Summary

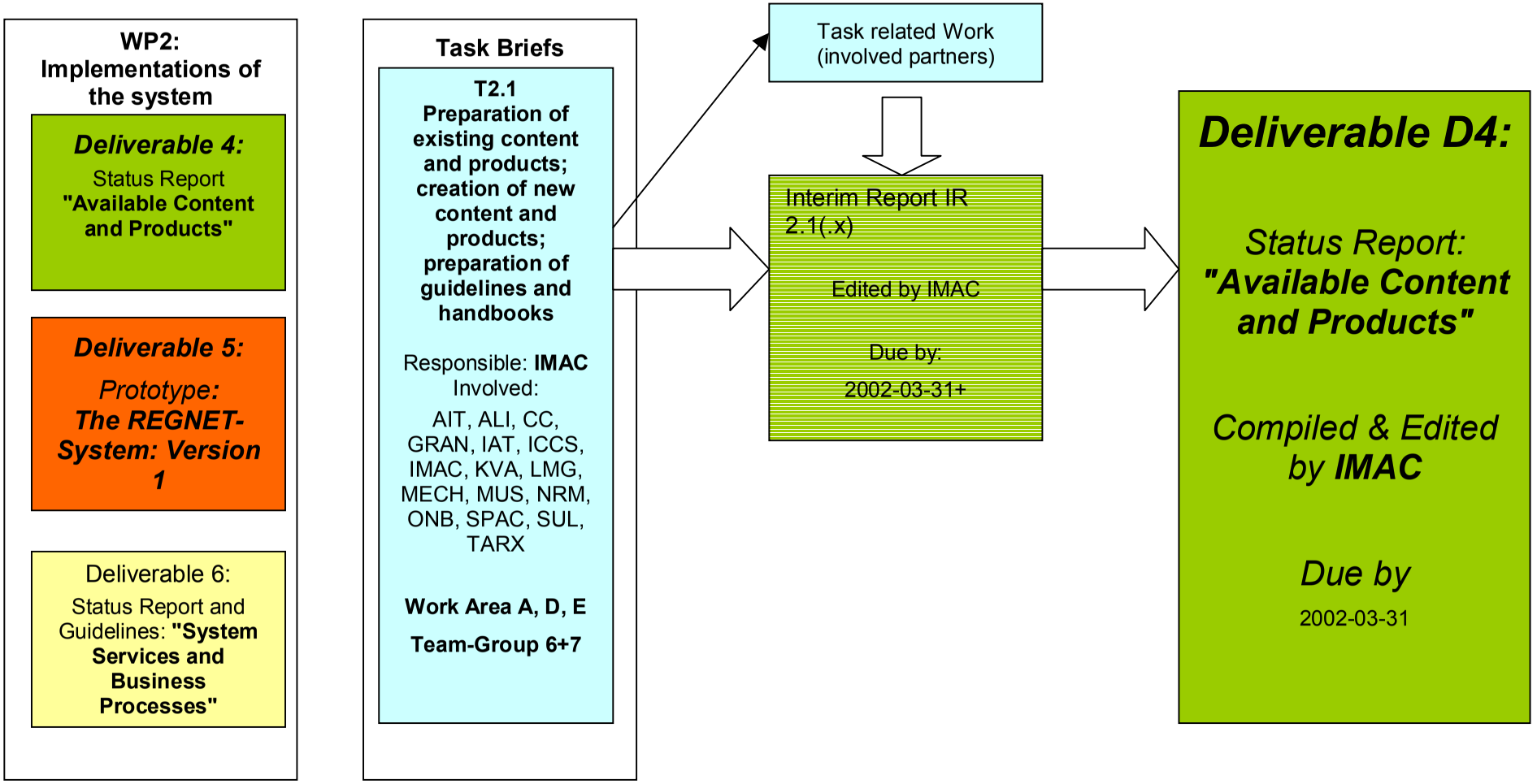
The focus of this document lies upon the "Definition of content to be provided" at one side and “The digitizing plans and activities” on the other side.

Situation

	Implementation of the System and Preparation of the Demonstration (WP2)				***Validation of the REGNET-Demonstrator and Preparation of the Demonstration Phase (WP3)***			
	Task	**Leader**	**Document**	**MM**	**Task**	**Leader**	**Document**	**MM**
Preparation of contents and products	2.1	IMAC	IR 2.1 → **D4**	42,5				
Validation of the REGNET-Demonstrator					3.1	IMAC	IR 3.1 → **D7**	18
Preparation of the Demonstration Phase					3.2	VALT	IR 3.2 → **D7**	33

Task 2.1 was split in two sub tasks in order to improve the management and to strengthen the importance of thematic content products within the project. The validation of the REGNET-Demonstrator and data developed and prepared will be handled in WP 3.

CONTENT ENGINEERING (Domain / Regional – specific)



1 Introduction

1.1 Purpose

According to the plans and the guidelines elaborated in Task 1.2 digital content available at the side of the content providers has to be prepared to be available for the REGNET demonstration phase. The already selected objects must be described ("bibliographic records") and thematic texts related to the themes agreed upon in WP 1 have to be written by the content providers. The type of loading data into the REGNET system will be elaborated and converted to the appropriate format within the system. Figure 1 outlines the main focus of work in WP 2.1 on the basis of the status reports given by all partners.

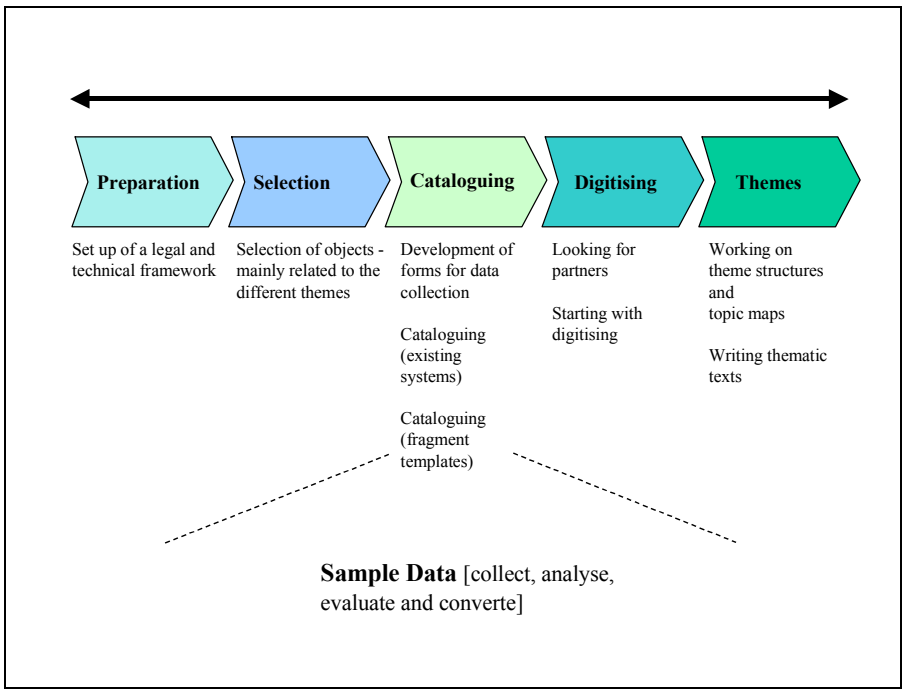


Figure 1: Main focus of work in WP 2.1

The work done by each partner (content provider) will be described shortly in chapter 2 in order to illustrate the status of content preparation. The content itself will be available in the REGNET demonstration system.

In order to increase the overview and to improve the task management the whole task was divided in two sub tasks: WP 2.1.1 Object preparation and digitizing and WP 2.1.2 Preparation of thematic content. WP 2.1.1 was dedicated to all issues related to cataloguing and digitizing as well as the provision of sample data in order to stimulate the development of appropriate tools. WP 2.1.2 should handle the thematic content production as defined in WP 1. The theme-based approach within REGNET was developed as one possibility to realize a thematic access to the REGNET system and to offer added value in connecting objects to each other. Besides the direct search in the catalogues and the navigation through predefined access points it should provide further access to digital content in the system. A theme is defined as a set of thematic texts and/or objects and thereby consists of already existing products (the objects in the collections of the content providers) and new products like thematic structures/views for navigation (as well as topic maps) and thematic texts for further information.

According to the theme concept and the theme-based work already done in WP 1, the sub task 2.1.2 is dedicated to the preparation of

- theme structures for all selected themes,

- fragments to be reused in different themes,
- thematic views for navigation (to be realised as topic maps, WP 2.2.4),

The delivery of all thematic products should follow the templates worked out by the theme expert group and was validated by an editorial team.

1.2 Overview and document structure

This report focuses on the description of the status of available content and products and the steps undertaken to convert existing data using project-related tools and standards. Starting from the description of the work done at the side of the content providers (status reports) the following elements are picked up:

- the available collections and the introduction of new partners,
- description of available online and offline tools for data entry and data conversion,
- further steps undertaken with regard to the thematic content production,
- examples of available content.

The results of this WP will include concrete guidelines for the usage of the tools available (online and offline) within the Appendix.

1.3 Work plan and partners

The directly involved partners and the assigned tasks are listed in Table 1.

Partner Name	*Acronym*	*Task*
Information & Management Consulting, Berlin, GERMANY	IMAC	Coordination of task. Provision of support in selection appropriate data structures. Recommendation of standards. Provision of sample data. Validating sample data. Work on the data entry component. Baseline for IR 2.1/D4, reporting.
AIT-Angewandte Informationstechnik GmbH, Graz, AUSTRIA	AIT	Coordination with technical development/data entry component. Contribution to IR 2.1 (Conceptional issues)
Österreichische Nationalbibliothek, Wien, AUSTRIA	ONB	Provision of sample data. Object description and digitizing. Data entry. Theme structure/fragment list for "Habsburg" (with special focus on Elisabeth). Thematic content production (fragments, object linkage). Contribution to IR 2.1
Stockholm University Library, Stockholm, SWEDEN	SUL	Provision of sample data. Object description and digitizing. Data entry. Validation of thematic content production (editorial group). Theme structure/fragment list for "Linnaeus". Thematic content production (fragments, object linkage). Contribution to IR 2.1
Lansmuseet pa Gotland, Visby, SWEDEN	LMG	Provision of sample data. Object description and digitizing. Data entry. Thematic content production (fragments, object linkage). Contribution to IR 2.1
Naturhistorika riksmuseet, Stockholm, SWEDEN	NRM	Provision of sample data. Object description and digitizing. Data entry. Thematic content production (fragments, object linkage). Contribution to IR 2.1

Kungl. Vetenskapsakademien Royal Swedish Academy of Science, Stockholm, SWEDEN	KVA	Provision of sample data. Object description and digitizing. Data entry. Thematic content production (fragments, object linkage). Contribution to IR 2.1
TARX nv, Hofstade, BELGIUM	TARX	Coordination with task 2.1.2. Support of other partners. Validation of thematic content production (editorial group). Provision of further templates, guidelines etc. if necessary. Support in generating of theme structures. Coordination with technical group. Topic MappingContribution to IR 2.1.
Stedelijke Musea Mechelen, Mechelen, BELGIUM	MECH	Provision of sample data. Object description and digitizing. Data entry. Theme structure/fragment list for "Saints" and "Gilt Leather". Thematic content production (fragments, object linkage).Contribution to IR 2.1
MUSEON, Den Haag, NETHERLANDS	MUS	Provision of sample data. Object description and digitizing. Data entry. Theme structure/fragment list for "Masks & Amulets" and "Measuring" and "Leather". Thematic content production (fragments, object linkage). Validation of thematic content production (editorial group). Contribution to IR 2.1.
Fratelli Alinari SpA, Florence, ITALY	ALI	Provision of sample data. Object description and digitizing. Data entry. Thematic content production (fragments, object linkage). Contribution to IR 2.1
Consorzio Civita, Rome, ITALY	CC	Integration of the new Italian museums. Documentation and digitizing plans for the Italian museums. Theme structure/fragment list for "Early Renaissance in Europe". Thematic content production (fragments, object linkage). Contribution to IR 2.1.
Institute of Computer and Communication Systems, Bulgarian Academy of Sciences, Sofia, BULGARIA	ICCS	Provision of sample data. Object description and digitizing. Data entry. Theme structure/fragment list for "Art Gallery Bulg./Russ." Thematic content production (fragments, object linkage), especially: working out a set of general context topic maps (art-painting-sculpture-graphics-...). Contribution to IR 2.1
Granollers City Council, Granollers, SPAIN	GRAN	Provision of sample data. Object description and digitizing. Data entry. Thematic content production (fragments, object linkage). Contribution to IR 2.1
IAT - Instituto Andaluz de Tecnologia, Sevilla, SPAIN	IAT	Contribution to IR 2.1
SPACE S.r.l., , Pratro, ITALY	SPACE	Contribution to IR 2.1

Table 1: Partners involved in WP 2.1

2 Status of work at the side of the content providers

Due to the fact that the main will be reflected in the content available in the REGNET system additionally a short report of work done and actual status should be given in this chapter 2.

2.1 ONB

Digitizing activities

ONB has harvested information about high quality archival scanners best suited for picture digitization, and has been able to finalize the purchase of a product which represents the state of the art in archival digitization, a Homrich Imaging Technik (HIT) vario digital XL working with Sinarback 23 and Phase One Power Phase FX digital backs, enabling the digitization of objects of smallest formats up to A0-format and larger, in best quality. The scanner and its peripheral components (vacuum table, computer, lights etc...) were installed in ONB. It necessitated the building of a working platform in concrete, to ensure best working conditions with a minimum of floor vibrations. Experiences with scanning and data exchange were made. ONB staff was trained on the new scanner unit, which is already connected to the ONB net. The new scanner unit was connected to the ONB net and ONB staff was trained on it. When the scanner was ready to be used (also for consumer and ONB internal digitization services), ONB digitized the complete historical Habsburg Genealogy tree according to Primisser (Pictures as well as historical text). Approx. The 110 full size pictures and thumb-nails thereof which are to be included in REGNET as an example for a Topic Map were delivered to AIT.

ONB also digitized 33 plant illustrations as a contribution to the Linnean theme (theme leader: SUL/KVA) and delivered them to SUL. ONB further attended a group meeting of the technical partners in Graz (18.-19.3.02) and initiated a separate meeting with AIT (28.03.2002, also in Graz) to discuss the integration of ONB catalogue data (which already exists in digitized form) into the REGNET system.

Thematic content production:

Structures and fragments were worked out according to ONB model digitization plan (Target Collections / Themes) and ONB themes inventory. ONB discussed and elaborated the structure of the Habsburg Genealogy according to Primisser and put it down in an excel spreadsheet, defining all the interrelations necessary for integration of the family tree in an exemplary Topic Map. As part of the content creation, also the historic text information was supplied (see above). For the Linnean theme, coordinated by the responsible Swedish partner SUL/KVA, ONB contributed 69 text fragments in English, German and French language in Dublin Core format and 33 plant illustrations. Furthermore, the complete virtual exhibition of “Empress Sissi” on the ONB-Website was delivered to AIT as a contribution to the Habsburg theme (approx. 100 pictures in full size and thumbnail format as well as related thematic text).

10 Top Pieces of ONB collections were selected. Picture data and related text fragments are ready to be imported in the “Tour d’horizon”-theme of REGNET.

2.2 SUL

The rare books collection of the Natural Sciences at Stockholm University Library is although small a collection of excellent standing and value, nationally and internationally. The collection is actually owned by the Royal Swedish Academy of Sciences (KVA) but is housed and cared for by Stockholm University Library. There is for instance a comprehensive collection of Carl von Linnés publications including works that he himself donated to the library of the Royal Swedish Academy. A large collection of mathematical history, a collection by and on Swedenborg and a collection called Bergianska Biblioteket containing approximately 5.340 works on natural sciences. These books are part of the national cultural heritage and should be available internationally. Some of the gems from this collection has been chosen for the REGNET system to be available to users world-wide.

Cataloguing and Digitizing

110 images have been digitized, selected from books in this collection. The images are available in TIFF and JPEG formats on CD-ROM and SU server. The images are predominantly chosen from books on botany and zoology and adhering to the Linnéan theme. Ca. 120 records have been catalogued according to the Dublin Core standard in English and Swedish (240 records in all). The records are image descriptions and fragment records on author /publishing detail, having on an average 3 subject categories. A preliminary linking has been done between header records and image description records, linking to LIBRIS records will be done when the REGNET system has a database. The records have been produced in word to match the fragment descriptions and will be converted in

the REGNET system. A file of around 1000 bibliographic records from the Natural Sciences Collection have also been sent to AIT for the bibliographic database of REGNET.

Thematic Content Production

SUL has worked on the Linnéan theme together with the Royal Swedish Academy of Sciences. A work-description on the theme was first of all sent out to MUSEON, ONB and ALINARI, partners that had expressed an interest to take part in the Linnean theme. Christer Karlsson, employed b KVA has produced thematic texts, scientifically sound, around the Linnean tradition in English and Swedish, in short and full version and done the linking to fragments on books and images. Records have been obtained from ONB in English, German and French and they have been linked to fragments on Linné. The images, chosen from ONB have been photographed and digitized from the outstanding collection of Rare Books at the Austrian National Library. These images were published earlier in Garden Eden. MUSEON has sent a description of their contribution to the Linnean theme, images and fragments on the House Fly , but have preferred to wait with the image description until there is a possibility to do so in the REGNET system.

2.3 KVA

KVA has devoted two man months to the REGNET project during the accounted time period. Besides the actual work with the content KVA attended the meeting in Barcelona. Four meetings have been held to coordinate the work on the Linnaeus theme with SUL since our work has been closely linked to their work.

Cataloguing and Digitizing

The scientific instrument collection database (#3.500 entries) has been exported to AIT as well as b&w digital images in low-resolution JPEG-format (#1.716) of the instruments. In the previous period a portrait collection was digitized in high-resolution TIFFF-format (#800). During this period an accompanying database has been produced. This database will be added later, since the information on the portrayed persons is not fully completed.

Thematic Content Production

KVA has worked together with SUL in producing the thematic texts on the Linnaeus theme. Thematic texts have been researched, prepared and written on the Linnaeus theme and on the measurement theme. Accompanying illustrations have been selected and scanned. The measurement theme texts (#14 fragments and #59 object/image descriptions) as well as digital images of the instruments have been sent to MUSEON who coordinates this theme. The Linnaeus texts are accounted for in the SUL report.

During this period Ph.D. student M.A. Christer Karlsson has been employed to do the research and the writing of fragments.

2.4 NRM

The Museum's collections include millions of specimens, and their size and quality place the Museum as one of the leading museums in Natural History. The collections provide the basis for the research carried out in the scientific departments and through loans and visits by guest researchers are constantly used by scientists and institutions throughout the world. It is a crucial role of the Museum to keep these collections available for international research, and to preserve them for future generations. Parts of the collections presented on the Museum's web pages are also registered in searchable databases. The museum collection is represented in botany, geology, paleontology, zoology.

The Museumshop

For the REGNET project the shop was chosen as a museum partner as an excellent opportunity to market the shop and take part in e-commerce. There are about 2000 objects in the shop. The assortment is based on animals, nature, human beings and space. The shop sells books, both for children and adults, some dissertations, textiles, fashion and accessoires, stones and minerals,

fossils, CDs, CD-ROM, videos, posters, postcards, calendars, pens, paperworks, t-shirts, ties, hats, umbrellas, toys, key-rings, puzzles, games, space-patches, material to press plants etc.

A selection of 113 objects from the shop have been chosen for the REGNET project. The selection of objects was done according to commercial requirements, that is to say best-selling articles.

Cataloguing and Digitizing

The objects in the shop have been photographed and digitized externally by a commercial photograph studio. The images are thus at a high standard, in the TIFFF-format on CD-ROM. The cataloguing of the objects has been done by museum staff according to Dublin Core standard in an Access database.

Personnel: The photographing and digitizing was left to a sub-contractor as the photographers studio in the museum had an immense work-overload and it was clear that the work would not be finished in time for the requirements for the REGNET project. Cataloguing was done by permanent staff in the museum 1.5 mm non cost

2.5 LMG

The Historical Museum of Gotland is one of the largest museums, attracting visitors from all over Sweden and abroad. Here you will meet 8000 years of the islands rich history – from the Stone Age to the Viking Age, from the Medieval Period to present day.

The merchandise in the Museum shop reflects the collections of the museum and the interesting nature, handicraft and art of the Gotland region. The shop has chosen some 110 objects for the REGNET project to present to customers world-wide. The objects have been photographed and digitized in JPEG-format. Records describing the objects are in Swedish and English in Superbase, compatible with Filemaker Pro, Excel and Access.

Thematic Content Production

LMG has provided input to the Saints theme. The Saints presented are images of medieval wooden sculptures. At present the Saints database comprises of 58 objects with pictures ready for use. The records describing the saints are in Swedish and will be translated into English.

2.6 GRAN

The kind of objects that make up the INcívics' catalogue are works of contemporary art. They are made for more than 30 artists in the field of visual arts (painting, sculpture, engrave, photography, video creation, net art ...). It is a contemporary collection, of emergent and young artists. They are living in the Granollers area. These artists have done some exhibitions in institutions and spaces from the city, like the municipal museum, the civic centers, and galleries. On the other side, this collection does not have a logical selection of art works, as for an academic purpose. We can find very different techniques involved in our way to be contemporaneus. At the moment we have 81 works, that means 104 pictures from 16 artists.

Cataloguing and Digitizing

The descriptions and images are available: all the pictures are digitized, almost all the thematic texts (curriculum vitaes of the artists), object description and the text from the artists speaking about their own job). For the REGNET demonstration system a first selection of the artists with their own files, this includes the images files and the thematic texts files are delivered. The data structures used are in accordance to the (offline data entry) templates and thereby could easily be converted into XML.

2.7 MECH

The participation of the municipal museums of Mechelen in the REGNET-project fits in with the general working of museums as it is described in the definition of ICOM. Especially for the basic functions of registration and unlocking of collections, the REGNET project is offering a lot of new

perspectives. Computerising and doing a scientific inventory belongs to the museological program. The call for enhanced physical presentation of museum-objects as well as the call for the virtual presentation of them becomes very compelling. At a time in which a lot of official and governmental authorities are developing digitized cultural databases in an active way the Mechlin museums want to anticipate. Via the participation in REGNET they want to develop new ways and tap new sources of public to present and distribute the objects and the scientific work. Besides, the participation in a european project and the international collaboration with regional museums is a special trump card for the official national recognition of the museum of Mechelen.

Cataloguing and Digitizing

The registration, classification and description of the collection are undoubtedly necessary, not only for the science, but in view of further developments characterized by substantial added value. It is impossible to manage the collection, to present and situate the objects, to make the very fascinating story of the city without a good knowledge and description of the different witnessing elements. The museums of Mechelen combine and synchronise in their workings the objectives of the REGNET-project with the inventory of the collection in the ADLIB-management system. On this moment a great part of the collection registration exists on paperbased catalogue cards. OCR tests are realized on this paper based catalogue and will be continued in collaboration with AIT and TARX. The first steps for computerized registration are set now. About 140 objects have got their description in the inventory. At this moment the ADLIB collection management system was upgraded in order to do exporting the catalogue data in XML-format. One hundred of these catalogue-records were delivered to REGNET in the XML-format. The basic registration is done by the administrative staff. This day to day registration and cataloguing is directly linked with REGNET-objectives. The activities related to REGNET means much more than registration. They are situated on different domains. First of all time-planning and guidelines are set up for the digitizing of the respective collections. In order to continue digitizing internally a study for the acquisition of a digitizing infrastructure was carried out and an appropriate system was ordered.

Thematic Content Production

Further on, new thematic concepts are worked out for the production and consultation of cultural content. The images and objects of the collection, addressed following the principle of themes and sub-themes receive at the same time with their description a scientific and automated registration, including the scanning of digital images. Mechelen has the management of the content production for the REGNET-themes saints and guilt-leather and contributes to the themes Habsburg, Tour d'horizon, and mearurering. A contribution to the museum-shopping part is envisaged. The exact items for the shop have to be defined in the near future. A forum with initiating texts was started on the REGNET web site in order to elicit the production of thematic texts for the themes saints and guilt-leather. The production of more than 50 thematic texts was carried out in Dutch concerning the SAINTS theme. Synchronisation efforts took place between MECH, GOTL, ALI and Argenti for thematic contributions concerning the saints theme.

In close collaboration with the Municipal Archives of Mechelen, 6 thematic texts were contributed to the Botany-Linnaeus theme. It concerned the Mechlin botanist Rembertus Dodoens.

2.8 MUS

Cataloguing and Digitizing

A working practice has been established: a team of 4 people are working on the development of the themes and different types of texts. Two persons are involved in logistics and photographing the objects. The digitized photographs are registrated by two employees and added to the collection management system.

Thematic Content Production

The MUSEON has worked on the themes: Leather, Measuring (time), Masks & Amulets. *Leather*: For this theme all necessary objects have been selected. All text fragments and theme fragments are ready (in Dutch only). The objects are now being photographed and digitized soon. Contributions from other partners are still welcome. *Measuring*: For this theme all necessary objects have been selected. All text fragments and theme fragments are ready (in Dutch and partly in English). The objects are

photographed and digitized. Contributions from KVA have been received, other contributions are still welcome. *Masks & Amulets:* For this theme most necessary objects have been selected. Most text fragments and theme fragments are ready (in Dutch only). Most objects are photographed and digitized. Contributions from other partners are still welcome.

MUSEON contributed several texts to the Linnaeus Theme. Regarding the Tour d'horizon a short list of objects is compiled. Due to the great variety of the collections great care is taken in selecting the objects. The object texts will be ready soon. All the objects on the short list are photographed and digitized.

Museumshop items

A contribution to the shop the museum part is being prepared. The exact items for the shop will be defined in the near future.

2.9 ICCS

The Union of Bulgarian Artists gathers artists from all kind of visual arts, grouped in 16 specialized sections: painting, graphic arts and illustration, sculpture, criticism, caricature, artistic spatial design, stage design, restoration, monumental arts, design, ceramics, graphic design, "13" - new forms of expression, wood-carving, artistic processing of metals and leather, textiles.

Cataloguing and Digitizing

For the object description a form has been developed. It has been passed to the Union of Bulgarian Artists management body for distribution and processing among their members. These forms have provided data about art objects. Having data about art objects, it has been used for cataloguing the objects according to Dublin Core standard. The art objects were digitized related to the filled forms of the providers – individuals and artists.

Thematic Content Production

With regard to the project the main focus of thematic content production was the development of Art Gallery topics: Preparation of thematic structure & fragments (Domain Artists) and the creation of predefined topics maps. For the Tour d'horizon a description of the 11 art categories were made in Bulgarian and English: painting, sculpture, graphic and illustration, design, wood-carving, textiles, ceramics, monumental art, new forms of expression, restoration, Icons. Samples of the art objects per 11 art categories are available in two demo versions. *Demo 1*: XML Dublin Core based data are worked out . Dublin Core standard is implemented as XML records on XML attribute level. *Demo 2:* XML data based on Dublin Core are worked out. The Dublin Core standard is implemented as XML records on XML element level. Both Demo versions are presented according to XSL files, worked out for demonstration purposes : www3.iccs.bas.bg (ICCS) and www.regnet.org.ru (SUSU-Russia). Search and retrieval functionality is supported according to Z39.50 standard. Demo solutions are available for the XML records and appropriate samples of XSL.

3 Collections for the demonstration phase

3.1 New partners in the project

During the 2nd phase of the project two new content partners could be integrated in the consortium. In the following an institutional profile will be given together with a short description of the collection. The work sheets(audit and digitizing plan) related to WP 1.1 and WP 1.2 for the partners (if available) could be found in Appendix 1.

3.1.1 Museo degli Argenti (ARG), Florence

The *Museo degli Argenti* or *Silverware Museum* is located in Pitti Palace, the majestic palace-fortress in the middle of Florence bought by Cosimo I Medici (1519-1574) from the Pitti family. At that time the building consisted only of the central portion, but during the following Centuries the Palace was the residence of the Grand Ducal Court and it continued to expand. With architectonic and decorative

adaptations, faithful to the taste of the successive princes: Medici, Lorraine and Savoy. Two wings were added around the courtyard in the XVIIth Century: the left one, the original *Appartamento d'Estate (Summer Apartment)* of the Grand Dukes, is now the seat of the Museum. The rooms overlooking the square of Pitti Square have always been used as hearing rooms and they contain magnificent frescoes. The decoration of the main room started during the half of the XVIIth Century by Giovanni da San Giovanni and his pupils Cecco Bravo, Ottavio Vannini e Francesco Furini, celebrating the patronage of the Medici under Lorenzo the Magnificent. The eclecticism of the Grand Ducal Court can be observed in the taste for decoration, such as the rare marble and hard-stones mosaics of the tables in the rooms of the Museum, or in the passion for collecting precious objects like: the Reliquiary attributed to a Parisian gold-smith after the design of Etienne Delaune, the vases from Lorenzo, the crystal from Francesco I, the ambers from Maria Magdalena from Austria and the ivory form Mattias Medici. During the half of the XIXth Century, the Grand Dukes left the residence and the stately halls were adapted for offices, with makeshift partitions. Only the room of Giovanni da San Giovanni remained the Silver Cabinet, were silver works from Salzburg and Würzburg were exhibited.

The *Museo degli Argenti* was officially created in the 1919, mostly with objects coming from the Bargello. In the museum various kinds of objects such as gems, cameos, precious stones, ivories, jewels and silver are gathered (about 5000 objects). Its lack of unity reflects the diversity of tastes, fashion and working methods of all the centuries of collecting which went to make it up. It is planned to introduce a database for cataloguing and in order to manage loans, restorations and to support the research. Actually descriptions exist in MS Word. The digitizing of the objects is just started, about 10% of the objects are already digitized (medium resolution, JPEG).

3.1.2 Capitoline Museums (CAP), Rome

The Capitoline Museums are the most ancient public collection in modern world. Their origin dates back to the donation by Pope Sixtus IV (1471), to the Roman People, of a set of classic bronzes - among others the Capitoline She-Wolf and the Spinario - previously kept in Lateran. Besides those bronzes, the fragments of the Acrolithic statue of Constantine , the Brutus Capitolinus and many other works now exhibited in the Palazzo dei Conservatori, show the original historic nature of the Capitoline Museums' collections. Only towards the second half of the 16th century, thanks to Pope Pius V, it was possible to acquire works artistically significant. Towards the end of the 17th century more and more important donations were preserved even in the Palazzo Nuovo, which was officially reserved for museum-purposes in 1733. That was the case of various sculptures, such as those from Albani's collection -busts of Roman emperors and philosophers, among others the Portrait of Augustus - the Capitoline Gaul , the Capitoline Venus and the Mosaic of the Doves. Those works of art and many others often came from ancient Roman families of noble origin and were often new elaboration of works from ancient Greece. In 1749 Capitoline Picture Gallery was open on the inside of the palazzo dei Conservatori. The Gallery originated from the union of two private collections (Sacchetti and Pio di Savoia) containing paintings from different schools and periods (from 14th to the 17th centuries). Tempera wooden panel paintings from the Middle Age to the Renaissance, Emilian (Garofalo, Dosso Dossi) and Venetian (Tiziano, Lorenzo Lotto, Veronese) school paintings from 16th century, Bolognese school painting from the 17th century (Guido Reni), can at present be admired; as well as Rome painting from the 17th century (Guercino, Caravaggio), baroque painting (Pietro da Cortona) and many others.

3.2 Available collections for the demonstration phase

The main goal of WP 2.1 was to provide sample and proved data in order to be able to describe the products and services at the end of this task. Table 2 shows the actual status of available data (14 April 2002). This data was processed (analysed, mapped to domain-specific standards and converted) in order to be integrated in the REGNET system (see chapter 4).

Partner Name	***Name of database***	***Format***
AIT	DMB (internal test database)	XML

AIT	Swiss Postercollection	XML
ONB	Picture archive - catalogue cards (including image)	Access TIFF, JPEG
ONB	Primisser (theme)	Excel JPEG
ONB	Theme (Elisabeth)	HTML JPEG
ORF	ARIS-HA (database for photos, films, sound and books)	Access
GRAN	object descriptions (art works), cv of artists	Word JPEG
SUL	Zool (bib. objects)	Marc/XML
SUL	Berg-R-Rf-Linne	Marc/XML
SUL	themes	Word
SUL	image related to themes	JPEG
KVA	Adamx	Access
LMG	LMG Saints	Access
LMG	(Museum) Shop1	Access (incl. 104 JPEG images)
LMG	Prislist	Access
NRM	Museum Shop images	TIFF
MECH	Mechelen object data	XML
MECH	(saints) Theme	Word
TARX	Theme: image metadata description, catalogue description (English and dutch) 4 object labels, 4 St. Angela themes	Excel Word
TARX	Images (related to theme)	JPEG
MUS	Themes (time, leather)	Word
ICCS	Themes	Word XML; DTD, XSL available
IMAC	Feldbrunnen (object data)	Excel
ICCS	Images related to themes	JPEG
SUSU	Themes	Word
SUSU	Object descriptions	Word
ALI	Themes in no special order AND these themes in thematic order(saints, travel, botanic, Musei Capitolini, Museo degli Argenti)	Word (2 XML)
ARG	Themes	Word

Table 2: Available data / databases (14 April 2002)

Currently the two AIT databases, the Swiss poster collection and the DMB internal test database, are transformed into REGNET data structure. The two databases are accessible and searchable through the portal Search & Retrieval. At present they contain mainly object data, Dublin Core records and person / organization information. TARX developed an off line data entry system, containing 4

templates: catalogue, object, image and thematic description. They contain respectively the metadata about physical catalogued objects, exhaustive textual descriptions about objects, digital images about objects and contextual thematic descriptions around objects. TARX developed also a spreadsheet to XML conversion program to put everything in XML. At The moment 4 partners are contributing in this way: MECH (for the thematic texts only, the catalogue records are exported from ADLIB and delivered already), TARX (for all sheets), GRAN (for all sheets) and MUS (for the description of the templates see chapter 5.2).

4 Preparation and conversion of data

4.1 The REGNET Data Structure

Basically we have three types of data in the system:

- PRIMARYDATA
  refers to repository metadata, data about persons/institutions, events, surrogates of real objects, ...
- SECONDARYDATA
  additional data for the system; data for processing PRIMARYDATA: e.g. stylesheets, thesauri, topicmaps, system configuration, ...
- PROCESSDATA
  workflow related process data: e.g. steps for object aquisition, ...

PRIMARYDATA is manly stored in the Cultural Heritage Data Management Node or eBusiness Data Management Node (e.g. PERSON's will also be in the Ontology node as 'users'). SECONDARYDATA and PROCESSDATA is stored in the Ontology System.

The XML documents in the system will all have the same basic structure:

```
<?xml version="1.0"?>
<regnet-document VERSION="1.0">
   <section name="-doctype-">
      <doctype>-doctype-</doctype>
      ...
      (data related to this document type)
      ...
   </section>
   <section name="-doctype-">
      <doctype>-doctype-</doctype>
      ...
      (data related to this document type)
      ...
   </section>
     ... (other sections)
</regnet-document>
```

Document types (= the -doctype- placeholder) will be a main concept of the whole data structure. We will now go into detail with this document types. For a basic schema of a REGNET document see Appendix 2.

4.2 Document Types

4.2.1 Document Types for PRIMARYDATA

When describing real objects we first have the 'concept' of the object. This concept is then usually mapped to a specific domain. Here you can see an example:

'Concept'	*Picture*	*Object*	*Bibliographic*	*e-Business*
Person	Photographer	Artist	Author	Supplier
Object	Picture	Peace of Art	Book	Item
Place	Creation Place	Location	Location	Address
Process	take photo	present	loan	ship
...				

Table 3: Example for the mapping of document types

This leads us to our document types (see also example below!).

- PERSON — data related to persons (they are further qualyfied e.g. author, owner, , ...)
- EVENT — events concerning the object
- OBJECT — object data
- SURROGAT — object surrogat data
- BIBILIOGRAPHIC — bibliographic data of the object
- INSTITUTION
- PLACE — creation place, current place, ...
- DC — Dublin Core representation of the object
- EBUSINESS — CATXML related data to this object
- ...

For PRIMARYDATA we have to perform conversions from the original data format into the REGNET data structure and back. Objects represented in exchange data formats will be transformed into an application data structure. The set of application data can/will contain further (application related) data about an object. For illustration see the example below.

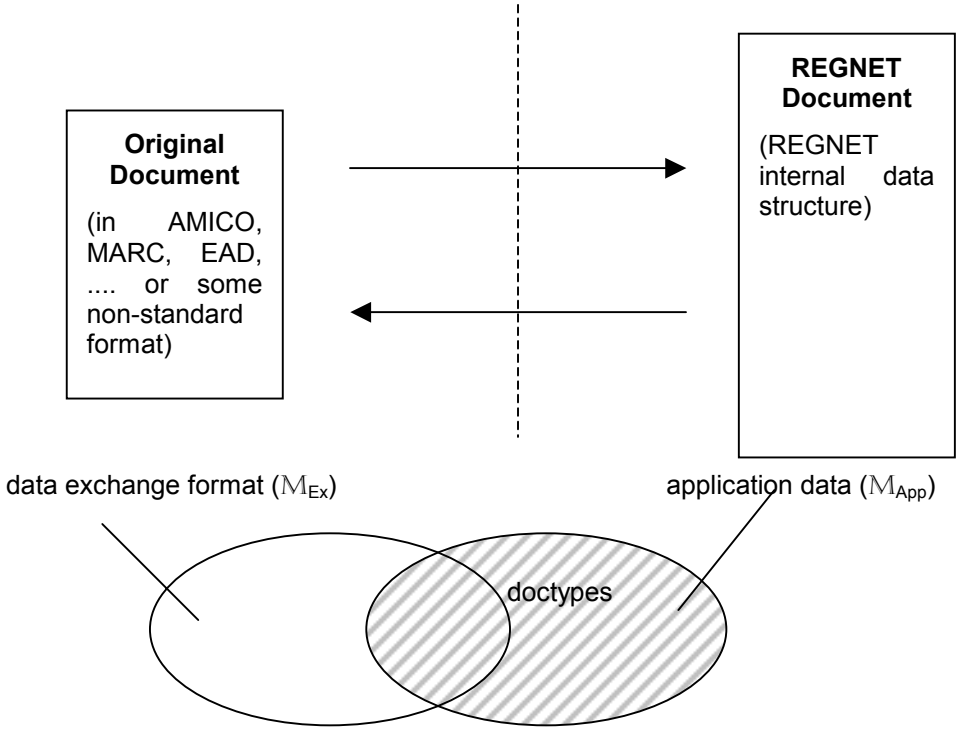


Figure 2: Conversion process into the REGNET data structure

Example 1:

Original Document: AMICO Record from the Swiss Poster Collection

```
<?xml version="1.0" encoding="UTF-8"?>
<am_object>
   <AID__amico_identifier>8</AID__amico_identifier>
   <am_titles>
      <OTG_object_title_name>
         <OTN__object_title_name>Burger King Gaststätte / Eindruck: Kränzlin am
                                 Marktplatz</OTN__object_title_name>
      </OTG_object_title_name>
   </am_titles>
   <am_appearance>
      <OPD__physical_description/>
      <OMG_materials_and_techniques>
         <OMD__materials_and_techniques_description/>
      </OMG_materials_and_techniques>
   </am_appearance>
   <am_creators>
      <CRG_creator>
         <CRT__creator_name_text>(G. Brauchle, Thal SG)</CRT__creator_name_text>
      </CRG_creator>
   </am_creators>

   <am_creation_dates>
      <OCG_creation_dates>
         <OCT__creation_date_text>1982 / 71</OCT__creation_date_text>
      </OCG_creation_dates>
   </am_creation_dates>
   <am_creation_place>
      <OCP__creation_place/>
   </am_creation_place>
   <am_owners>
      <OOG_owner>
         <OON__owner_name>Interdistra AG Vaduz</OON__owner_name>
         <OOC__owner_credit_line/>
      </OOG_owner>
      <ORG_rights_copyright>
         <ORS__copyright_statement>
            (Eidenbenz &amp;amp; Co. St. Gallen)
         </ORS__copyright_statement>
      </ORG_rights_copyright>
   </am_owners>
   <am_media_metadata>
      <XRE_dc_relation>
         <XRY__dc_relation_type>hasVersion</XRY__dc_relation_type>
         <XRI__dc_relation_identifier>h
            http://www.csceurope.org/plakat/cd00.050/t_s_dsc05579.jpg
         </XRI__dc_relation_identifier>
      </XRE_dc_relation>
      <XRE_dc_relation>
         <XRY__dc_relation_type>isVersionOf</XRY__dc_relation_type>
         <XRI__dc_relation_identifier>
            http://www.csceurope.org/plakat/cd00.010/s_t_s_dsc05579.jpg
         </XRI__dc_relation_identifier>
      </XRE_dc_relation>
      <XRS__dc_rights/>
   </am_media_metadata>
</am_object>
```

REGNET Document:

Here you can see how the resulting REGNET document may look like. As mentioned before: the real transformations have to be developed. Also DTDs / SCHEMAs for each document type must be provided.

```
<?xml version="1.0" encoding="UTF-8"?>
<regnet-document VERSION="1.0">
   <section name="PERSON">
      <doctype>PERSON</doctype>
      <role>creator</role>
      <name>G. Brauchle</name>
   </section>
   <section name="PERSON">
      <doctype>PERSON</doctype>
      <role>owner</role>
      <name>Interdistra AG Vaduz</name>
   </section>
   <section name="OBJECT">
      <doctype>OBJECT</doctype>
      <resourcetype>IMAGE</resourcetype>
      <csc>CSC-AT</csc>
      <ooa>05579</ooa>
      <crt>(G. Brauchle, Thal SG)</crt>
      <otn>Burger King Gaststätte / Eindruck: Kränzlin am Marktplatz</otn>
      <oct>1982 / 71</oct>
      <oon>Interdistra AG Vaduz</oon>
      <ors>(Eidenbenz &amp;amp; Co. St. Gallen)</ors>
      <xri>http://csc002.cscaustria.at/plakat/cd00.050/t_s_dsc05579.jpg</xri>
      <xri>http://csc002.cscaustria.at/plakat/cd00.010/s_t_s_dsc05579.jpg</xri>
   </section>

   <section name="DC">
      <doctype>DC</doctype>
      <title>Burger King Gaststätte / Eindruck: Kränzlin am Marktplatz</title>
      <creator>(G. Brauchle, Thal SG)</creator>
      <subject/>
      <description/>
      <publisher>Interdistra AG Vaduz</publisher>
      <contributor/>
      <date>1982</date>
      <type>image</type>
      <format/>
      <source/>
      <language>de</language>
      <relation>http://csc002.cscaustria.at/plakat/cd00.050/t_s_dsc05579.jpg
      </relation>
      <relation>http://csc002.cscaustria.at/plakat/cd00.010/s_t_s_dsc05579.jpg
      </relation>
      <coverage/>
      <rights>(Eidenbenz &amp;amp; Co. St. Gallen)</rights>
   </section>
   <section name="EBUSINESS">
      ...
         CatXML information related to this object
      ...
   </section>
</regnet-document>
```

4.2.2 Document Types for SECONDARYDATA

SECONDARYDATA comprises all kind of additional data for supporting the system. The document types are:

- THESAURUS
- TOPICMAP
- LOOKUPTABLE e.g. country codes, currency codes, ...
- DTD
- SCHEMA
- STYLESHEET
- INTERFACE WSDL definitions
- QUERYTYPE definition of query formats
- CONFIGURATION
- DATABASE description of a database + additional info about content
- PERSON the same person structure as in PRIMARYDATA but other roles (e.g. user, admin, ...

4.2.3 Document Types for PROCESSDATA

In PROCESSDATA documents workflows will be described. It can be used to control and monitor the system.

- It is an new approach: proposal
- Reflects all supported business functions (from D3 – Part 2)
- Reflects application processes

Document Types for PROCESSDATA are:

- RETRIEVAL
- PORTAL
- EBUSINESS
- PUBLISHING
- ...

Standards for modeling workflows and processes were published from the Workflow Management Coalition (WfMC) (see. http://www.wfmc.org/standards/docs.htm).

- The Process Definition Interchange Interface includes a common meta model for describing process definitions.
- An XML-Process Definition Language is already defined through and DTD.

Here you can see an XPDL.DTD snippet:

```
...
<!ELEMENT ParticipantType EMPTY>
<!ATTLIST ParticipantType
   Type (RESOURCE_SET | RESOURCE | ROLE | ORGANIZATIONAL_UNIT | HUMAN | SYSTEM)
   #REQUIRED>
...
```

- Definition of ParticipantTypes to a process.

- what do we have to define: Resources, Roles, Organisational units, ...

Such a workflow definition should be done for all components in the REGNET System (e.g. for Publishing Component).

4.3 REGNET DTD

The purpose of a Document Type Definition is to define the legal building blocks of an XML document. It defines the document structure with a list of legal elements. DTDs are typical secondary REGNET data. To validate REGNET data a DTD for each (object-) database and for each Document type has to be defined.

Example: the News DTD (news to fill in the News section of the Portal):

<!ELEMENT RN_News (date_of issuing, date_of_withdrawing, news_date, news_title, news_text, subject, news_link)>

<!ELEMENT date_of_issuing (#PCDATA)>

<!ELEMENT date_of_withdrawing (#PCDATA)>

<!ELEMENT news_date (#PCDATA)>

<!ELEMENT news_title (#PCDATA)>

<!ELEMENT news_text (#PCDATA)>

<!ELEMENT subject (#PCDATA)>

<!ELEMENT news_link (#PCDATA)>

4.4 Data Conversion

For transformation from 'original data' into 'REGNET data' we will have to:

- define DTDs/SCHEMAs for all document types in REGNET. They will be the targets for the transformations.
- define transformations (generally for exchange metadata formats).
- check general transformations against provided collections and extend the transformation if necessary.
- provide a set of tools for the transformation process, based on the original data storage format (e.g. plain text, database, xml-formated, …) and available platforms.

4.4.1 Description of data conversion

Content providers prepare data that will be imported into the REGNET Cultural Heritage Data Management in various formats: MS Access, Marc, Word, Excel, XML, etc.

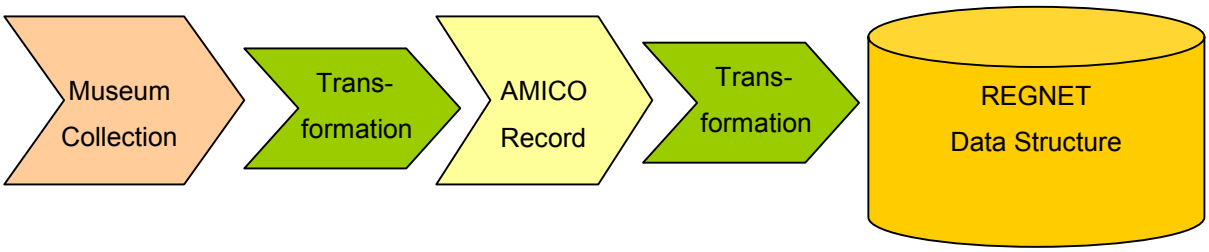


Figure 3: The conversion process

Files in non-XML-format have to be transformed into XML.

1. Transformation from existing databases: export Tools provided by database products and JAVA applicataions (e.g. jXTransform from DataDirect®) accessing the database via JDBC are used. With jXTransform you can specify SQL queries and transform the output into XML structures.
2. Transformation from text, Word, Excel files: simple JAVA applications and VB scripts are used to transform formatted files into XML structures.

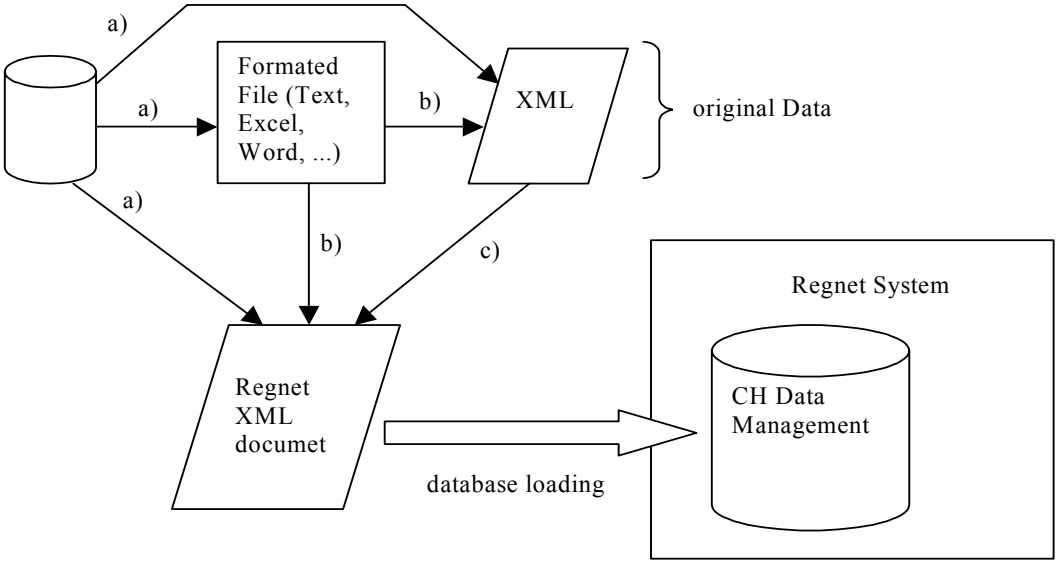


Figure 4: Data Transformations

Mapping of data fields to Dublin Core and REGNET data structure: Original data is mapped to Amico data structure and Amico data is mapped to Dublin Core and REGNET data structure. A first version of a crosswalk between different exchange data formats has already been developed (see the crosswalks in the Appendix 3). The following Excel sheet demonstrates exactly the procedures for mapping and transforming the Swiss poster collection that are further on defined in XSLT and performed. The files produced by this transformations will be imported to the REGNET CH Data Management system.

Note: there are data fields containing persons as well as institutions and it is not possible to decide automatically which name refers to a person and which name refers to an institution. For this reason the section NAME was created.

Poster Collection			Object (Amico)				Dublin Core		
Pos	TagName	Note	Pos	Tag Name	Hint	Value	Pos	TagName	Value
1	ID		1	ooa	owner accession number	1	1	Title	3, 15
2	Entwurf	remove-brackets	2	crt	creator name text	2	2	Creator	2
3	Titel	remove-brackets	3	crq	creator qualifier	Designer	3	Subject	
4	Jahr		4	crt	creator name text	6	4	Description	
5	Auftraggeber	remove-brackets	5	crq	creator qualifier	Printer	5	Publisher	Plakatsammlung Basel
6	Drucker	remove-brackets	6	crt	creator name text	14	6	Contributor	6, 14
7	Technik	remove-brackets	7	crq	creator qualifier	Photographer	7	Date	4
8	Format		8	otn	object title name	3, 15	8	Type	Image
9	URL_Scan		9	oct	creation date text	4	9	Format	
10	URL_Bild_klein		10	ocs	creation date start	4 (4 Zeichen)	10	Identifier	
11	URL_Bild_gross		11	cxp	context related person	5	11	Source	
12	Datum		12	omd	material technique	7	12	Language	de (/ fr)
13	Ort	remove-brackets	13	med	measurement_dimension	8	13	Relation	11,10,9
14	Foto	remove-brackets	14	xri	relation	11	14	Coverage	
Poster Collection			**Object (Amico)**				**Dublin Core**		
Pos	**TagName**	**Note**	**Pos**	**Tag**	**Hint**	**Value**	**Pos**	**TagName**	**Value**

				Name					
15	Titel2	remove-brackets	15	xri	relation	10	15	Rights	
16	iNummer		16	xri	relation	9			
			17	cxs	context related site place	13			
			18	dcd	documented catalogue date	12			
			NAME						
			(xNL?)						
			Pos	**TagName**	**Hint**	**Value**			
			1	nte	name text	2, 5, 6, 14			
			2	nqu	name qualifier	General (, Person, Organisation)			
			3	nro	name role	Designer, Orderer, Printer, Photographer			
			PLACE						
			?						
			Pos	**TagName**	**Hint**	**Value**			
			1	pro	place role	Place of usage			
			2	pte	place text	13			

Table 4: Procedures for mapping and transforming of the Swiss Postercollection

5 Data Entry Tools

5.1 Online Data Entry

5.1.1 General description

The present REGNET data input tool with TeXtML Server at its core provides access for end users to the databases through common user interfaces in the internet. (e.g. http://www.digipark.at/plakat)

Three basic functionalities are available:

1. **Search and Edit**: the user defines the search criteria (all document types or one special document type chooses a search term from a list, and/or defines further search criteria).
2. **Add a new record**: the user clicks the “add a document” button and chooses a document type. Then the system displays the data entry mask for this document type and the user adds data. By clicking “apply modifications” the new abstract/record will be added to the database.
3. **Delete a record**: the user performs any search, displays a record from the search result be and clicks the “remove” button. The system will ask for confirmation and then deletes the record.

Data can be transformed and imported as a batch by the administrator in following formats:

- XML our preferred format (http://www.w3c.org/XML)
- MS Access
- MS Excel (for automatic transformation from Excel to XML the data has to be formatted as described hereunder (see “A simple example …)).
- CSV (comma separated values - the lowest common denominator text format is easily parsed and often adequate, particularly for “flat” or “tabular” data).

This proceeding is mainly used for exchange of data with existing systems.

A simple example for preparing data for importing into REGNET System:

An Excel sheet is generated by the content provider: the names of the data fields have to be in the first row (not in the first column!):

PK	*Nr.*	*First Name*	*(Sur)name*	*Title*
1	1	Rudolph I		Römischer König
2		Anna	von Hohenberg	
3		Agnes		Herzogin von Burgund
4	2	Albrecht I		Römischer König
5		Elisabeth		Gräfin von Tirol und Goerz
6	3	Hartmann		
7		Johanna		Prinzessin von England
8	4	Rudolph II		
9		Agnes	von Böhmen	
10	5	Mechtilde		
11		Ludwig	der Strenge	Herzog von Bayern

Table 5: Example from ONB (Clipping from the Primisser contribution)

Starting from an Excel sheet in the format described above a macro can convert data into appropriate XML files. Further steps of data conversion and importing are described in the chapter data conversion. For the actual version of the guidelines see Appendix 4.

5.1.2 Example: Data Entry for the REGNET Portal

In order to manage and collect data that will be added to the REGNET Portal a data entry facility is provided. Regarding the contents of this data entry tool it is mainly a portal related tool, but regarding its functionality it is a good example for the application in general (http://www.digipark.at/RNDataInput). At present there is one user defined (username: regnet, password: member).

REGNET partners will use this tool to enter relevant information for the portal. Data is stored internally in TeXtML Server and regularly exported automatically to the concerning areas in the portal. At present there are seven document types defined (according to our considerations during the Graz meeting). Data fields that are one group are in a separate paragraph.

News	***(Contribution to a) Newsletter***	***Event***
Date of issuing	Name of Newsletter	Date of issuing
Date of withdrawing	Issue	Date of withdrawing
News date	Author	Event Date(s)
News title	Title	Event Title
News text	Newsletter text	Description
Language	Language	Tickets
URI	URI	Language
Subject	Subject	Venue
	Newsletter (contribution) Category	URI
	Image	E-mail
		Contact Name
		Contact Address
		Phone
		Image
		Event type
Link	***Address***	***(Institution) Profile***
URI	URI	URI
Link Title	Person Title	Institution name
Description	First name	Description
Language	Last name	Admission
Link category	Organisation	Open
	Language	Language
	PO Box	Open
	Street	Contact Name
	Postal Code	Contact Address
	Place	Phone
	Country	E-mail
	Phone	image
	Fax	
	E-mail	
Collection		
Collection name		
Description		

Language Collection type		

Table 6: Document types and data fields

Creator of the record and date of creation are recorded internally.

5.1.3 Multilinguality

To manage the problem of multilingual contributions we have grouped the data fields that could be in various languages:

E.g.: doctype EVENT:

Date(s), venue, uri, contact information and event type will not change if the language changes. Event title, description, language code and the short description of availability of tickets have to be translated. The user duplicates this data field group and adds a translation (he can also add a third version or a fourth etc.).

5.2 Tools for the offline data entry

5.2.1 Introduction

The aim of the development of an off line data entry system for the REGNET-system is threefold.

1. Collection of content examples at an early stage of the project (until the availability of on line tools) and introduction of new parameters corresponding with the REGNET-objectives related to thematic descriptions and e-Business.
2. Possibility to set up a kind of collection management for those partners not disposing of a standard collection management system.
3. Availability of data input facilities for those not possessing or not willing to use, on line connections.

5.2.2 First version

In order to collect in an early phase of the project examples of the content that the content providers of REGNET intended to deliver an off line data entry system was developed based on a template within a word processor.

The template contained a subset of the fields of a standard collection management system supplemented with those fields required to fulfill the objectives of REGNET. This concerns mainly the possibility to create exhaustive texts containing thematic descriptions. The latter have only indirect relations to catalogued objects in contradiction with catalogue and object descriptions, which pertain directly to the objects.

The following templates reflect the first approach.

Thematic Description (Fragment)	
Thematic description name	title
Language	en for English, nl for Dutch, …
Audience level	generic, children, students, scientific, …
Size level	short or full
Author	
Contribution	

Date	
Modified by	
Modif. date	
Modif. description	
Copyright	
Cost	
Content	Text with embedded references A reference is a number: e.g. (1) = reference to link 1
Links	(1): description of an action and/or URL of resource referenced by (1) (2): description of an action and/or URL of resource referenced by (2) …..
Primary keywords	most relevant
Secondary keywords	second order
Number	digital object identifier

Table 7: Template for thematic descriptions

Object Image (fragment)	
Object image name	title
Language	en for English, nl for Dutch, …
Audience level	generic, children, students, scientific, …
Size level	dimension, resolution
Author	the person who made this
Contribution	
Date	
Modified by	
Modif. date	
Modif. description	
Copyright	
Cost	
Content	Text with embedded references A reference is a number: e.g. (1) = reference to link 1
Links	(1): description of an action and/or URL of resource referenced by (1) (2): description of an action and/or URL of resource referenced by (2) …..

Primary keywords	most relevant
Secondary keywords	second order
Number	digital object identifier

Table 8: Template for object images

Object Description	
Object description	title
Language	en for English, nl for Dutch, …
Audience level	generic, children, students, scientific, …
Size level	short or full
Author	
Contribution	
Date	
Modified by	
Modif. date	
Modif. description	
Copyright	
Cost	
Content	Text with embedded references A reference is a number: e.g. (1) = reference to link 1 Metadata : Dublin Core, AMICO standard, ….. In case of sellable object: Additional item of “Price”
Links	(1): description of an action and/or URL of resource referenced by (1) (2): description of an action and/or URL of resource referenced by (2) …..
Primary keywords	most relevant
Secondary keywords	second order
Number	digital object identifier

Table 9: Template for object descriptions

5.2.3 Second version

The experience with the first version was positive with regards to the collected content but provoked some difficulties for automating the conversion from the word processor template towards the internal data structures of REGNET. Therefore it was decided to develop a second version for off line data entry based on a more structured and record oriented model. It became the spreadsheet model, carried out in 2 phases.

5.2.3.1 Phase 1

Scope: This document contains the description of the new version of the guidelines and tools for the data entry of catalogue, object, image and thematic descriptions together with a file naming convention as adopted by TARX and MECH for the storage of the produced files. The latter can be taken as example or inspiration for all content providers producing content in file structures. In view of the availability of a full-blown data entry tool, these templates are still an intermediate version but their structured nature is such that, once their content is captured by the system, neither retyping nor copy/paste has to be performed. The REGNET system can take care of handling most of the changes in the future. These templates are in the first place intended for these content providers not possessing a Collection Management System (CMS), for the types of data not addressed by CMSs or other systems (e.g. the thematic descriptions and the multilingual and multilevel issues). This data entry facility contains all necessary fields required for a proper operation within REGNET. The data entry fields will be mapped directly into the REGNET data structures. The content providers possessing a CMS will normally export the concerned data for import in REGNET where a conversion will take place into the REGNET data structures. Once the full-blown data entry system is available, supplementary data, not foreseen in the CMS functions but required within REGNET, can be added (e.g. cost of digital image or description, validation data, etc.).

Rationale: Based on the outcome of the presentations and discussions from the Barcelona meeting (January 2002), the latest content production examples from the content providers, the experiences with the first set of data entry templates and the document RN_T220v01_AIT_Datastructure, the following methodology will be adopted to form the base of the data entry approach. In order to cope with current and future needs we need to create data sets, records or files representing data for cataloguing, object and thematic descriptions and digital images supported by files containing the respective meta data.

The AMICO data specifications will be taken as basic guideline to realize this. This means in practice the generation of a separate data set for catalogue description files (AMICO catalogue records), object and thematic description files and image files (AMICO media files) and the corresponding meta data files (AMICO meta data records).

Applied to REGNET, this results in 7 types of files:

1. “object description” containing a substantial textual description about a physical object (outside of the catalogue description)
2. “thematic description” containing a substantial textual description about a thematic aspect
3. “image” containing a digital image
4. “catalogue description meta data” concerning a physical object (= equivalent to a (part of a) collection management system)
5. “object description meta data” concerning an object description
6. “image meta data” concerning a digital image
7. “thematic description meta data” concerning a thematic description

File types 1 and 2 can be generated by a standard text editor (.txt Unicode). File type 3 can be generated by scanning, a digital camera and a graphical software tool (.JPEG). File types 4, 5, 6 and 7 can be generated via a data entry template (.xls in this case).

The following table shows the 2- and 3-letter identifiers that will be used as unique references for the template and/or filing types. They can be used also in the file naming convention (see later).

Filing type identifiers	***Meta data data entry via template***	***Content data entry via text editor/image capture***
object description	odm	od
thematic description	tdm	td
image	imm	im

catalogue description	cdm	N/A
	excel-file	text or image file

Table 10: Template for object descriptions

Within some of the meta data entry templates, the following fields (which are specific for REGNET) permit the production and displaying of the same information in all combinations of different languages and different levels.

- content level (target audience)
- size level (in view of screen surface, number of KB)
- resolution level (for images)
- language

Other specific REGNET fields are included in order to cope with e-Business requirements (e.g. price) and editorial needs (e.g. validation).

Linking between resources: Links between resources (e.g. from thematic description to image and vice versa) can be established through the insertion (embedding) of numbers between brackets such as (1), (2), etc. within textual information. The address of the referenced resource is kept in the links-field of the meta data description with the corresponding number. Those links-fields correspond more *or less with the "related resources" field of meta data standards.*

Templates: The templates can be found in the files (available on the project server):

RN_T220v01_Template_odm.xls

RN_T220v01_Template_tdm.xls

RN_T220v01_Template_imm.xls

RN_T220v01_Template_cdm.xls

File naming convention: The TARX/MECH file naming convention is taking into account the following parameters:

- producer of the file (acronym of content provider)
- title, subtitle(s) and characteristics (Angela_life, replica)
- type of content (image, description, meta data, …)
- content level (target audience)
- size level (abstract, one page, 600x400 pixels, …)
- resolution level (75 dpi)
- language

Full http-addresses and directory names are completely omitted. They should be added and managed by the REGNET system once the files reside in the REGNET data structures. All file names should be in English.

General lay out of the file naming for object description:

nn_ti(_repeatable)_od_cl_sl_la.txt

nn acronym of the content provider

ti title of the content, repeated as far as necessary with additional information (sub-title or additional characteristics; one could mention for instance that this is a replica by adding "_replica")

od object description

cl content level (cl1, cl2, … see lower)

sl size level (sl1, sl2, … see lower)

la	language (ISO-code)
txt	Unicode text
cl1	children – youth
cl2	general (basic reference)
cl3	extended (some knowledge required)
cl4	scientific
sl1	abstract (words – few lines – one paragraph)
sl2	full screen - one page A4
sl3	several pages - article
sl4	… - book

Example : tarx_angela01_od_cl2_sl1_en.txt

An object description (od) by TARX (tarx) in English (en) and in text format (txt) pertaining to the object catalogued as “angela01” (e.g. sculpture representing St. Angela) with content level 2 (cl2) and size level 1 (sl1).

General lay out of the file naming for thematic description:

nn_ti(_repeatable)_td_cl_sl_la.txt

nn	acronym of the content provider
ti	title of the content, repeated as far as necessary with additional information (sub-title or additional characteristics)
td	thematic description
cl	content level (cl1, cl2, … see lower)
sl	size level (sl1, sl2, … see lower)
la	language (ISO-code)
txt	Unicode text

cl1	children – youth
cl2	general (basic reference)
cl3	extended (some knowledge required)
cl4	scientific

sl1	abstract (words – few lines – one paragraph)
sl2	full screen - one page A4
sl3	several pages - article
sl4	… - book

Example : mech_christopher_life_td_cl2_sl2_nl.txt

A thematic description (td) by MECH (mech) in Dutch (nl) and in text format (txt) pertaining to the life of St. Christopher with content level 2 (cl2) and size level 2 (sl2).

General lay out of the file naming for a digital image:

nn_ti(_repeatable)_im_rl_sl.jpg

nn	acronym of the content provider
ti	title of the content, repeated as far as necessary with additional information (sub-title or additional characteristics; for instance "_frontview" or "_detail")
im	image
rl	resolution level (r1, rl2, … see lower)
sl	size level (sl1, sl2, … see lower)
jpg	jpeg

rl1	around 75 dpi
rl2	around 150 dpi
rl3	around 300 dpi
rl4	around 600 dpi

sl1	thumbnail
sl2	300 x 200 pixels
sl3	600 x 400 pixels
sl4	…

Example : mech_holymary01_im_rl3_sl1.jpg

An image (im) produced by Mechelen (mech) in jpeg format (JPEG) with resolution level 3 and size level 1 pertaining to the object catalogued as "holymary01"

General lay out of the file naming for catalogue description meta data:

nn_ti(_repeatable)_cdm_la.xls

nn	acronym of the content provider
ti	title of the content, repeated as far as necessary with additional information (sub-title or additional characteristics)
cdm	catalogue description meta data
la	language (ISO-code)
xls	excel

Example : tarx_angela01_cdm_nl.xls

A file delivered by TARX containing catalogue description meta data in Dutch about the object with identifier “angela01”.

General lay out of the file naming for object description meta data.

nn_ti(_repeatable)_odm_cl_sl_la.xls

This file naming follows exactly the same structure as the corresponding object description except for the file type and extension (here it is odm instead of od and xls instead of txt).

Example : tarx_angela01_odm_cl2_sl1_en.xls

The object description meta data (odm) by TARX (tarx) in English (en) and in excel-format (xls) pertaining to the object catalogued as “angela01” (e.g. sculpture representing St. Angela) with content level 2 (cl2) and size level 1 (sl1).

General lay out of the file naming for thematic description meta data.

nn_ti(_repeatable)_tdm_cl_sl_la.xls

This file naming follows exactly the same structure as the corresponding thematic description except for the file type and extension (here it is tdm instead of td and xls instead of txt).

Example : mech_christopher_life_tdm_cl2_sl2_nl.xls

The thematic description meta data (tdm) by MECH (mech) in Dutch (nl) and in excel format (xls) pertaining to the life of St. Christopher with content level 2 (cl2) and size level 2 (sl2).

General lay out of the file naming for image meta data.

nn_ti(_repeatable)_imm_rl_sl_la.xls

This file naming follows exactly the same structure as the corresponding image description except for the file type and extension (here it is imm instead of im and xls instead of jpg).

Example : mech_holymary01_imm_rl3_sl1.xls

The image meta data (imm) produced by Mechelen (mech) in Dutch (nl) and in excel format (xls) for an image with resolution level 3 and size level 1 pertaining to the object catalogued as “holymary01”

5.2.3.2 Phase 2

The method proposed in phase 1 provoked a huge amount of individual files. For large quantities it became difficult to keep a general overview and a sufficient level for coherence and consistency checks. Therefore it was decided to combine the meta data records into one spreadsheet book with different sheets. With some slight changes in the fields lay out, the overall methodology remained the same.

Naming convention for off line data entry of records and files within REGNET: This section describes one possible methodology for naming files and meta data records in a coherent and consistent way.

The naming is taking into account the following parameters:

- producer of the file (acronym of content provider)
- title, subtitle(s) and characteristics (Angela_life, replica, existing inventory number of a collection object)
- type of content (catalogue, object, image, thematic description); this is a pivotal item, a reference: before this parameter are the parameters containing editors, names, titles, numbers (identification) and after this parameter reside the "scoping" parameters such as language and the levels.
- content level (target audience)
- size level (abstract, one page, 600 pixels, …)
- resolution level (150 dpi)
- language

Full http-addresses and directory names are completely omitted. They should be added and managed by the REGNET system once the files reside in the REGNET data structures. All file names should be preferably in English, also for files containing other languages.

General lay out of the naming for the catalogue description meta data record:

nn_ti(_repeatable)_cd_la

nn acronym of the content provider

ti title of the content, freely repeated as far as necessary with additional information (sub-title or additional characteristics)

cd catalogue description meta data, this is a fixed denomination

la language (ISO-code)

Example : tarx_angela01_cd_nl

A record delivered by TARX containing catalogue description meta data in Dutch (nl) about the object catalogued as "angela01" (e.g. sculpture representing St. Angela).

General lay out of the naming for the object description meta data record.

nn_ti(_repeatable)_od_cl_sl_la

nn acronym of the content provider

ti title of the content, freely repeated as far as necessary with additional information (sub-title or additional characteristics; one could mention for instance that this is a replica by adding "_replica")

od object description, this is a fixed denomination

cl content level (cl1, cl2, … see guide)

sl size level (sl1, sl2, … see guide)

la language (ISO-code)

Example : tarx_angela01_od_cl2_sl1_en

A record delivered by TARX containing object description meta data in English (en) pertaining to the object catalogued as “angela01” with content level 2 (cl2) and size level 1 (sl1).

General lay out of the naming for the object description file:

nn_ti(_repeatable)_od_cl_sl_la.xxx

This file naming follows exactly the same structure as the corresponding meta data record with the addition of the used encoding (xxx).

Example : tarx_angela01_od_cl2_sl1_en.txt

An object description (od) by TARX (tarx) in English (en) and in text format (txt) pertaining to the object catalogued as “angela01” (e.g. sculpture representing St. Angela) with content level 2 (cl2) and size level 1 (sl1).

General lay out of the naming for the thematic description meta data record.

nn_ti(_repeatable)_td_cl_sl_la

nn acronym of the content provider

ti title of the content, freely repeated as far as necessary with additional information (sub-title or additional characteristics)

td thematic description, this is a fixed denomination

cl content level (cl1, cl2, … see guide)

sl size level (sl1, sl2, … see guide)

la language (ISO-code)

Example : mech_christopher_life_td_cl2_sl2_nl

The thematic description meta data record by MECH (mech) in Dutch (nl) pertaining to the life of St. Christopher with content level 2 (cl2) and size level 2 (sl2).

General lay out of the naming for the thematic description file:

nn_ti(_repeatable)_td_cl_sl_la.xxx

This file naming follows exactly the same structure as the corresponding meta data record with the addition of the used encoding (xxx).

Example : mech_christopher_life_td_cl2_sl2_nl.txt

A thematic description (td) by MECH (mech) in Dutch (nl) and in text format (txt) pertaining to the life of St. Christopher with content level 2 (cl2) and size level 2 (sl2).

General lay out of the naming for the image meta data record.

nn_ti(_repeatable)_im(_nr)_rl_sl_la

nn	acronym of the content provider
ti	title of the content, freely repeated as far as necessary with additional information (sub-title or additional characteristics; for instance "_frontview" or "_detail")
im	image, this is a fixed denomination
nr	optional number in case there are several images of the same object
rl	resolution level (r1, rl2, … see guide)
sl	size level (sl1, sl2, … see guide)
xxx	the used encoding: jpeg, …

Example : mech_holymary01_im_rl3_sl1_nl

The image meta data (im) produced by Mechelen (mech) in Dutch (nl) for an image with resolution level 3 and size level 1 pertaining to the object catalogued as "holymary01"

General lay out of the file naming for a digital image:

nn_ti(_repeatable)_im(_nr)_rl_sl.xxx

This file naming follows exactly the same structure as the corresponding meta data record with the omission of the language code and the addition of the used encoding (xxx).

Example : mech_holymary01_im_rl3_sl1.jpg

An image (im) produced by Mechelen (mech) in jpeg format (jpg) with resolution level 3 and size level 1 pertaining to the object catalogued as "holymary01"

Lay out of templates with examples

<catalogue_description>						
<identifier>	<language>					
tarx_angela01_cd_en	en					
</identifier>	</language>					
<title>	<object_type>	<mat_techn>	<dimension>	<dim_value>	<dim_unit>	
Saint Angela	glass door panel	glass	width	50	cm	
</title>	original	etch	heigth	200	cm	
	</object_type>	</mat_techn>	depth	0,2	cm	
			</dimension>	</dim_value>	</dim_unit>	
<description>						
</description>						
<location>	<author>	<author_org>	<creation_date>	<contributor>	<contrib_org>	<comment>
Convent of the Ursulines; Tildonk Belgium	unknown	unknown	end 19th century			
</location>	</author>	</author_org>	</creation_date>	</contributor>	</contrib_org>	</comment>
<relation_link>	<primary_keyw>	<secondary_keyw>	<rights>	<price>	<currency>	
(1): tarx_angela01_od_cl2_sl1_en	saint	glass	Vereniging van de religieuzen Ursulinen van Tildonk vzw	N/A	Euro	
(2): tarx_angela01_im_rl2_sl3_en	angela	etch	</rights>	</price>	</currency>	
</relation_link>	</primary_keyw>	</secondary_keyw>				

<original documenting and validation>	<doc_author>	<doc_auth_org>	<doc_date>	<doc_validator>	<doc_val_org>	<d_val_date>	<d_comment>
	Sermeus, Rosette	Tarx nv Belgium	2002-02-16	Haesaerts, Vic	Tarx nv Belgium	2002-02-18	
</original documenting and validation>	</doc_author>	</doc_auth_org>	</doc_date>	</doc_validator>	</doc_val_org>	</d_val_date>	</d_comment>
<docum. modification and validation>	<mod_author>	<mod_auth_org>	<mod_date>	<mod_validator>	<mod_val_org>	<m_val_date>	<m_comment>
</docum. modification and validation>	</mod_author>	</mod_auth_org>	</mod_date>	</mod_validator>	</mod_val_org>	</m_val_date>	</m_comment>
</catalogue_description>							
<catalogue_description>							
new template							
</catalogue_description>							

Table 11: Template for the catalogue description

<object_description>						
<identifier>	<content_level>	<size_level>	<language>	<encoding>	<file_size>	<file_size_unit>
tarx_angela01_od_cl2_sl1_en	cl2	sl1	en	doc	19	kb
</identifier>	</content_level>	</size_level>	</language>	</encoding>	</file_size>	</file_size_unit>
<title>	<description>					
Saint Angela						
</title>	</description>					
<resource_locator>	<author>	<author_org>	<creation_date>	<contributor>	<contrib_org>	<comment>
tarx_angela01_od_cl2_sl1_en.doc	Sermeus, Rosette	Tarx nv Belgium	2001-12-10			
</resource_locator>	</author>	</author_org>	</creation_date>	</contributor>	</contrib_org>	</comment>

<relation_link>	<primary_keyw>	<secondary_keyw>	<rights>	<price>	<currency>	
(1): tarx_angela01_cd_en	saint	glass	Tarx nv Belgium	0	Euro	
(2): tarx_angela01_im_en	angela	etch	</rights>	</price>	</currency>	
</relation_link>	</primary_keyw>	</secondary_keyw>				
<original documenting and validation>	<doc_author>	<doc_auth_org>	<doc_date>	<doc_validator>	<doc_val_org>	<d_val_date>	<d_comment>
	Sermeus, Rosette	Tarx nv Belgium	2002-02-16	Haesaerts, Vic	Tarx nv Belgium	2002-02-18	
</original documenting and validation>	</doc_author>	</doc_auth_org>	</doc_date>	</doc_validator>	</doc_val_org>	</d_val_date>	</d_comment>
<docum. modification and validation>	<mod_author>	<mod_auth_org>	<mod_date>	<mod_validator>	<mod_val_org>	<m_val_date>	<m_comment>
</docum. modification and validation>	</mod_author>	</mod_auth_org>	</mod_date>	</mod_validator>	</mod_val_org>	</m_val_date>	</m_comment>
</object_description>							
<object_description>							
new template							
</object_description>							

Table 12: Template for the object description

<image_description>

<identifier>	<resol_level>	<size_level>	<language>	<encoding>	<file_size>	<file_size_unit>
tarx_angela01_im_rl2_sl4_en	rl2	sl4	en	jpg	52	kb
</identifier>	</resol_level>	</size_level>	</language>	</encoding>	</file_size>	</file_size_unit>
<title>	<color_palette>	<color_metric>	<dimension>	<dim_value>	<dim_unit>	<resol_value>
Saint Angela	RGB	24	width	709	pixels	200

</title>	</color_palette>	</color_metric>	heigth	1112	pixels	</resol_value>	
			</dimension>	</dim_value>	</dim_unit>		
<resource_locator>	<author>	<author_org>	<creation_date>	<contributor>	<contrib_org>	<comment>	
tarx_angela01_im_rl2_sl4.jpg	Sermeus, Rosette	Tarx nv Belgium	2001-12-10				
</resource_locator>	</author>	</author_org>	</creation_date>	</contributor>	</contrib_org>	</comment>	
<relation_link>	<primary_keyw>	<secondary_keyw>	<rights>	<price>	<currency>		
(1): tarx_angela01_cd_en	saint	glass	Tarx nv Belgium	0	Euro		
(2): tarx_angela01_od_cl2_sl1_en	angela	etch	</rights>	</price>	</currency>		
(3): tarx_angela_life_td_cl2_sl1_en	</primary_keyw>	</secondary_keyw>					
</relation_link>							
<original documenting and validation>	<doc_author>	<doc_auth_org>	<doc_date>	<doc_validator>	<doc_val_org>	<d_val_date>	<d_comment>
	Sermeus, Rosette	Tarx nv Belgium	2002-02-16	Haesaerts, Vic	Tarx nv Belgium	2002-02-18	
</original documenting and validation>	</doc_author>	</doc_auth_org>	</doc_date>	</doc_validator>	</doc_val_org>	</d_val_date>	</d_comment>
<docum. modification and validation>	<mod_author>	<mod_auth_org>	<mod_date>	<mod_validator>	<mod_val_org>	<m_val_date>	<m_comment>
</docum. modification and validation>	</mod_author>	</mod_auth_org>	</mod_date>	</mod_validator>	</mod_val_org>	</m_val_date>	</m_comment>

</image_description>

<image_description>

new template

</image_description>

Table 13: Template for the object description

<thematic_description>

<identifier>	<content_level>	<size_level>	<language>	<encoding>	<file_size>	<file_size_unit>
tarx_angela01_tdod_cl2_sl1_en	cl2	sl1	en	doc	19	kb
</identifier>	</content_level>	</size_level>	</language>	</encoding>	</file_size>	</file_size_unit>
<title>	<description>					
The life of Saint Angela						
</title>	</description>					
<resource_locator>	<author>	<author_org>	<creation_date>	<contributor>	<contrib_org>	<comment>
tarx_angela_life_td_cl2_sl1_en.doc	Sermeus, Rosette	Tarx nv Belgium	2001-12-10			
</resource_locator>	</author>	</author_org>	</creation_date>	</contributor>	</contrib_org>	</comment>
<relation_link>	<primary_keyw>	<secondary_keyw>	<rights>	<price>	<currency>	
(1): tarx_angela01_cd_en	saint	education	Tarx nv Belgium	0	Euro	
(2): tarx_angela01_im_en	angela	boarding schools	</rights>	</price>	</currency>	
(3): tarx_angela02_cd_en	</primary_keyw>	ursula				
(4): tarx_angela02_im_en		ursulines				
(5): tarx_angela_life_td_cl2_sl2_en		Desenzano, Italy				
</relation_link>		</secondary_keyw>				

<original documenting and validation>	<doc_author>	<doc_auth_org>	<doc_date>	<doc_validator>	<doc_val_org>	<d_val_date>	<d_comment>
	Sermeus, Rosette	Tarx nv Belgium	2002-02-16	Haesaerts, Vic	Tarx nv Belgium	2002-02-18	
</original documenting and validation>	</doc_author>	</doc_auth_org>	</doc_date>	</doc_validator>	</doc_val_org>	</d_val_date>	</d_comment>
<docum. modification and validation>	<mod_author>	<mod_auth_org>	<mod_date>	<mod_validator>	<mod_val_org>	<m_val_date>	<m_comment>

</docum. modification and validation>	</mod_author>	</mod_auth_org>	</mod_date>	</mod_validator>	</mod_val_org>	</m_val_date>	</m_comment>

</thematic_description>

<thematic_description>

new template

</thematic_description>

Table 14: Template for the thematic description

For Guidelines to use the templates see Appendix 5.

6 Presentation of data (examples)

6.1 Thematic content – the Saints theme as one example

6.1.1 Scope and synopsis

This paper contains a pragmatic methodology for the production of content concerning the SAINTS theme. It is based on the results of the study of all aspects of REGNET content production and a series of trial and error iterations by MECH and TARX. The involved persons are: Heidi De Nijn (MECH), Lieve Lettany (MECH), Rosette Sermeus (TARX) and Vic Haesaerts (TARX).

The rationale behind this approach was to produce fragmented content through file structures that can be used into any type of final data entry and database system that will be chosen within REGNET.

Synopsis: Within the REGNET-project, “SAINTS” is one of the central themes around which a number of related themes are set up in order to form a domain wherein pieces of information (we call them fragments in a general way) containing cultural content will be produced. This has to be done in such a manner that these pieces of information can be associated to a theme or several themes but also can contain references to each other. Because the SAINTS theme represents, de facto, a very broad range of information sources, priority will be given to some subsets of saints and saints related items. The availability of saints related objects within the respective CIOs is one influencing factor. Another one is the level of communality with other (than Roman-Catholic) religions and cults (see PROTECTION- and FERTILITY-sections in the SAINTS, MASKS, AMULETS and (GILT) LEATHER themes).

Different types of fragments are envisaged, the most important being thematic descriptions, object descriptions and object images. Other types could be: audio, video, web sites, etc. An important requirement in setting up these different types of fragments is their reusability, extensibility and scalability within different contexts. Whereas the object image- and object descriptions-fragments pertain specifically to the respective collections of the cultural institutions and organizations (CIOs), the thematic descriptions contain descriptive information, offering a more educational, scientific and contextual setting for the different objects.

The theme leadership for SAINTS resides in the Stedelijke Musea Mechelen (MECH). Because of the broad character of this theme, a lot of other REGNET-partners have committed to contribute to this theme. The following points aim to present some methodology and guidelines for the production of SAINTS-related content. The following pages contain the detailed structure of the information that will be produced primarily by MECH. It is preferable that contributions from other partners should be based upon this structure. We think that the currently proposed methodology is flexible enough to be used in most of the translating or matching processes that have to be carried out for final REGNET-compatibility.

6.1.2 Nomenclature

The following nomenclature will be used throughout this paper.

Theme: something worth to be spoken, written about; different theme levels are possible.

Theme scheme: the set of related themes that will be considered to form the SAINTS-theme (current boundaries of the domain that we want to cover).

Object description: texts consisting of different types and versions, but all pertaining to the same object.

Image: an image concerning an object.

Thematic description: mainly textual information on an aspect of a theme (definition, life, cult, but also history, style, etc.). The main aim of thematic descriptions is to give contextual and background information for the collection objects presented in REGNET. Graphical support can be given by pointing to an object or a set of objects.

Fragment: an information resource in general for a theme, in practice an object description, a thematic description or an image.

Because we will use the topic map paradigm for the development of the REGNET system the following terms will also be used together with their mapping to the terms mentioned in the previous paragraph.

Topic: corresponds with theme

Association: corresponds with the relations between themes

Occurrence: corresponds with object descriptions and thematic descriptions (fragments)

Other terms will be used in the topic map language but the three mentioned above are the main ones.

6.1.3 List of the set of initial themes for SAINTS

In order to obtain a manageable content for the REGNET-demo, the initial theme scheme contains a limited number of related themes. Special attention was given to some themes which are closely related to other central themes such as MASKS and LEATHER in order to demonstrate reusability of fragments in different thematic structures. The aim of this list is to define the first thematic domain for SAINTS wherein contributions can be produced. In principle there is no limit to this list. REGNET partners can add supplementary themes, especially individual saints.

SAINT related themes - topics		
SAINT	PATRON SAINT	PROTECTION
FERTILITY	URSULINES	PLAGUE
TRAVELLING	PILGRIMAGE	LEATHER-WORKERS
ST. ANTHONY	ST. ROCHE	ST. SEBASTIAN
ST. ADRIAN	ST. CHRISTOPHER	ST. ANNE
HOLY MARY	ST. JAMES THE GREATER	ST. CRISPIN
ST. CRISPINIAN	ST. ANGELA	ST. URSULA

Table 15: List of themes for Saints

The choice of those themes is based on:

- the availability of a substantial number of related objects within CIOs
- the extensibility of the treated theme (there are a lot of saints)
- the high degree of linking to other thematic domains
- high potential involvement of several European regions

These themes will become topics in the topic map design.

Until now we consider these themes as individual stand alone themes against which we are going to produce a considerable number of related object descriptions and images. At the same time we want to produce also thematic descriptions related to those themes. These thematic descriptions express more supporting contextual information about the themes. They do not relate directly to the objects but they can make use of them as examples.

Objects can be used within thematic descriptions as supporting material. They are addressed via links as can be seen in the examples in annex. These links must be set up in such a way that the REGNET

system can make the appropriate choice of the best suited object description and image depending on the profile of the end user and the type of the used equipment. The best way to do this is make reference to the entity ANTHONY01 for instance which contains different descriptions and images of the object ANTHONY01. Another way to proceed is to make a query of available objects of St. Anthony and to make the best choice.

The primary keywords will be restricted to the ones that correspond directly with a theme or topic that will be present in the *topic* map. As far as possible, existing thesauri will be used for expressing these *keywords*. There is more freedom for the choice of the (less important) secondary keywords, but they also have to be compliant as far as possible with existing thesauri.

6.1.4 Relations between themes – topics (associations)

Some themes, not all, have relations between each other. These relations can be hierarchical or specific. The following table expresses the first set of relationships between the SAINT-themes. Because we are entering here the domain of semantic nets and knowledge engineering, attention should be paid to the correct and appropriate application of these matters. Some formal agreement about this should be obtained. E.g. the relations must be very explicit, not simply stating that there is some kind of relationship. In topic map language, these relations are called ”associations”.

ENGLISH			DUTCH		
Topic	Association	Topic	Topic	Association	Topic
Anthony	is-a	Saint	Antonius	is-een	Heilige
Roche			Rochus		
Sebastian			Sebastianus		
Adrian			Adrianus		
Christopher			Christophorus		
Anne			Anna		
Maria			Maria		
Crispin			Crispinus		
Crispinian			Crispinianus		
James-the-Greater			Jacobus-de-Meerdere		
Angela			Angela		
Ursula			Ursula		
Anne	is-mother-of	Maria	Anna	is-moeder-van	Maria
Crispin	is-brother-of	Crispinian	Crispinus	is-broeder-van	Crispinianus
Angela	was-influenced-by	Ursula	Angela	werd-beïnvloed-door	Ursula
Angela	founded	Ursulines	Angela	stichtte	Ursulinen
Anthony	is-patron-saint-against	Plague	Antonius	is-beschermheilige-tegen	Pest
Roche			Rochus		
Sebastian			Sebastianus		
Adrian			Adrianus		
Christopher			Christophorus		
Christopher	is-patron-	Travellers	Christophorus	is-patroonheilige-	Reizigers

	saint-of			van	
James-the-Greater		Pilgrims	Jacobus-de-Meerdere		Pelgrims
Ursula		Ursulines	Ursula		Ursulinen
Crispin		Leather-workers	Cripinus		Leerbewerkers
Crispinian			Crispinianus		
Anne	is-symbol-of	Fertility	Anna	is-symbool-van	Vruchtbaarheid
Maria			Maria		
Plague	is-subject-for	Protection	Pest		Beschermin g
Travelling			Reizen		

The right definition of associations is not an easy task. Supplementary information of played roles in associations is also important. E.g. in “Anthony is patron saint against the plague”, the role of “Anthony” is “protection” and the role of “plague” is “disease”.

Additional input from other partners is highly desirable. The editorial team can make some final decisions.

6.1.5 Relations between themes (topics) and object images, object descriptions and thematic descriptions (occurrences).

Themes - Objects.

Referring to the data specification of the AMICO-model, all information related to an object will be grouped into an entity (files, records, fields, XML) having a unique identifier for this object.

E.g. the entity ANTHONY01 will contain all information about the object ANTHONY01 of the MECH collection:

When the REGNET system is referring to this topic, it can reach all information about the object. Depending on the context (desired level of end user, location or profile (language) of end user, consultation equipment of end user), the system can make the appropriate choice between the available information that suits best the end user and his equipment. E.g. the extended description in English of an object in a MECH museum for an interested visitor from England on the Internet. Or the basic version in Dutch for a Flemish speaking person, visiting physically a MECH museum using his portable device for consultation (Internet WAP, Bluetooth, etc.).

CONSIDERATION POINT: In order to realise this technically, we could follow several ways. One could be to consider the above-mentioned entity ANTHONY01 as one occurrence of ANTHONY. Or we could make from ANTHONY01 a topic with an association to the topic ANTHONY (ANTHONY01 “is an object representing” ANTHONY). The different files related to ANTHONY01 become then occurrences of ANTHONY01. The ultimate choice will influence the mechanisms used for database management, search and retrieval, etc.

Themes – Thematic descriptions

To a great extent, for themes the same can be applied as for object descriptions. The multilevel and multilingual solutions for objects are equally valid for thematic descriptions.

The biggest difference lies in the relationship data-metadata. In thematic descriptions there is a direct relationship with the end product, i.e. the text. In object descriptions we have two levels of relationships: one of the author of the description of an object and one of the author or creator of the object itself. Nesting of (metadata) descriptions can be envisaged.

7 References

AMICO	Art Museum Image Consortium	http://www.amico.org/
CDWA	Categories for the Description of Works of Art	http://www.getty.edu/research/institute/standards/cdwa
CIDOC	The International Committee for Documentation of the International Council of Museums (ICOM-CIDOC) / Le Comité international pour la documentation du Conseil international des musées (ICOM-CIDOC)	http://www.cidoc.icom.org/guide/guideint.htm
EAD	Encoded Archival Description	http://lcweb.loc.gov/ead/
MARC	Machine Readable Cataloguing Format	http://lcweb.loc.gov/marc/
MDA	UK Museum Documentation Standard (Spectrum)	http://www.mda.org.uk/spectrum.htm
W3 Consortium-XML		http://www.w3c.org/XML
WfMC	Workflow Management Coalition (WfMC) Standard for search and retrieval of bibliographic data	http://www.wfmc.org/standards/docs.htm

List of Figures

List of Tables

Table of Appendices