On the Way to a Legal Sharing of Web Applications in NLP

Victoria Arranz, Olivier Hamon
ELDA/ELRA
55-57 rue Brillat Savarin, 75013 Paris, France
E-mail: {arranz ; hamon}@elda.org

Abstract

For some years now, web services have been employed in Natural Language Processing (NLP) for a number of uses and within a number of sub-areas. Web services allow users to gain access to distant applications without having the need to install them on their local machines. A large paradigm of advantages can be obtained from a practical and development point of view. However, the legal aspects behind this sharing should not be neglected and should be openly discussed so as to understand the implications behind such data exchanges and tool uses. In the framework of PANACEA, this paper highlights the different points involved and describes the work done in order to handle all the legal aspects behind those points.

Keywords: IPR issues; Web Services; Language Resources

1. Introduction

For some years now, web services have been employed in Natural Language Processing (NLP) for a number of uses and within a number of sub-areas. Web services allow users to gain access to distant applications without having the need to install them on their local machines. A large paradigm of advantages can be obtained from a practical and development point of view. However, the legal aspects behind this sharing should not be neglected and should be openly discussed so as to understand the implications behind such data exchanges and tool uses.

A number of European initiatives have been looking at the legal aspects of data sharing these past few years (such as META-NET[1] and CLARIN[2]), but this has been done from a repository and language resource (LR) point of view (Choukri et al., 2012), clearing out the licensing conditions between the LR centre and its potential user.

For instance, with the advice and collaboration of legal experts, META-NET has defined a number of licenses which allow for the sharing of language resources in the above-mentioned scenario. A not-to-be-neglected big concern of the different initiatives has been to ensure not only the right to read the content of a LR, but also to transform it and to share it, together with any derivatives, to interested third parties. Bearing this in mind, a variety of licensing user cases have been defined and later implemented into license templates that members/users can choose from, according to their needs.

CLARIN has also worked on designing a licensing and authorization schema for their network of digital repositories (Lindén, 2010).

However, what happens when licensing aspects need to go further than this one-to-one LR acquisition? Despite the wide coverage of the licensing schemas proposed by these initiatives, none of them has put into place a licensing schema which covers the multiple needs of a web service based LR production platform, i.e. a factory of language resources.

In the EU-FP7 PANACEA[3] project (7FP-ITC-248064), a large part of the work effort is devoted to the development of a platform (Poch et al., 2012) dealing with web services and workflows (i.e. chains of web services). One of the main objectives is to allow developers to share their applications without having to give any access to the source code or to an outdated executable. From a user's point of view, the fact of being able to run an application (or a combination of them) and obtain its output without the drawback of dealing with any installation issues is already a big advantage. However, in order to reach that state with all uses clearly defined, a number of questions need to be answered.

During the project and the setting up of web services, unavoidable questions have arisen based on the different needs of the platform. Some of these needs concern the following:

- The input to the web services and workflows;
- The temporary data and storage on servers;
- The usage of applications;
- The implications on the development;
- The output data;
- The different licenses and disclaimers.

In this paper, we highlight those different points while trying to solve all legal aspects involved in as simple a way as possible. First, we summarize the context of web services and workflows from a user's point of view. Next, we focus on the different challenges regarding the Intellectual Property Rights (IPR) when sharing web services and workflows. Finally, we draw some conclusions regarding this sharing of web services and workflows in a legal framework.

2. From Web Services and Workflows to Legal Web Applications

Deploying a web service is a handy way to share an application without dealing with any installation, download and maintenance issues. However, this does not mean that users can play with applications without taking into consideration the usage rights behind them. In what regards workflows, users need to be aware of the same

[1] http://www.meta-net.eu

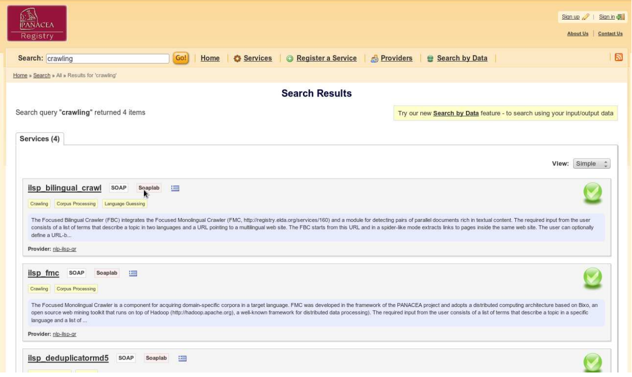
[2] http://www.clarin.eu

[3] Platform for Automatic Normalized Annotation and Cost-Effective Acquisition of Language Resources for Human Language Technologies.

limitations but at a different level: Intellectual Property Rights (IPR) issues exist for each web service of the chain and for each language resource obtained and produced!
In PANACEA, web services are collected within a catalogue[4] named the *Registry*, which is based on the BioCatalogue[5] tool (Belhajjame et al., 2008). This registry makes web services visible to the community as well as allows providers to add new web services that any interested user may wish to use. Therefore, users can:

a) browse the different web services available,
b) provide new web services and
c) use available web services, either within a workflow or for a single usage, depending on their needs.

Figure 1 provides a snapshot of the Registry.



Home » Search » All » Results for 'crawling'

Search Results

Search query "crawling" returned 4 items

Try our new Search by Data feature - to search using your input/output data

Services (4)

ilsp_bilingual_crawl

The Focused Bilingual Crawler (FBC) integrates the Focused Monolingual Crawler (FMC, http://registry.elda.org/services/160) and a module for detecting pairs of parallel documents rich in textual content. The required input from the user consists of a list of terms that describe a topic in two languages and a URL pointing to a multilingual web site. The FBC starts from this URL and in a spider-like mode extracts links to pages inside the same web site. The user can optionally define a URL-b...

Provider: nlp-ilsp-gr

ilsp_fmc

The Focused Monolingual Crawler is a component for acquiring domain-specific corpora in a target language. FMC was developed in the framework of the PANACEA project and adopts a distributed computing architecture based on Bixo, an open source web mining toolkit that runs on top of Hadoop (http://hadoop.apache.org), a well-known framework for distributed data processing). The required input from the user consists of a list of terms that describe a topic in a specific language and a list of ...

Provider: nlp-ilsp-gr

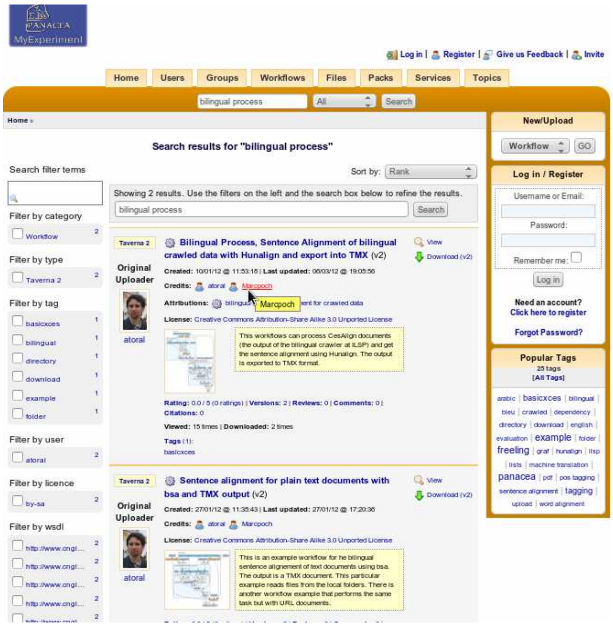
ilsp_deduplicatormd5

Figure 1: PANACEA *Registry* snapshot

Likewise, workflows are also collected within a catalogue[6], based on the myExperiment[7] tool (De Roure et al., 2008), and they offer the same features as the above-mentioned web services, i.e., browsing, contributing with new workflows or using them through the Taverna software (Hull et al., 2006).
Figure 2 illustrates this catalogue and the manner the workflows are listed.
With regard to the legal situation of the tools within these catalogues, a clear legal framework is being defined. This will allow for IPR issues to be clearly stated with regard to the different content types within such catalogues and within the PANACEA platform as a whole. This legal framework has a double role, both informative and active, thus ensuring that users are:

- well aware of the rights behind the applications integrated in the web services,
- well equipped with the necessary documents (licenses, disclaimers, description documents) they may need for a usage of the services and the platform as user-friendly and as simple as possible.

Furthermore, all this is also applicable to the rights and conditions behind the input and output data that circulate around the PANACEA platform.


PANACEA
MyExperiment
Log in | Register | Give us Feedback | Invite
Home
Users
Groups
Workflows
Files
Packs
Services
Topics
Search
Home »
Sort by: Rank
Search
Taverna 2
View
Download (v2)
New/Upload
Workflow GO
Log in / Register
Log in
Marcpoch
Popular Tags
25 tags
[All Tags]
arabic | basicxces | bilingual | bleu | crawled | dependency | directory | download | english | evaluation | example | folder | freeling | graf | hunalign | ilsp | lists | machine translation | panacea | pdf | pos tagging | sentence alignment | tagging | upload | word alignment

bilingual process | All

Search results for "bilingual process"

Search filter terms

Filter by category
☐ Workflow 2

Filter by type
☐ Taverna 2 2

Filter by tag
☐ basicxces 1
☐ bilingual 1
☐ directory 1
☐ download 1
☐ example 1
☐ folder 1

Filter by user
☐ atoral 2

Filter by licence
☐ by-sa 2

Filter by wsdl
☐ http://www.cngl... 2
☐ http://www.cngl... 2
☐ http://www.cngl... 2
☐ http://www.cngl... 2
☐ http://www.cngl... 2

Showing 2 results. Use the filters on the left and the search box below to refine the results.

bilingual process

Bilingual Process, Sentence Alignment of bilingual crawled data with Hunalign and export into TMX (v2)

Original Uploader: atoral

Created: 10/01/12 @ 11:53:18 | Last updated: 06/03/12 @ 19:05:56

Credits: atoral Marcpoch

Attributions: bilingua... ...ent for crawled data

License: Creative Commons Attribution-Share Alike 3.0 Unported License

This workflows can process CesAlign documents (the output of the bilingual crawler at ILSP) and get the sentence alignment using Hunalign. The output is exported to TMX format.

Rating: 0.0 / 5 (0 ratings) | Versions: 2 | Reviews: 0 | Comments: 0 | Citations: 0

Viewed: 15 times | Downloaded: 2 times

Tags (1): basicxces

Sentence alignment for plain text documents with bsa and TMX output (v2)

Original Uploader: atoral

Created: 27/01/12 @ 11:35:43 | Last updated: 27/01/12 @ 17:20:36

Credits: atoral Marcpoch

License: Creative Commons Attribution-Share Alike 3.0 Unported License

This is an example workflow for he bilingual sentence alignement of text documents using bsa. The output is a TMX document. This particular example reads files from the local folders. There is another workflow example that performs the same task but with URL documents.

Username or Email:
Password:
Remember me: ☐

Need an account? Click here to register

Forgot Password?

Figure 2: PANACEA *myExperiment* snapshot

3. Establishing a Legal Framework

Some of the legal issues linked to the usage of web services and workflows are quite challenging. In particular, in PANACEA, the issues are related to the automatic production of language resources within a web platform. For instance, we have to handle the data coming from the Internet, the combination of data and software *via* web services, the combination of different web services through a workflow or the management of derivative products.

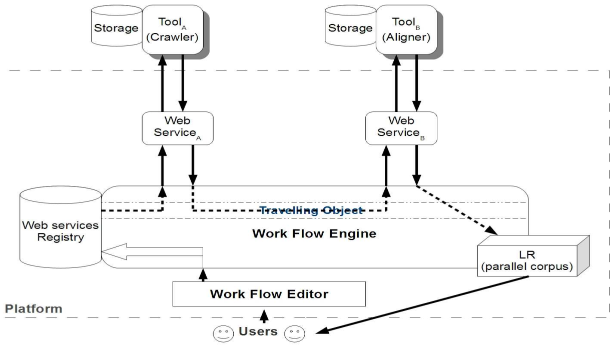


Figure 3: Starting point for the needs

Figure 3 depicts the starting point of the study, with a very simple definition of the PANACEA platform structure. This structure already points out the kind of parameters to take into consideration (e.g., tools, data, web services, workflows, catalogues, output LRs, users…) and it also sets up the basics for the multiple element relationships and combinations that can take place and that will require clearing out. Once the different elements were analysed, we were faced with Figure 3 converted into Figure 4, the latter providing us with all the questions and worries that the platform users may run into (points here derived from discussions with the actual project partners and potential platform users).

[4] http://registry.elda.org
[5] http://www.biocatalogue.org
[6] http://myexperiment.elda.org
[7] http://www.myexperiment.org

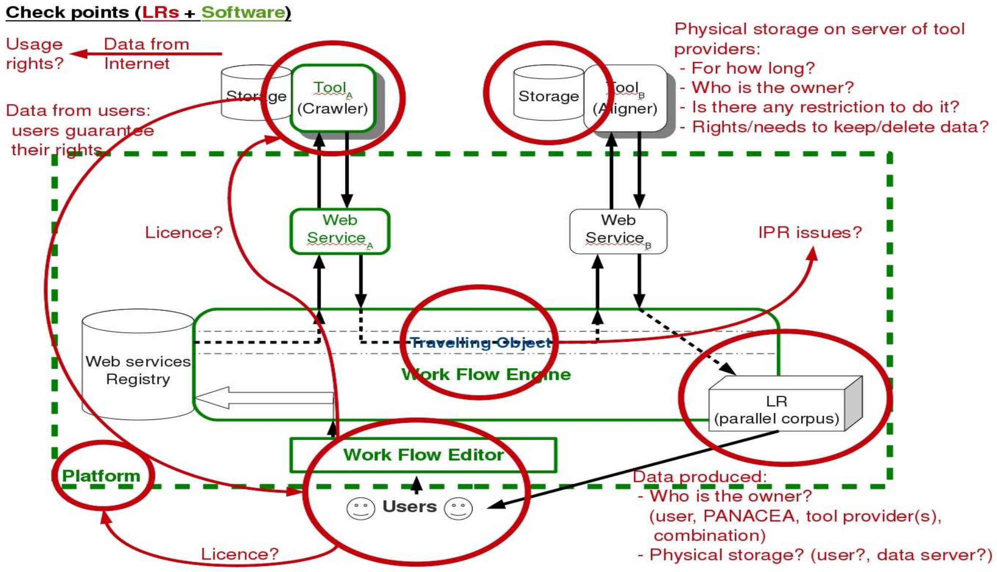

Check points (LRs + Software)
Usage rights?
Data from Internet
Storage
Tool_A (Crawler)
Storage
Tool_B (Aligner)
Physical storage on server of tool providers:
- For how long?
- Who is the owner?
- Is there any restriction to do it?
- Rights/needs to keep/delete data?
Data from users: users guarantee their rights
Licence?
Web Service_A
Web Service_B
IPR issues?
Web services Registry
Travelling Object
Work Flow Engine
LR (parallel corpus)
Work Flow Editor
Platform
Users
Data produced:
- Who is the owner? (user, PANACEA, tool provider(s), combination)
- Physical storage? (user?, data server?)
Licence?

Figure 4: Legal concerns to be handled within the platform

These questions and worries (that within the figure are referred to as check points) regard the following information:

- The input to the web services and workflows;
- The temporary data and storage on servers;
- The usage of applications;
- The implications on the development;
- The output data;
- The different licenses and disclaimers.

The following sections elaborate on all these points in more detail.

3.1 Input to Web Services and Workflows

Two main types of data sources may be provided by a user to a web service: either data coming from the Internet (for instance, when using web services for crawling), or material which is already "owned" by the user (or rather, "in the user's hands"). In either case, the usage of the data is restricted by some rights.

These rights are generally well-known when talking about already available language resources. However, this is a completely different story when facing Internet sources. For the latter, the user should make sure that (s)he has the right to crawl such data. In order to do so, (s)he may need to obtain an authorization to use the material when these are to be employed, for instance, in the training of a commercial application.

3.1.1. Case Study: Internet Data Crawling

In the framework of PANACEA, we have carried out a case study on crawled data so as to:

- analyse quantitatively the full implications behind its use, in particular with the perspective of future massive data handling;
- provide the means for users to do so themselves;
- describe the procedure and execution cost clearly.

As anticipated, the task has been very demanding and time-consuming since the procedure consists on the following steps:

- Locating all sources and contact points: this is relatively simple when planning to approach a few sources, but very costly when considering hundreds/thousands of them.
- Studying terms and conditions: a web site may contain public data which can already be used. We need to see whether data use and future distribution (of the data or any derivative product) are at all possible.
- Approaching providers: once established that data sources need to be approached (on a case per case basis), the efforts required may vary from a few simple exchanges (for providers willing to contribute to R&D, for instance) to endless discussions to define data use and conditions, among others.

Thus, it can be concluded that the complexity behind this "data usage right obtaining" lies on the following parameters:

- It is source-dependent: complexity may raise if a particular institutional source is hard to reach and likewise for a blog-data owner (some blog owners change blogs very frequently and previous blogs are left as "orphans" on the cyber space).
- Negotiation duration is generally long: from a study conducted on the authorization discussions conducted within the project we concluded that these could last as little as 1 day or go up to almost 1 year. Table 1 provides the exact details for this analysis, with regard to both the monolingual and the bilingual (or parallel) data crawled within the project. The average duration ranged between 66 days (for monolingual data) and 176 (for the parallel one). Multilingual sources have proven to be more complicated and longer to negotiate. Rather often, the reason for this is that the data owners are more sceptical about sharing it for the sake of research, being aware of the higher production cost and, as a

consequence, potential value of the data. In any case, we are pleased to say that a large number of data providers have agreed to share their web resources with the project and with the R&D community[8]. These corpora will be available shortly through the ELRA Catalogue[9].

- Difficult access to some institutions & blogs: finding a contact point or getting through to the right person may be complicated. In the case of some Web sites, contact takes place exclusively through some forms to fill in. Reaching a human with a full name may be far from trivial!
- Need to be reassured of no ownership right infringement: many data owners fear a misuse of their data. Unfortunately, some of them (a smaller number) refuse to allow data use for a usage different from that it was intended, in particular when it implies data manipulation (such as cleaning or editing in order to generate aligned corpora). In these cases, data providers are explained what the data will be used for, in the sense of "for language engineering", without any further interfering or tampering with their content. For example, organisations using their Web sites for the dissemination of their political activities may be wary of the potential use of their data content.
- Need to understand data use: "what is HLT?" a large number of users has not heard about Human Language Technologies, which means that some technology education is required during the data authorisation discussions.

Duration (in days)	**Monolingual data**	**Parallel data**
Shortest	1	8
Longest	339	344
Average duration	66	176

Table 1: Negotiation duration

Last but not least, in order to allow potential data-crawling users to negotiate themselves the right to use such data, appropriate authorisation letter templates are available that users can easily customise and have signed for their own purposes.

The complexity of such tasks has also been confirmed with collaborating projects like ACCURAT (Tadić, 2011), who rather decided that the endeavouring of such negotiations needed to be left up to the final user.

3.1.2. Data Provided by Platform Users

Regarding sources provided by the PANACEA platform user, (s)he must guarantee such rights (in an implicit way with the PANACEA platform). This is established as such within the Terms of Use of the platform. It is the sole responsibility of the input provider to check and ensure that (s)he has the right to use the input data (s)he provides to the platform.

As a reference, in the case of META-NET, data owners are asked to sign a depositor's agreement[10], given that the META-NET repository (META-SHARE[11]) carries out storing and sharing activities with such resources, something the PANACEA platform does not foresee.

Both ELRA[12] and LDC[13], as institutions with a long experience in the sharing of LRs at a European and American level, respectively, have executed such kind of activities as their main role for many years now. Both of them hold Distribution licenses that the data providers sign with the distribution entity to grant them the right to share these data.

All these points and other related ones are being duly indicated within the PANACEA platform to avoid any misunderstanding. The users will be provided with clear statements so as to know how to handle every scenario.

3.2 Temporary Data and Storage on Servers

The usage of web services and workflows implies the storage of data on the servers where the web service(s) is/are located. These are generally referred to as temporary data, e.g. source data sent by the user, results data sent back to him/her and potential intermediate processed data. It may seem obvious that such data should not remain on a web service server, simply because the web service provider is not the owner and does not have the right to use the data. However, it may be useful for the user to keep the data on the distant server for a certain time, even after the process is over (for instance, to retrieve the data should the user lose it). The duration of storage is then the main parameter, although users may choose not to send proprietary data stored on a peer server.

In that regard, the PANACEA platform displays a temporary-file deletion disclaimer (on the catalogue of web services and thus for each web service) stipulating that "Temporary files will not be used by anyone but the actual user of the input data that generated them." and that temporary files would be automatically deleted from the server after a certain number of days, free to the web service provider to indicate how many. Therefore, service providers must guarantee the privacy of the data used.

An actual implementation of such disclaimer for one of the web services within the platform reads as follows:

Temporary files deletion disclaimer

Temporary files may be generated by the various processes for their needs and operations. Temporary files will not be used by anyone but the actual user of the input data that generated them. This is part of our data protection policy aimed at safeguarding the owner rights on the data travelling through the web services. Temporary files will be automatically deleted from the system after 2 days, even if they are not accessible to anyone but the actual user. It is the sole responsibility of the input provider to check and ensure that (s)he has the right to use the input data provided to the platform. No access or use of the temporary files will be allowed other than stipulated in this disclaimer.

[8] The full list of kind contributors can be found at http://panacea-lr.eu/en/links/acknowledgements/

[9] http://catalog.elra.info

[10] http://www.meta-net.eu/meta-share/licenses

[11] http://www.meta-share.eu

[12] http://www.elra.info

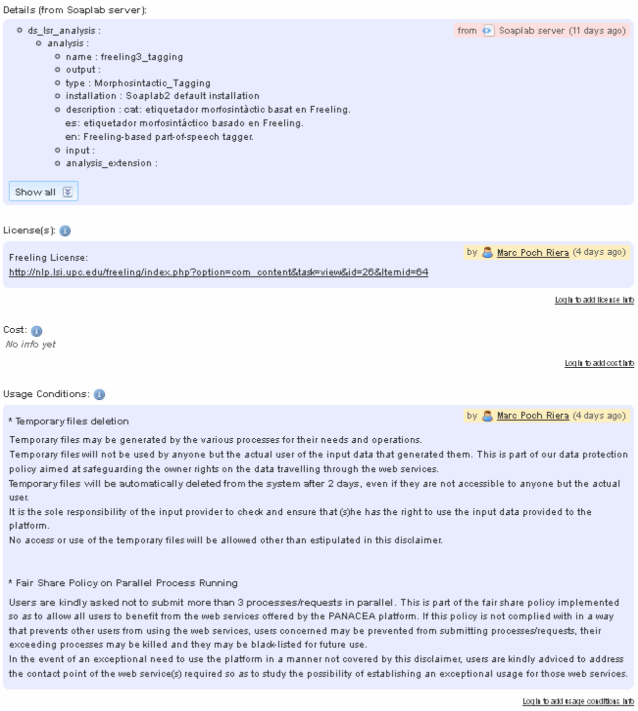
[13] http://www.ldc.upenn.edu/

3.3 Usage of Applications

Regarding the usage of an application integrated in a web service, there also exists a strong relationship between the web service provider and the user. Indeed, when a web service provider is not the owner of the application, he must guarantee that the usage of the application(s) provided respects the usage rights of the application owner and that all IPR issues have been cleared between them. The user shall consider it so.

In particular, the provider must follow the redistribution specification in the application license. For that purpose, the web service provider will offer all relevant legal documentation on the platform (on the space allocated for this purpose within each web service page), comprising application license or link to it, usage restrictions/conditions documentation (if relevant), etc. Figure 5 below illustrates how this information is being provided within the platform. In this case, the web service provider is giving the URL pointing to the service source license (together with other information such as disclaimers for user conditions)[14].

Needless to say that web services may also have a fee. This can also be specified on the web service page, together with the type of license to be signed. However, at this stage of the project, the handling of such payments has not been fully managed, but it is planned for the final version of the platform.



Details (from Soaplab server):

- ds_lsr_analysis :
  - analysis :
    - name : freeling3_tagging
    - output :
    - type : Morphosintactic_Tagging
    - installation : Soaplab2 default installation
    - description : cat: etiquetador morfosintàctic basat en Freeling.
      es: etiquetador morfosintáctico basado en Freeling.
      en: Freeling-based part-of-speech tagger.
    - input :
    - analysis_extension :

License(s):

Freeling License:
http://nlp.lsi.upc.edu/freeling/index.php?option=com_content&task=view&id=26&Itemid=64

Cost:
No info yet

Usage Conditions:

* Temporary files deletion

Temporary files may be generated by the various processes for their needs and operations.
Temporary files will not be used by anyone but the actual user of the input data that generated them. This is part of our data protection policy aimed at safeguarding the owner rights on the data travelling through the web services.
Temporary files will be automatically deleted from the system after 2 days, even if they are not accessible to anyone but the actual user.
It is the sole responsibility of the input provider to check and ensure that (s)he has the right to use the input data provided to the platform.
No access or use of the temporary files will be allowed other than estipulated in this disclaimer.

* Fair Share Policy on Parallel Process Running

Users are kindly asked not to submit more than 3 processes/requests in parallel. This is part of the fair share policy implemented so as to allow all users to benefit from the web services offered by the PANACEA platform. If this policy is not complied with in a way that prevents other users from using the web services, users concerned may be prevented from submitting processes/requests, their exceeding processes may be killed and they may be black-listed for future use.
In the event of an exceptional need to use the platform in a manner not covered by this disclaimer, users are kindly adviced to address the contact point of the web service(s) required so as to study the possibility of establishing an exceptional usage for those web services.

from Soaplab server (11 days ago)
Show all
by Marc Poch Riera (4 days ago)
Log in to add license info
Log in to add cost info
by Marc Poch Riera (4 days ago)
Log in to add usage conditions info

Figure 5: Sample of licensing information for web services

As it can also be observed in Figure 5, from a practical application-usage point of view, one further disclaimer has been put into place. This states the fair use of the platform and delimits the number of processes that can be submitted in parallel. The exact details are as follows:

Fair Share Policy on Parallel Process Running

Users are kindly asked not to submit more than 3 processes/requests in parallel. This is part of the fair share policy implemented so as to allow all users to benefit from the web services offered by the PANACEA platform. If this policy is not complied with in a way that prevents other users from using the web services, users concerned may be prevented from submitting processes/requests, their exceeding processes may be killed and they may be black-listed for future use.

In the event of an exceptional need to use the platform in a manner not covered by this disclaimer, users are kindly adviced to address the contact point of the web service(s) required so as to study the possibility of establishing an exceptional usage for those web services.

3.4 Implications on the Development

When dealing with workflows, data are not only stored on the server of the different web services, but are also "traveling" between web services. To guarantee the privacy of the data transferred from one web service to another, the transfer protocol must be secured so as to avoid any security bridge. Indeed, data going from one server to another (e.g. in the case of a workflow process) or from a client machine to a server (e.g. in the case of a single web service process) should be secure enough so as not to be corrupted or retrieved by a third user. In PANACEA, this process is secured by using SOAP[15] (Simple Object Access Protocol), which allows to reach a sufficient level of security since SOAP transports data using both SMTP and HTTP (and potentially HTTPS).

3.5 Output Data

The owner of the web services and workflows results may be subject to question. From the different entities who are involved in the process, that is, the user, the web service provider or the workflow provider, all of them may seem to have some rights over the resulting output. However, the context should be the same as the one faced when dealing with applications on a one-to-one basis. The difference lies on the complexity imposed by the chaining of applications and data, which must be supervised by a clear stating of usage rights and limitations within the platform. This means the following:

- *With regard to the usage of web services*: these rights and limitations are stated on the page of the web service itself (cf. Figure 5 for an example), by means of:
  - License(s);
  - Temporary files deletion disclaimer;
  - Fair Share Policy on Parallel Process Running.
- *With regard to the usage of workflows* (cf. Figure 6 for an example[16]): these represent a chain of web services and so as to use them, the rights and limitations for every component web service need to be respected. In order to ensure this, each workflow will provide this information on its page, as it is currently done for web services. Moreover, other relevant legal information will

[14] For further reference, this particular service can be found at http://registry.elda.org/services/237.

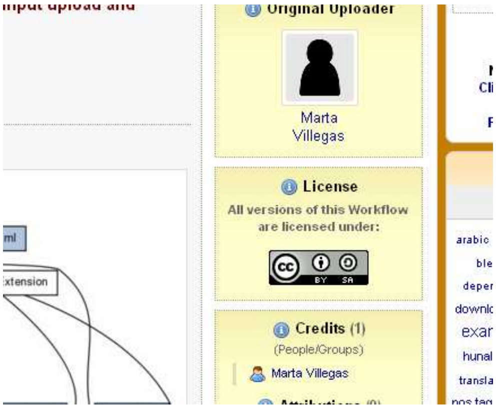
[15] http://www.w3.org/2002/07/soap-translation/soap12-part0.html

[16] This workflow can be found at http://myexperiment.elda.org/workflows/46.

be also displayed (e.g., disclaimers).

Yet, the case of workflows is a particular one as it offers a web service combinatorial richness which needs to be secured on the legal aspects side. For example, when a workflow user wishes to change one of the web services within the workflow to a different one, or to one of his own, this means that the relevant license needs to be agreed upon too. This has not been implemented in the platform yet, but work is planned to allow this license switching.

Last but not least, a workflow has a workflow owner and his sharing of the workflow is done under certain rights and conditions too. This is so indicated on the workflow page, as we can see in Figure 6 with the CC license specified for that purpose.



All versions of this Workflow are licensed under:

Figure 6: Sample of legal information for workflows

4. Conclusions

This paper aims at providing an overview of the legal implications behind a platform of web services and workflows where data and applications from different sources and with different destinations co-habit and interact. The paper describes the context and principles considered, together with the solutions and measures currently being implemented.

As it can be observed, the different needs of the platform call for different solutions. These needs have been studied in detail within the project with the help of legal advice. All aspects are being currently defined and implemented within the platform so as to make sure users find all necessary legal reference when intending to use the platform.

Such definition is part of a larger exploitation plan, which also foresees the future of the platform in its different case scenarios.

The legal framework defined in this work goes through a number of issues which represent the “questions and worries” that any potential user of the platform may bump into. These issues look into web services, workflows and their input and output data, as well as aspects concerning temporary data, traveling objects and security. For that purpose, we detail the restrictions, licenses and disclaimers established for the applications, web services and workflows within their catalogues, as well as for the different data handled.

5. Acknowledgements

This work has been carried out under the financing of the EC FP7 project PANACEA, Grant Agreement 248064. We would like to thank our colleagues Jérémy Leixa and Priscille Schneller for the hard work done within the PANACEA project.

Our thanks also to the other PANACEA partners, for their opinions and feedback.

Once again, we would also like to thank the kind contribution of all those generous data providers who have supported PANACEA and the HLT community by donating their data. Their names and contributions are listed in http://panacea-lr.eu/en/links/acknowledgements/.

6. References

Belhajjame, K., Goble, C., Tanoh, F., Bhagat, J., Wolstencroft, K., Stevens, R., Nzuobontane, E., McWilliam, H., Laurent, T., Lopez, R. (2008). BioCatalogue: A Curated Web Service Registry for the Life Science Community. In *Microsoft eScience Conference.*

Choukri, K., Piperidis, S., Tsiavos, P., Patrikakos, T., Gavrilidou, M., Weitzmann, J. H., (2012). META-SHARE: Licenses, Legal, IPR and Licensing issues, T4ME Deliverable 6.1.3, February 2012.

De Roure, D., Goble, C., Stevens. R. (2008). The Design and Realisation of the myExperiment Virtual Research Environment for Social Sharing of Workflows. *Future Generation Computer Systems*, vol. 25, pp. 561-567.

Hull, D., Wolstencroft, K., Stevens, R., Goble, C., Pocock, M., Li, P., Oinn. T. (2006). "Taverna: a tool for building and running workflows of services". In *Nucleic Acids Research, vol. 34, iss. Web Server* issue, pp. 729-732.

Lindén, K. (2010). "A report including Model Licensing Templates and Authorization and Authentication Scheme". CLARIN deliverable D7S-2.1.

Poch, M., Toral, A., Hamon, O., Quochi, V., Bel, N. (2012). "Towards a User-Friendly Platform for Building Language Resources based on Web Services". In *Proceedings of LREC 2012*, Istanbul, Turkey.

Tadić, M. (2011). "Analysis on IPR of project results". ACCURAT Deliverable D6.7. December, 2011.