



D5.3-6

Handling IPR and ethical issues

Document information

Title	Handling IPR and ethical issues
ID	CLARINPLUS-D5.3-6
Author(s)	Hanne Fersøe, Sussi Olsen
Contractual Delivery Date	2017-06-30
Actual Delivery Date	2017-06-29
Distribution	Public
Document status in workplan	Deliverable

Project information

Project name	CLARIN-PLUS
Project number	676529
Call	H2020-INFRADEV-1-2015-1
Duration	2015-09-01 – 2017-08-31
Website	www.clarin.eu
Contact address	contact-clarinplus@clarin.eu

Table of contents

1	Executive Summary	2
2	Introduction	2
2.1	The Challenges to be Handled.....	2
2.1.1	IPR Issues	2
2.1.2	Legal Issues Concerning Personal Data	3
2.1.3	Authentication and Authorization.....	3
2.2	Legal Issues Committee	3
2.2.1	Legal Information Platform	4
2.2.2	CLIC White Papers.....	4
3	Licensing	4
3.1	Roles.....	4
3.2	Deposition License Agreements.....	5
3.3	End-user License Agreements	5
3.4	CLARIN License Categories	5
3.4.1	PUB.....	5
3.4.2	ACA.....	5
3.4.3	RES	6
3.4.4	Additional Requirements	6
3.5	CLARIN License Category Calculator	6
4	Terms of Service and Privacy Policy	6
5	Lessons Learnt	7
6	Conclusion	7
	References	7

1 Executive Summary

This best practise document gives a brief overview of the regulatory and contractual framework of the CLARIN infrastructure.

The term *regulatory framework* refers to the set of legal provisions that cover the protection of copyright and of personal data. These provisions are expressed in national legislation throughout Europe, but they are not identical from one country to the next.

The term *contractual framework* refers to the set of contracts or license agreements developed in CLARIN for the members to use when adding content to the infrastructure.

A much more thorough introduction to the regulatory framework can be found at the CLARIN Legal Information Platform available at <https://www.clarin.eu/content/legal-information-platform>, edited by Paweł Kamocki and Erik Ketzan.

2 Introduction

The CLARIN mission is to give researchers in the Humanities and Social Sciences access to services offering language-based resources, technology and expertise. This mission is implemented by providing a research infrastructure at a European level. The infrastructure is based on nine basic principles, called pillars laid out in the CLARIN mission statement at <https://www.clarin.eu/content/mission-and-strategy>.

This document focuses on pillar *b*, Legal Issues:

No other restrictions on the use for research purposes should apply than those following from confidentiality, privacy or ethical considerations. The rights and legitimate interests of data owners should be protected at all times.

The purpose of this document is to provide a best practice guide to the handling of these issues when CLARIN ERIC members intend to provide content for use through their national CLARIN centre(s).

2.1 The Challenges to be Handled

2.1.1 IPR Issues

IPR stands for *Intellectual Property Right*, which covers the concepts copyright, related rights and industrial property.

The basic rule in intellectual property law is that all original material is protected by law. The legislation, which regulates these matters, is not identical throughout European and other countries, but the basic principle of data protection is implemented throughout. The owner of the data has, by law, the right to decide who can have lawful access to the material and for what purposes and under which conditions. Such right or rights are granted through agreements between the relevant parties. In CLARIN we refer to all such agreements as *license agreements*.

An overview of IPR law can be found in 'Report about New IPR Challenges'. DASISH Work Package 6, Deliverable D6.1, June 2013, available online at http://dasish.eu/publications/projectreports/D6.1_final.pdf.

As mentioned above, the European legislation on IPR differs throughout the countries, so apart from reading this guide CLARIN ERIC members should know about their national legislation. A way to gain such knowledge is by contacting the relevant national authorities, universities or libraries.

CLARIN favours the principles of Open Source and Open Access, cf. CLARIN Statutes Article 22, available at <https://www.clarin.eu/sites/default/files/OJ-2012-136-EU-Decision.pdf>. However, resources generally are protected by IPR and CLARIN has a framework of license agreements to handle this kind of data protection.

2.1.2 Legal Issues Concerning Personal Data

Special legal issues arise when personal data are involved, concerning the protection of such personal data. The legal framework around the use of personal data for research is complicated. In the EU, there is the General Data Protections Regulation¹ but CLARIN ERIC members should also know about their national legal framework regarding personal data. A list of the national data protection authorities for EU countries can be found at: http://ec.europa.eu/justice/data-protection/article-29/structure/data-protection-authorities/index_en.htm.

Researchers should be aware that the legal framework applies to material collected in any way, including from the internet where persons may be identified, *e.g.*, from social media such as Twitter and Facebook.

A thorough description of ethical issues can be found in 'Report about New IPR Challenges'. DASISH Work Package 6, Deliverable D6.1, June 2013
http://dasish.eu/publications/projectreports/D6.1_final.pdf

Agreements involving the publication of material containing personal data are not easy to put into a standard such as a license agreement. These agreements are usually individual, but informed consent from the person(s) involved is always required, and often also the anonymization of all personal data.²

2.1.3 Authentication and Authorization

The handling of legal and ethical issues in a technical framework such as CLARIN requires implemented procedures for *authentication* (that the user can be recognized and verified) and *authorization* (that a user is authorized to access the platform). These issues and best practices for their implementation are not described in this paper, but is available on the CLARIN web site at <https://www.clarin.eu/content/service-provider-federation> and <https://www.clarin.eu/content/federated-identity>.

2.2 Legal Issues Committee

CLARIN has a Legal Issues Committee, CLIC, see <https://www.clarin.eu/governance/legal-issues-committee>. The main responsibility of the committee is to advise the Board of Directors on all issues related to IPR, privacy protection and ethical matters, as well as legal issues related to access policies and their implementation. Every CLARIN ERIC member consortium appoints one expert to the committee. The board of directors can invite additional members.

¹ See <http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679>.

² In June 2015 the Council of the EU agreed on a proposal for a new EU data protection regulation. Information about the proposed regulation can be found at www.eudataprotectionregulation.com. The main elements of the agreement are a) An enhanced level of data protection, b) Increased business opportunities in the digital single market, c) More and better tools to enforce compliance with the data protection rules, and d) Guarantees regarding transfers of personal data outside the EU.

2.2.1 Legal Information Platform

In 2017, the CLIC launched a *Legal Information Platform*, see <https://www.clarin.eu/content/legal-information-platform>. This website contains extensive information on legal information for language resources, including Copyright and Related Rights, Copyright Exceptions, Orphan Works, Related Rights and Databases, Licensing Practice,, Public Licenses for Data and Software, License Chooser Tools, Personal Data Protection, Changing Laws in the EU, including the Text & Data Mining Exception in the UK, the EC's proposal for the Directive on Copyright in the Digital Single Market, as well as an extensive bibliography on legal issues for language resources.

2.2.2 CLIC White Papers

The White Paper Series is a venue for the Open Access publication of commentary and scholarship on legal issues and language science under the editorial direction of the CLIC. To date, two papers have been published:

- Creative Commons and Language Resources: General Issues and What's New in CC 4.0 - Pawel Kamocki & Erik Ketzan
- Guidelines for Building Language Corpora Under German Law: Guidelines by the DFG Review Board on Linguistics

3 Licensing

This chapter describes the basic licensing framework in CLARIN³. The framework is based on the roles, described in 3.1 below, that are involved in depositing, providing and using CLARIN resources, and on the basic license agreement types, PUB, ACA and RES described in 3.2 and 3.3.

Since the development of the CLARIN licensing framework, other licensing frameworks, *e.g.*, GNU, Creative Commons, META-SHARE, have become internationally adopted. Examples of well-known end-user licenses classified according to the CLARIN categories examples can be found at the CLARIN license categories website at <https://www.clarin.eu/content/license-categories>.

In order to help content providers and service providers choose the appropriate license agreement a tool has been developed, the so-called license category calculator, described in 3.5 below.

3.1 Roles

In the CLARIN licensing framework the following roles are involved:

- *Resource Owner*, who possess the copyright for the data, *e.g.*, a writer, a researcher collecting data from informants, or a tool developer;
- *Resource Provider*, who has acquired rights from the resource owner, *e.g.*, a publishing house or a university;
- *Service Provider*, who provides the technical infrastructure required for hosting data and for user authentication, *e.g.*, a CLARIN technical centre; and
- *End-User*, who uses the data provided through the CLARIN centre.

³ A discussion about details in the license framework can be found at <http://www.ep.liu.se/ecp/article.asp?issue=123&article=002>.

3.2 Deposition License Agreements

A *Deposition License Agreement* (DELA) regulates the legal relationship between the resource provider, who deposits the data, and the service provider, including the end-user license conditions required by the depositor⁴.

For each resource deposited with a CLARIN centre, a DELA must be made.

CLARIN DELA templates are available, but must be adapted to the specific national legislation: <https://www.clarin.eu/content/licenses-agreements-legal-terms>.

Licenses from internationally established licensing frameworks, *e.g.*, GNU, META-SHARE and Creative Commons are acceptable as end-user license conditions indicated in the DELAs.

3.3 End-user License Agreements

An end-user license agreement (EULA) regulates the use of the resources by the end-users. A EULA is selected by the resource provider and consequently stated in the DELA.

Each resource offered through CLARIN must come with a EULA.

Licenses from internationally established licensing frameworks, *e.g.*, GNU, META-SHARE and Creative Commons will be accepted as well.

3.4 CLARIN License Categories

There are three basic access categories, PUB, ACA and RES, and each of these can be supplemented with additional requirements. For more details, see the comprehensive description at <http://hdl.handle.net/11372/DOC-62> (PART III), and the CLARIN ERIC website at <https://www.clarin.eu/content/license-categories>.

3.4.1 PUB

PUB means *Public Use*, indicating that no authentication and no authorization required. It is the least restricted license type and a CLARIN template for it is available at <https://www.clarin.eu/content/licenses-agreements-legal-terms>. The template must be adapted to the specific national legislation. Legacy DELAs may include non-standard provisions, but it is recommended that new resources are categorized with the standard provisions.

3.4.2 ACA

ACA means *Academic Use*, indicating that a resource with this license can only be used for research purposes. The ACA license relies on the implementation of user authentication and authorization. A CLARIN template for the license is available at <https://www.clarin.eu/content/licenses-agreements-legal-terms>. The template must be adapted to the specific national legislation. Legacy DELAs may include non-standard provisions, but it is recommended that new resources are categorized with the standard provisions.

⁴ The agreements made between the resource owner and the resource provider are in principle no concern of the CLARIN technical centre. CLARIN assumes that the entity offering the material for deposition has acquired the adequate rights to do so.

3.4.3 RES

RES means Restricted Use. This license type would typically require direct one-to-one contact between the resource provider or resource owner and the end user. RES would be the typical license category to use for material that contains personal data. An adaptable CLARIN template for it is available at <https://www.clarin.eu/content/licenses-agreements-legal-terms>.

3.4.4 Additional Requirements

The three basic license categories can be supplemented with a long row of additional requirements, so-called “laundry tags”. These comprise statements about commercial use, requirements about re-deposition of modified versions of tools or data *etc.* The possible additional requirements are listed below each of the three main categories on the aforementioned website.

3.5 CLARIN License Category Calculator

As more experience was gained in the member countries, the need for more additional requirements to the three basic categories and the need for combinations of these requirements became clear. All centres are recommended to use the License Category Calculator, which is available at the CLARIN website, see <https://www.clarin.eu/content/clarin-license-category-calculator>.

The user interface to the calculator is a web form, where users are asked a series of yes/no questions. The questions are divided into five sections. The first section contains five questions about Access Conditions, the next section asks six questions about General Usage Conditions, and the last section asks four questions about Distribution Conditions. For each question that the user answers, the calculated result appears immediately at the top of the form.

It can be very complex to handle combinations of Additional Requirements manually. It is therefore highly recommended to use the calculator.

A more advanced version of the calculator is available at the Finnish CLARIN website, see <http://www.helsinki.fi/finclarin/calculator/ClarinLicenseCategory.html>. In addition to calculating the license, this calculator also generates the relevant DELA, EULA and ToS documents.

4 Terms of Service and Privacy Policy

Terms of Service (ToS) is also often referred to as terms of use. These terms contain a general description of the rules and conditions for the use of the website or the service, and are aimed at laymen.

A *Privacy Policy* describes how the personal data of the end-user will be processed.

A CLARIN centre must offer both a ToS and a Privacy Policy.

Templates are available, but must be adapted to the specific national legislation, see <https://www.clarin.eu/content/licenses-agreements-legal-terms>.

5 Lessons Learnt

In the CLARIN centres at the national level license issues are handled in very different ways. Generally, CLARIN centres do not give much explanation. This may be wise considering that the subject area is complex and that such communication is aimed at laymen. No CLARIN centre would want to complicate things for its depositors and end users. Some countries have implemented IT-systems for depositing material, *e.g.*, Norway, see <https://repo.clarino.uib.no/xmlui/page/deposit>, where the depositor has to go through a procedure, which includes a licensing step.

6 Conclusion

We have seen that there are two elements in the handling of IPR and Ethical issues: One is understanding or at least having an overview of the legal framework governing these issues:

- National and international IPR law
- Protection of personal data.

The second is an understanding of how these legal frameworks should be implemented, technically. Throughout this report we have underlined the most basic elements for the handling of these issues in CLARIN:

- The handling of legal and ethical Issues in a technical framework such as CLARIN requires implemented procedures for user authentication and authorization.
- For each resource deposited with a CLARIN centre, a Deposition License Agreement (DELA) must be made.
- Each resource offered through CLARIN must come with an End-User License Agreement (EULA).
- Licenses from internationally established licensing frameworks, *e.g.*, GNU, META-SHARE and Creative Commons are acceptable as well.
- It can be very complex to handle combinations of additional requirements manually. It is therefore highly recommended to use the CLARIN license calculator.
- A CLARIN centre must offer both a Terms of Service and a Privacy Policy.

References

<https://creativecommons.org/>

<https://www.gnu.org/licenses/licenses.en.html>

<http://www.meta-net.eu/meta-share/licenses>