Report on Web Services

Marc Kemps-Snijders (MPI Nijmegen), Nuria Bel (UPF Barcelona)
March 2009

This document is based in particular on the collaboration with the experts from RACAI, USheffield, ULeipzig, UTübingen, UStuttgart, DFKI, UWroclaw, BBAW, UPF and MPI.

Content

1. Introduction

Web Services are the elementary building blocks in service-oriented architectures which will play in the open and interoperable CLARIN scenario a key role. Before entering a detailed discussion it seems to be necessary to clarify some terms:

- Tools are denote pieces of code that are bundled together to fulfill a certain set of functions and that work stand-alone, i.e. carry out this function in a certain environment (operating system, interpreter etc) by using some input and some parameters without interacting with other tools.
- Web-Sites are presentations of some information that is encoded in HTML - be it generated in a static or dynamic manner - by making use of the HTTP protocol to transport the information. Often these presentations are enriched by scripting components (Javascript), applets or other components that are executed on the user's desktop computer or notebook and that interact with services offered by some service center. Web Sites can offer interaction features
- Web Applications carry out a function for the user and that offer user interaction possibilities for entering data, visualizing data etc. These web application typically have user client and server components that interact with each other via the web, i.e. using the HTTP protocol. There is no principle difference between web-sites with active components and web applications.
- Web Services are software components that are being invoked by web applications to fulfill a specific sub-task that is needed to solve a problem, but that is not part of the code of the application but being offered by some service provider and being executed on some server. Web services themselves can activate other web services. To interact Web services will typically interact by making use of the HTTP protocol and additional interface specification and messaging standards.
- Service-Oriented-Architecture (SOA) is the cover term that describes the distributed system of web services and web applications that can invoke each other based on proper agreements on interaction standards. Due to its distributed character SOA are seen as being a very flexible framework to combine various software components to new more powerful applications.
- Workflow systems are web applications that allow users to combine web applications and web services to processing chains. They allow users to look into registries of web services and applications, to select them and combine them. In addition these workflow applications allow users to select data resources and parameter files that serve as inputs to the various processing steps in a chain. Workflows can become complex in particular when asynchronous interaction types are involved. When one service is waiting on the service of a preceding one to start operation we speak about synchronous interaction. When a service is waiting on the results of several preceding services all having different speeds then we have asynchronous interaction the handling of which is more complicated. This workflow construction work can either be done by a scripting language or by graphical programming manners. It is known that graphical programming is in general not as powerful, but much more simple for the naive user.

This note describes the state of discussion and activities about web services and workflow systems (WS&WF) in a summarizing form. Therefore it is not yet a full requirements specification document as it is requested as deliverable D2.6. The note is based on the following activities that took place until now:

- first description of the task at the CLARIN kickoff meeting — Nijmegen, June 2008
- D-SPIN/CLARIN workshop on Web Services & Workflows — Munich, November 2008
- Subproject meetings based on national funds — various in Sp and De
- WP2 Expert workshop on WS&WF — Sheffield, January 2009

- WP2 Workshop Oxford, February 2009

During all the time various bilateral meetings took place between the core partners of the Working Group dealing with Web Services and Workflow issues. The work is under the leading of Nuria Bel (UPF) and Marc Kemps-Snijders.

In general we can say that quite a number of clarifications have been achieved and some elaborations of the web services and workflow field have been achieved. However, we need to conclude that a number of questions that are addressed in this working group have not yet been solved and are also not fully understood. This area is less mature compared for example to the metadata area.

In the following we will first (1) describe the Line of Work which will include a motivation of the various steps, the results and the steps already planned, (2) the constraints on these activities as specified by other activities in CLARIN, (3) architectural specifications, (4) some facts which we can already specify at this moment and (5) some remarks about the risks.

2. Line of Work

The working group is organized along 3 activities: (1) workshops to listen to work reports and to carry out abstractions and harmonization, (2) concrete bottom-up driven implementation activities in small teams and (3) specification and implementation of a CLARIN prototypical system.

Workshops

For the D-SPIN workshop a number of internal and external experts were invited who have already been working on web services and workflow issues for a number of years. The primary intention was to get a broad overview about approaches and the problems associated with them. The pillars of an open CLARIN web services infrastructure were discussed. Existing suggestions for metadata such as UDDI and ebXML were evaluated and compared with the domain specific requirements which emerge for example from profile matching wishes. Suggestions of how to solve the Import/Export problem were compared, in particular the strong typing as suggested by UIMA and the weak typing as being done in GATE. Finally standard workflow solutions such as JBPM and BPEL solutions were compared where Teamware represents as a domain specific application. It is expected that the open source JPM solution will be able to deliver a work flow engine that is capable of processing both JPDL and BPEL process definitions.

The follow-up expert workshop in Sheffield brought together those teams from CLARIN that are already working on WS&WF issues. Their approaches and solutions were analyzed in detail with the goal to determine the basis for the CLARIN show cases. Approaches and solutions were presented and discussed in detail as well as the opportunities for the partners that are involved.

The WP2 workshop in Oxford was meant to again listen to a number of reports of concrete activities, to summarize the state of the discussions, to discuss harmonization activities and to link up with the other activities in WP2. A convergence with respect to a number of conclusions was achieved (see below).

Reports from the first two workshops have been made available, the one about the Oxford meeting is in progress. However, a list of conclusions has been distributed already.

The following workshop activities are planned that are directly meant to support the WS&WF work:

- In April there will be a workshop organized by WP5 in Athens where a number of typical processing chains will be analyzed in detail with respect to the flow of linguistic information. This analysis work will deepen the understanding about the requirements in particular for import/export issues.

- In May/June there will be a workshop organized by WP2 in Barcelona where experts will discuss the results from the Athens workshop and try to transform the results into requirement specifications. It is intended to invite experts from leading US projects and institutions (CMU, Bamboo) to understand their directions.

Implementation Activities

While the workshops always have a harmonization and abstraction aspect which can be seen as a top-down approach, it was clear from the beginning that we also need bottom-up activities, i.e. work on concrete solutions. There are mainly two reasons for doing this: (1) We need practical work to improve our own understanding of the problems we are faced with. (2) To master the problems in the construction phase CLARIN will need a group of young experts that bring together sufficient knowledge from linguistics and a decent knowledge about the practical IT aspects. This can only be achieved by a "learning-by-doing" approach.

Currently we see the following teams working on WS&WF matters:

- Experts from Sheffield are transforming their GATE components to web-services and they are working on an advanced graphical workflow tool called Teamware.
- Experts from Romania are also transforming some web-based services to web services and they are working on issues such as automatic detection of formats.
- Experts from Leipzig, Tübingen and Stuttgart are working on a joint project on implementing a typical processing chain bringing together web services and resources from the different institutes.
- Experts from Barcelona working on web services implementations of their language technology and about generic workflow tools.
- Experts from DFKI who created their Heart of Gold architecture for web services and who are now looking for ways to adapt to CLARIN.
- Experts from the MPI who are offering metadata, ISOcat DCR and LMF-based lexicon services as web services and who need to improve these services.
- Experts from Wroclaw who want to turn their tools into web services.

Some of these activities are running under national and institutional funds.

Prototypical System

It the task of the workshops to bring these experts together, determine common action lines, work out harmonization directions and specify a common EU-level CLARIN show case. The paragraph on conclusions may give an impression about the state of this discussion.

3. Relations to other CLARIN Activities

We can see a few activities that have a relevance for the WS&WF working group, but that are dealt with by other work packages and working groups. Workshops and personal interactions ensure that there is a continuous exchange of ideas, approaches and solutions. We can refer in particular to the following issues: (1) centres of storage and computation, (2) metadata to describe tools and services, (3) single

sign-on principles, (4) approaches for semantic interoperability, (5) analysis of use cases and (6) IPR issues and business models.

Centres

Much of the data used in CLARIN will be stored in institutional repositories and archives, the change rate is predominantly low. For many operations users will carry out work on data resources resulting in annotations, i.e. the change rates are much higher. Therefore CLARIN must ensure that users have access to temporal workspaces where they can carry out their operations and store created data in a temporary fashion. End users will typically use this work space for day to day work. It must be easily possible to push results forward from private workspaces to shared repositories. However, the rules of operation in workspaces needs to be the same to not create irritations with laymen as users. Sharing resources by making them visible and accessible is one of the core pillars of CLARIN. All resources within a user's private workspace must therefore also be associated with at least the minimal amount of required metadata information. This metadata can initially be obtained through the user uploading the resource into his/her private workspace.

Obviously we need centres that can offer workspaces and computational capacities to allow arbitrary researchers to execute the offered services. It is obvious that there are currently no solutions, since this is coupled with investments. In the CLARIN operational phase this issue needs to be solved probably in collaboration with the high performance computer centres in Europe. This collaboration will also need to come solutions of how to quickly transfer large amount of data to the processes. Yet the grid solutions do not seem to work seamlessly and efficiently. Big companies are implementing cloud computing techniques - basically a revival of the central computer idea where a smart dispatcher software distributes requests to a large of cores. CLARIN will seek to establish a close collaboration with initiatives such as DEISA.

Metadata

The CLARIN infrastructure is about language resources in a general sense. It includes both data resources, services and tools. All these are described in metadata[1] registries which provide an overview of all resources and their general characteristics. The CLARIN meta data infrastructure will be based on the flexible component architecture as describe in CLARIN EB-2008-5 (Metadata Infrastructure for Language Resources and Technology). Metadata is both intended for both human consumption as well as automated support tasks. Particularly for the latter tasks technical information, such as format, encoding, descriptions of input/output formats becomes vitally important. These automated support tasks are intended to provide users guidance through the workflow specification process. The lower the technical expertise level of the users, the more automated support must be provided.

The specific metadata fields that need to be recorded for each of the resources, services or tools may not be based on a static schema and will strongly depend on the type of resource, service or tool being described and the manner in which they are used. At least a basic level of interoperability should be provided by mapping information in meta data elements to well defined concepts, such as for example listed in the Data Category Registry.

[1] Metadata here is meant in the sense of restricted keyword type of descriptions and not in the general sense which covers all sorts of annotations.

The WS&WF working group will rely on the outcome of the metadata working group on (a) the specification of the elements with which all language resources and technology components will be described and contribute to this process and (b) will interface with the metadata infrastructure architecture which is described elsewhere. A UML diagram explains the essentials of the architecture.

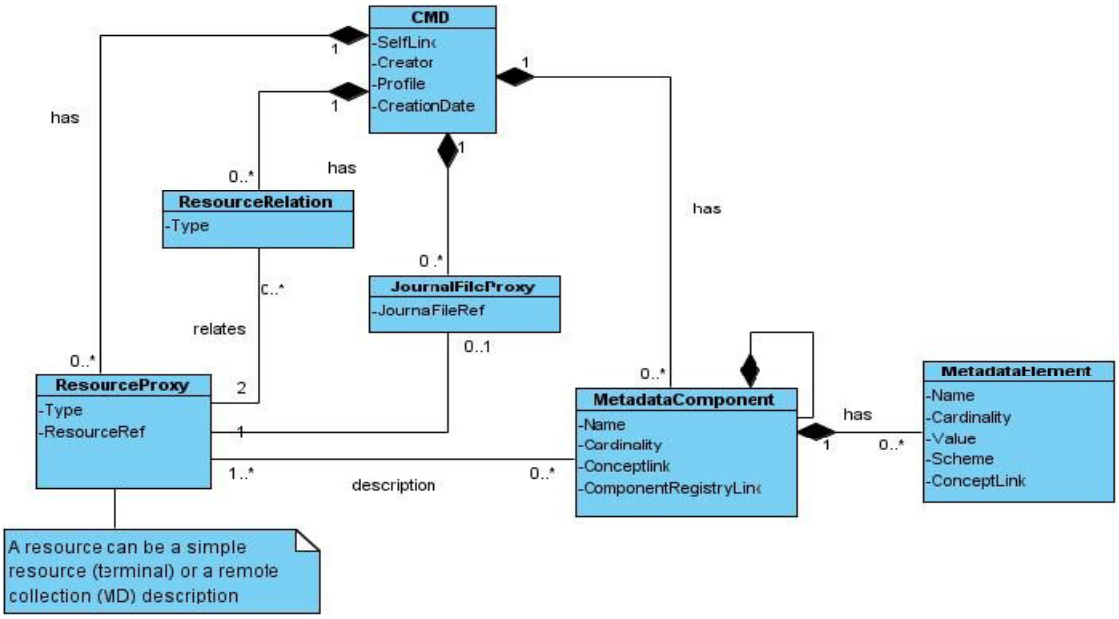

Figure 1: Core elements of the CLARIN metadata architecture

CLARIN web service registries need to adopt this architecture.

Single Sign-On Issues

The WS&WF working group will currently not tackle the task of how to pass on user credentials in a SOA to support single identity and single sign-on principles as intended in CLARIN. For the preparation phase we will assume that we can bypass these problems. Currently, the MPI is studying these issues together with the Dutch Big Grid project by using Short Lived Certificates. This project will report to CLARIN about the perspectives.

Semantic Interoperability

All aspects with respect to semantic interoperability at the level of semantic tags are dealt with by the ISO data category registry initiative ISOcat. Web services will be provided to access this registry. The services will be described at the ISOcat site (www.isocat.org). We expect that more needs to be done, in particular the infrastructure for relation management needs to be worked out by ISO.

Humanities Use Cases and Linguistic Properties

The WS&WF working group will heavily rely on the analysis work of work packages 3 on typical applications used within the humanities and on the work of work package 5 which will analyze the linguistic properties on typical processing chains as they occur when using language resources and technology in the LRT domain. From WP5 we expect a much deeper insight with respect to

- the types of chains - long and short ones from various areas of the LRT domain

- the variation in import/export formats (lexical and annotation structures)
- the variation in linguistic encoding per type of operation
- the specifications and selections to be made when accessing complex resources

The coming WP5 workshop will give valuable input and also the call for proposals in WP3 will provide insights. Finally we expect recommendations on best practices and standards with respect to the mentioned issues.

IPR Issues and Business Models

The WS&WF working group will ignore all IPR issues and business model issues in the preparatory phase, since these would complicate matters and distract us from tackling the difficult technical problems to solve first. Here we will collaborate with WP7.

3. Architecture Specifications

In this chapter we will describe a number of architectural problems we are faced with and discuss solutions. The dimensions which have already been discussed are: (1) Web Service registries with metadata description for discovery and profile matching purposes, (2) provenance information that allows to trace history, (3) service interaction to determine the key data structures involved, (4) aspects of workflow systems, (5) import/export aspects and (6) web services wrapping.

Web service registries

Services are registered in registries for web service, web applications and tools that require some level of metadata to support both human and automated look up. Automated loop up is required by advanced web service composition support tasks, such as suggesting a suitable transformer service from a specific application domain. Well known examples of web service registries are UDDI and ebXML, but it is known that these suggestions failed for various reasons. One certainly is that they were meant to contain discipline independent semantics. Several other web service registries, such as eGov-Share, XRoads, have been built which are trying to overcome the limitations of UDDI. For CLARIN the discipline specific CLARIN meta data infrastructure will be used to provide a flexible and extensible web service metadata model which is based on about a decade of experience in the field. For interoperability purposes a semantic mapping of the metadata descriptions to minimal UDDI should be provided. This would allow us to take full advantage of UDDI integration into middleware tools. Thus UDDI has a kind of similar function as Dublin Core in the area of data resources.

These metadata descriptions may not be overloaded by process specific information such as for logging and journaling and by parameter files that are not used for search purposes. However the metadata descriptions associated with resulting resources and services should have pointers this kind of extended information.

The data model of a web service registry distinguishes between functional data, i.e. data associated with operations and non functional information about the service, i.e. service publishing data. Functional data consists of the operations that services are able to perform and the parameters that are associated with these operations. Operations may for example be specific REST or SOAP operations a service provides. These operations service interfaces may be described by interface specification languages, e.g. WADL (for REST) or WSDL (SOAP and REST). The success of these interface specification languages also depends upon middleware components that are capable of using these specifications and invoking the

service operation in the correct manner. As a consequence some service types may not be immediately usable since adequate interface specification or middleware tool support may be lacking.

The metadata descriptions of a service need to include pointers to the corresponding interface specifications.

Provenance

In processing chains at every step a new data resource or a new version of a data resource is being produced, i.e. each processing step adds information as a result of its operations. The term "Provenance" is a data resource that contains a detailed and cumulative description of the processing history. For scientific research provenance data must be documented in a sufficient manner to allow reproducibility of the results. As such provenance data is of primary importance to CLARIN as data becomes available through orchestration of services. It seems clear that in order to reproduce these processing chains all relevant data must be described. Often it will also be important for a service embedded in the processing chain to check what kind of operation was carried out at a certain step beforehand.

Although a number of provenance systems, for example PASOA and KARMA may be used to assist in capturing provenance data there is currently no specification for the provenance data within CLARIN. The minimal set of provenance data for workflows will at least contain the process type and the initial input variables specifications and values. For each individual service or tool at least the service identifier and input variables should be recorded. A pointer to provenance data is included in the CLARIN metadata description data structure, shown above, through the JournalFileProxy class.

Service interaction

From the above discussion a basic interaction model can be derived that describes service interaction.

When a user interacts with a service or workflow one of the first steps is that he/she will select one or a set of resources to operate on. For highly structured resources such as lexica or annotations the information to be used could be contained in a specific sub-structure. From a CLARIN point of view the resource will always be described by metadata which will point to provenance information when the resource is already a result of previous processing steps. When a service operation is performed the new resource should also be tagged with a CLARIN metadata description and it can be stored in a user's private workspace or institutional repository or archive. Some of the metadata for the created resource may be taken from the original resource. The result will always be a new resource: a) in the case of stand off annotation the added data is contained in a new resource closely related with the source resource; b) in the case that the results are put into the same file one would have a new version requiring a new metadata description. Provenance data, such as additional parameters are stored in the provenance file associated with the new resource.

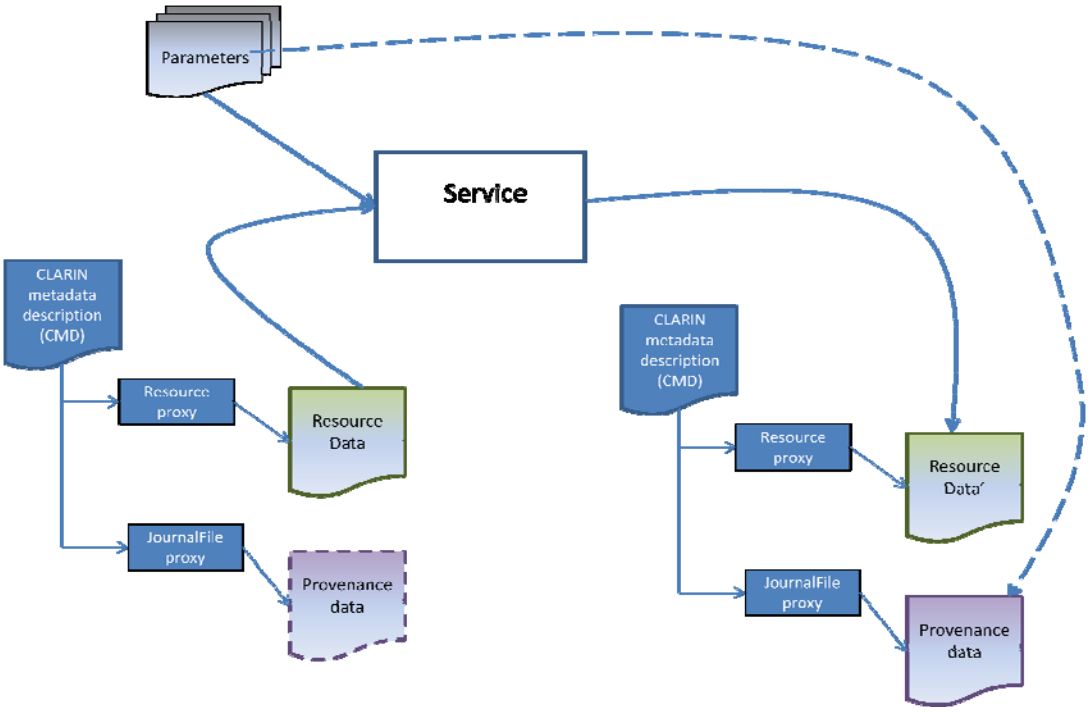

Figure 2: Basic service interaction model

The ability to gather provenance data and construct meta data specifications for each service interaction is one that must be built into the CLARIN infrastructure.

Workflow systems

In general workflow systems describe traditional processing chains which in most cases are synchronous. In most language technology applications, however, the algorithms are not smart enough to cope with all language phenomena for example, i.e. there will be steps including manual correction steps where humans need to be involved forming an element in a processing chain. Human involvement has a number of consequences. First, tools must be able to participate in workflow processes. In general they need to be able to accept a task, get the relevant data (resource (fragment) and additional parameters) and either return the result or signal cancellation of the task. This will require that tools participating in the workflow scenarios must support some type of service interface, so it partially behaves like a service. Secondly, particularly the human/tool interaction may typically take long time to complete. Our participant may just have gone on a 2 week holiday trip, which makes this type of interaction typically asynchronous. To avoid long running processes some level of persistency and dependable routing is necessary to temporarily halt execution. Persistency is typically needed to store process state, such as the current task and process variables.

Dependable routing builds on the ability to relate incoming information back to a halted process instances. For this additional information is passed to the service being invoked which supplies this information back along with the result. This incoming result message (from the point of view of the process engine) is then to be related to the correct task in the correct process instance for which persistency look up capabilities are used. Services supporting asynchronous communication must therefore have the capability to accept process instance related information as part of the payload of the

message that allows the calling process to continue the process instance. This message format needs to be specified in the CLARIN context.

Each workflow process definition itself can be associated with a metadata description and be registered as a service component. The process definition file includes all relevant information that is important for executing the process repeatedly. The metadata description of the workflow process maintains a pointer to the process definition resource. Yet we lack a proper specification in CLARIN about the content of such a process definition resource. This also needs to be studied by carrying out the various bottom-up driven studies.

Import/Export Aspects

The large variety in import/export formats is forming the greatest obstacle for allowing users to easily establishing processing chains, since this area is heavily underspecified. We need to distinguish the formal structure aspects from the linguistic encoding aspects. With respect to the structures we can see that powerful representational schemes are emerging which CLARIN should adapt if they turn out to be successful. For lexica the Lexical Markup Framework has been standardized by ISO. For complex annotation structures two similar suggestions have been worked out with Linguistic Annotation Format and Graphical Annotation Format. CLARIN will observe the further standardization and verification process. Yet we cannot speak about well-tested and proven formats, although it is agreed that formats with a strong representational power are needed to reduce the import/export complexity.

Of course still many other legacy formats are being used which brings up the question of how to do transformations. The problem is schematically indicated in the following two figures.

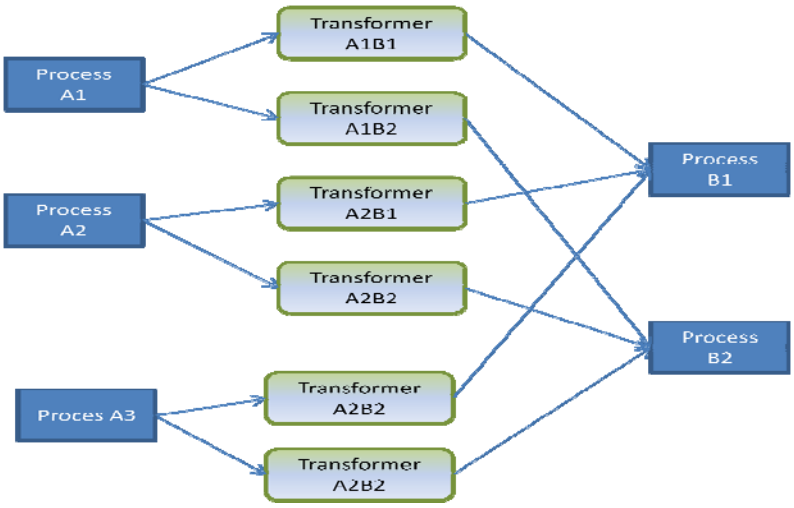

Figure 3: Transformations using point-to-point connections

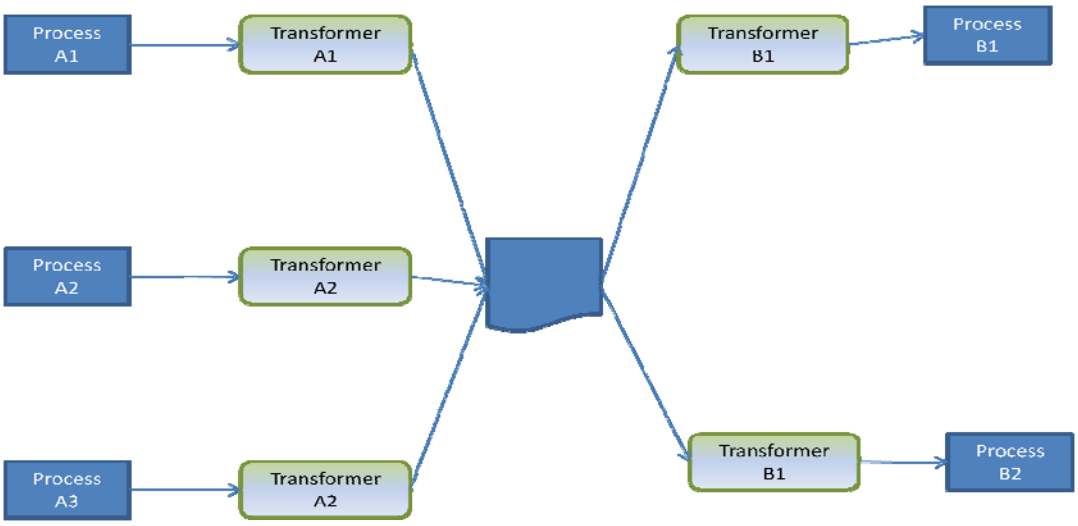

Figure 4: Transformation using a pivot model.

Services usually produce information in a proprietary format (structures and encoding) which is at the heart of interoperability problems. It is obvious that using pivot formats which are formats with a high representational power will be the only feasible way to go for CLARIN. We cannot expect that all existing code is being rewritten, but what we can expect that there will be transformers that transform legacy formats into these recognized pivot formats. Due to the complexity of the problem CLARIN does not assume that there will be only one such pivot format for a particular resource type, but already a reduction to a few would reduce the effort. Another transformer will take care that a follow-up service can consume the resource.

In addition to the structural aspects the transformation at the linguistic encoding level needs to be solved. Here we can again distinguish between two layers: (1) the terms that different linguists and their tools are using and (2) the pattern in which the encodings are written into an annotation layer. As an example we can refer to the output of tokenizers or the way syntax trees are encoded in brackets. Here we still lack a good overview requiring a deeper analysis as it is planned in WP5.

An eScience seminar of the Max Planck Society and a number of IEEE eScience conferences revealed that these problems are most difficult to solve for many disciplines. The heterogeneity is enormous and due to the scientific dynamics we will be confronted with new linguistic encoding types in future as well.

Web Service Wrapping

It is obvious that the "open" CLARIN scenario will require packing existing services to make them CLARIN web services. The code carrying out this "packing" will be called wrappers. The wrappers will include some pre- and post-processing code. The pre-processing part is responsible for reading metadata descriptions, reading the process definition file, resolving all pointers to access parameter and journaling information, claiming compute resources etc. The post-processing part is responsible for the inverse operations, i.e. create new metadata descriptions, an appropriate journaling file, write the result resource into a workspace, turning over control, etc. It is the task of WP2 in CLARIN to develop a prototypical wrapper in the preparatory phase.

The concrete implementation activities of processing chains in CLARIN will stepwise need to adopt the open CLARIN framework and from these we will improve the specifications stepwise. It is the intention in

the CLARIN preparation phase to provide a first wrapper for web services based on improved specifications.

4. Conclusions and State of Discussion

In this chapter we want to summarize the requirements and facts which already now can be stated and which therefore will be guidelines for the WS&WF work in CLARIN.

Requirements for an Open CLARIN Scenario

- Interaction with CLARIN archives/repositories that are registered in a centre registry. Workflow editors and run time workflow system components will need to interact with CLARIN repositories to access data and services. Data resources must be read from and written to private workspaces, institutional repositories or archives.
- Interaction with web services. Web services are described in the web service registry, which also contains pointers to interface specifications for invoking the web service.
    - SOAP
    - REST
    - Others, XML/RPC, REST 'like' services
- Tool interaction. In many cases no automated tools are available for performing a specific task. In a processing chain consisting of a tokenizer and a POS tagger for example it very likely that not for all languages a POS tagger may be available. In these cases it may that a human is required to manually annotate the resource.
- User interaction. Some services require additional information that is collected from the user at run time. One example is the token set to be used by a tokenizer.
- Execution of processing chains requires collaboration between various users providing several areas of expertise.
- A workflow specification language should at least support general features such as decision elements (if/then), loops (while/for each), forks and joins, exception handling.
- Synchronous and asynchronous call support. Some processes may take a very long time to complete. If a request is sent to a user to manually annotate a document then this user may well be on holiday for some time before he/she is able to comply with the request.

Facts

- No one will expect from CLARIN that we solve the difficult problems in the remaining two years which we are partly faced with. What we need is a number of useful show cases and theoretical analyses from which we can derive what needs to be done in the construction phase and how much the construction will cost.
- To achieve this it is obvious that the web services start adopting those aspects that have been consolidated within CLARIN.
    - Here in particular we refer to the metadata discussion above. The component model and the open registries for metadata of resources and tools/services need to be adopted. The whole metadata architecture has been described and discussed and can be included in the WS+WF plans. An expert group will work on categories with help of which tools and services can be described.
    - In the same way WS+WF should not bother with distributed AAI at this moment. Some development work and tests together with grid experts are being carried out currently with SLCS (Short Lived Credential Service) and we will report on the results at the end of this year.
    -
- The WS+WF work can focus on the things which are different for the linguistics domain such as
    - how to specify APIs for REST- and WSDL-based services for our domain to which the metadata records will point to

    - where and how to store the history of parameters used in workflow chains
    - how to harmonize import/export format aspects
- Metadata is stable and describes tools/services. It is not overloaded with process data. Each new result resource produced in a workflow chain will need to have a metadata description
- CLARIN needs to look into standards which are currently being worked out in ISO TC37/SC4. A deviation can only be supported in CLARIN if there are severe arguments.
- It is supported that there are multiple parallel activities: (1) there will be national activities based on national funds; (2) there needs to be at least one cross-national project that can act as show case.
- For the cross-national showcase a general purpose workflow engine should be chosen, user friendliness by providing a graphical interface is secondary. Here the ideas from Sheffield and the tests from D-SPIN should be taken up to formulate requirements.
- It is important that all activities report to Núria and Marc to ensure a convergence and to allow for a "CLARIN" branding.

Risk Statements

In underspecified areas it is always wise to make a risk analysis. So here we want to indicate a few problems to be expected:

- Every sub-group working on web-services will work independent of each other and ignoring CLARIN practices, common experiences and standards.
- The big players with long-time experience in our field do not move at all to anticipate the "open" CLARIN framework. Open here means in particular to support open registries and to not expect that a tool builder will need to modify existing services to be CLARIN compliant. This does however not imply that each service is automatically CLARIN compatible. Conversion methods must be delivered dealing with input/output specifications.
- The expectation to come to one single pivot format for annotation structures may turn out to be not feasible. Almost orthogonal is the specification of the linguistic encoding covered in the annotation structure. This has to be defined within parameter descriptions. Current metadata can only be of limited use since they will not go further as to specify mime types, typical tag sets and algorithms used.

5. References

Acronyms

Reference	Abbreviation of	Link
[APA]	Alliance for Permanent Access	http://www.alliancepermanentaccess.eu
[ARK]	Archival Resource Key	http://www.cdlib.org/inside/diglib/ark/
[BPEL]	Business Process Execution Language	http://en.wikipedia.org/wiki/WS-BPEL
[CGN]	Corpus Gesproken Nederlands	http://lands.let.kun.nl/cgn/
[DAM-LR]	Distributed Access Management for Language Resources	http://www.dam-lr.eu/
[DC]	Dublin Core	http://dublincore.org/
[DCAM]		http://dublincore.org/documents/abstract-model/
[DC-DS-XML]		http://dublincore.org/documents/dc-ds-xml/
[DC-TEXT]		http://dublincore.org/documents/dc-text/
[DEISA]	Distributed European Infrastructure for Supercomputing Applications	http://www.deisa.eu/
[DFKI]		http://www.language-archives.org/archive/dfki.de
[DOBES]		

[DOBES]	Dokumentation Bedrohter Sprachen	http://www.mpi.nl/dobes
[DOI]	Digital Object Identifier	http://www.doi.org/
[EAD]	Encoded Archival Description,	http://en.wikipedia.org/w/index.php?title=Encoded_Archival_Description&oldid=250469911
[ebXML]	e-business XML	http http://www.ebxml.org/
[EGEE]	Enabling Grids for E-sciencE	http://www.eu-egee.org/
[e-IRG]	e-Infrastructure Reflection Group	http://www.e-irg.eu/
[ELDA UC]	Universal Catalogue	http://universal.elra.info/
[ENABLER]		http://www.ilsp.gr/enabler/
[ESF]	European Science Foundation Second Learner Study	http://books.google.de/books?id=g292tXMX4tgC&pg=PA1&lpg=PA1&dq=esf+Second+learner+perdue&source=bl&ots=WKi3GUQQP6&sig=n7QSWy3StXvD06nMfAzY7GBbm9w&hl=de&sa=X&oi=book_result&resnum=3&ct=result#PPP1,M1
[FIDAS]	Fieldwork Data Sustainability Project	http://www.apsr.edu.au/fidas/fidas_report.pdf
[HS]	Handle System	http://www.handle.net/
[ICONCLASS]		http://en.wikipedia.org/wiki/Iconclass62
[INTERA]	Integrated European language data Repository Area	http://www.mpi.nl/intera/
[ISOcat]		http://www.isocat.org
[LAF]	Linguistic Annotation Framework	
[LMF]	Lexical Markup Framework	http://en.wikipedia.org/w/index.php?title=Lexical_Markup_Framework&oldid=255448197
[METATAG]		http://en.wikipedia.org/w/index.php?title=Meta_element&oldid=256779491
[METS]	Metadata Encoding and Transmission Standard	http://en.wikipedia.org/wiki/METS
[MILE]		http://www.mileproject.eu/
[MPEG7]		http://en.wikipedia.org/w/index.php?title=MPEG-7&oldid=241494600
[NLSR]		http://registry.dfki.de/
[OAIS]	Open Archival Information System	http://en.wikipedia.org/wiki/Open_Archival_Information_System
[OASIS]	Organization for the Advancement of Structured Information Standards	http://www.oasis-open.org/
[ODD]	One Document Does all	http://www.tei-c.org/wiki/index.php/ODD
[OLAC]	Open Language Archives Community	http://www.language-archives.org/
[PMH]	Protocol for Metadata Harvesting	http://www.openarchives.org/pmh/
[REST]	Representational State Transfer	http://www.ics.uci.edu/~fielding/pubs/dissertation/rest_arch_style.htm
[SCHEMAS]		http://www.schemas-forum.org/
[SOAP]	Simple Object Access Protocol	http://www.w3.org/TR/soap12-part1/
[SRU]	Search/Retrieve via URL	http://www.loc.gov/standards/sru/
[SRW]	Search/Retrieve Web Service	http://en.wikipedia.org/wiki/Search/Retrieve_Web_Service
[TEI]	Text Encoding Initiative	http://www.tei-c.org/

[UDDI]	Universal Description Discovery and Integration	http://www.oasis-open.org/committees/uddi-spec/doc/tcspecs.htm
[WADL]	Web Application Description Language	https://wadl.dev.java.net/wadl20090202.pdf
[WSDL]	Web Services Description Language	http://www.w3.org/TR/wsdl20
[Z39.50]		http://en.wikipedia.org/wiki/Z39.50

Related Documents

[CLARIN_NEWS_4]	CLARIN news letter Dec 2008	http://www.clarin.eu/filemanager/active?fid=231
[CLARIN_WS_NOTE]	CLARIN note on web services	http://www.clarin.eu/filemanager/active?fid=270
[D-SPIN_PRES]	D-SPIN workshop report and Presentations	http://www.clarin.eu/wp2/wg-26/wg-26 documents/web-service-presentations-at-the-wp2-workshop-in-oxford