ISSN 2003-2382

License Clearance Tool: A holistic technical solution promoting open IP and open innovation practices among research communities

By Marianna Katrakazi, Panagiota Koltsida, Eleni Toli, Prodromos Tsiavos. Athena Research & Innovation Center

ABSTRACT

Open Science (OS) movement, remote collaboration among research communities and the increased quantity of new content, data and resources, have made it clear that traditional licensing schemes require new tools that would combine technical and legal features, as well as techno-licensing tools. In fact, the diversity of open licenses deprived from standardization frequently leads to situations where more than one open licenses with different or conflicting terms apply at the same time, and hence it gives rise to license compatibility concerns. This creates a legal uncertainty that may discourage authors, scientists, and researchers from releasing their work under an open license. In this paper, we identify legal and technological barriers that pose a challenge in adopting open science practices; thereafter, we present a new tool, named License Clearance Tool (LCT), which has been developed by the Athena RC (Greece) as part of the National Initiatives for Open Science in Europe – NI4OS-Europe (https://ni4os.eu/), a European project that contributes to the European Open Science Cloud (EOSC) by supporting its activities in Southeast Europe. LCT is an open-source tool, which provides a holistic approach addressing IP issues. LCT focuses on automating the clearance of Intellectual Property Rights (IPR) by ensuring the compatibility among different licenses included in the same resource and assists users on the selection of the most suitable license by providing a content summary of them with respect to permissions, prohibitions, and obligations in relation to the user needs. It is intended to support mainly researchers and non-legal experts in general to publish in FAIR/open modes.

1 INTRODUCTION

The advent of low-cost Information and Communication Technologies (ICT) and the World Wide Web in the early 1990s led to increased generation of new content and knowledge, as it allowed the collection of large amounts of data and information that could be easily used, copied, modified, or distributed for further use, often with no or without significant financial or technical barriers.[1] For the first time in the history of humanity such an extended collaboration between researchers and the production of collaborative research outcomes had been made possible[2] and new opportunities emerged for scientists and researchers to publish and share the content of research projects, scientific papers and large data sets.[3] Such developments to a great extent followed collaboration patterns found in Free / Open Source Software (FOSS)[4] communities. FOSS practiced a licensing model based on a premise of sharing and collaboration rather than exclusion and direct exchange.[5] In case of content, reusability of copyrighted works was achieved through open content licenses.

The Creative Commons (CC) initiative, which was initially set up in 2002, contains a set of various licenses that allow people to share their copyrighted work to be copied, edited, built upon, etc., while retaining the copyright to the original work; CC provides six core licences, each of which allow stakeholders to use the original work in different ways. While there are different CC licences, all CC licences include certain standard rights and obligations. CC initiative constitutes one of the most successful open content licensing schemes and provides authors with a great variety of licenses for literary, musical or audiovisual works enabling them to choose the most appropriate one that meets his/her needs. CC initiative aims to make copyright content more 'active' by ensuring that content can be reutilized with a minimum of transactional effort.[6] Thus, the emergence of FOSS and open content licenses together with ICT revolution has brought a new economic model for the sharing of digital resources and the reusability of existing knowledge.[7] It has also created new challenges and opportunities for Open Access movement, as defined in the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003).[8]

The Open access movement constitutes an essential attribute of Open Science (OS) and aims to make scientific

knowledge openly accessible in ways that maximize its value to science and society. Researches can benefit from the greater scrutiny offered by open science, as it allows a more accurate verification of research results, whereas authors experience an increase in the number of citations their works receive in the open access environment.[9] At a global level, international institutions and other bodies have taken many initiatives aiming to implement Open Science mechanisms and some countries have made efforts to adapt legal frameworks and implement policies encouraging greater openness in science. At the EU level, the European Commission has placed a great emphasis on the adoption of OS practices during the last years. Indicatively, several pieces of EU legislation were adopted in order to facilitate the reuse of research data, such as Public Sector Information Directive (PSI) and the EU Copyright Directive[10]. Furthermore, the open-access policy of Horizon 2020 projects provides for open-access to publications by default.[11] Additionally, in 2018 the European Open Science Cloud (EOSC) was launched in the context of the broader Digital Single Market strategy, which constitutes a pan-European federation of data infrastructures supported by the EC, Member States and research communities. EOSC aspires to provide a solid framework for collaboration and the pooling of resources at European, national, regional and institutional levels. EOSC highly promotes the use of open licenses and all stakeholders wishing to contribute to EOSC are highly encouraged to use open content and open software licenses.[12]

There is thus a growing need to develop legal and technological solutions to cater not only for the increased knowledge sharing, but also to allow scientific practices supporting openness and collaboration to flourish. However, the proliferation of FOSS and Open Science projects led to a series of issues of what became known as the problem of the fragmentation of the commons', i.e. the creation of multiple licensing schemes that were not necessarily compatible with each other.[13] For a resource provider, choosing the appropriate license for a combined resource or choosing the appropriate licensed resources for a combination is a difficult process, given that it involves choosing a license compliant with all the licenses of combined resources.[14] The paper discusses how issues of commons fragmentation or licensing compatibility can be tackled through a combination of licensing and technological tools, what we call in this paper, a techno-licensing approach. LCT aims to help researchers, universities and other stakeholders to freely use and share their ideas without any legal uncertainty related to the licensing scheme applicable to them and to contribute to the establishment and sustainability of EOSC, where all researchers, innovators, companies and citizens can publish, find and re-use data, tools and services for research, innovation and educational purposes.

2 LEGAL CHALLENGES

The legal challenges in the new research environment created by the Open Science movement have been the driver for the development of the LCT

a) Large sets of open licenses with different or contradictory content.

The wide spectrum of actions available in the digital era has affected conventional Intellectual Property (IP) licensing practices and highlighted the need for alternatives to the mainstream models of sharing copyrighted material in a lawful manner. In this context, a series of different open licenses emerged that allowed the free use and dissemination of copyrighted content. Since then, many

1 Tim Berners-Lee, *Weaving the Web: The Past, Present and Future of the World Wide Web by Its Inventor* (Orion Business 1999).

2 Don Tapscott and Anthony D. Williams, Wikinomics: *How Mass Collaboration Changes Everything* (New York: Portfolio 2008); Thomas L. Friedman, *The World is Flat: A Brief History of the Twenty-first Century* (Farrar, Straus, and Giroux 2005).

3 Organisation for Economic Co-operation and Development (OECD), "Making Open Science a Reality" (2015) OECD Science, Technology and Industry Policy Papers, No. 25, OECD Publishing, Paris. http://dx.doi.org/10.1787/5jrs2f963zs1-en.

4 Richard Stallman, "The GNU Operating System and the Free Software Movement." in Chris DiBona, Sam Ockman & Mark Stone, Open Sources: Voices From the Revolution, (O'Reilly & Associates Inc, 1999).

5 Yochai Benkler, "Coase's Penguin, or, Linux and the Nature of the Firm" (2002) https://doi.org/10.2307/1562247.

6 Brian Fitzgerald, "Open Content Licensing (OCL) for Open Educational Resources" https://www.oecd.org/education/ceri/38645489.pdf.

7 Yochai Benkler, "The Wealth of Networks: How Social Production Transforms Markets and Freedom" (2006) https://doi.org/10.1177/1084713807301373.

8 Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003), https://openaccess.mpg.de/Berlin-Declaration.

9 Fitzgerald (n 6).

10 Directive (EU) 2019/790 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC (EU Copyright Directive) [2019] OJ L 130, 17.5.2019, p. 92–125.

11 European Commission, "Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020" (2013) http://www.gsrt.gr/EOX/files/h2020-hi-oa-pilot-guide_en.pdf.

12 EOSC Rules of Participation (2021), Art. 5.

13 Niva Elkin-Koren, "Exploring Creative Commons: A Skeptical View of a Worthy Pursuit. The Future of the Public Domain (P. Bernt Hugenholtz & Lucie Guibault, eds.)", Kluwer Law International (2006) http://ssrn.com/paper=885466.

14 Benjamin Moreau and Patricia Serrano-Alvarado and Matthieu Perrin and Emmanuel Desmontils, " Modelling the Compatibility of Licenses" 16th Extended Semantic Web Conference (ESWC2019) (2019) DOI:10.1007/978-3-030-21348-0_17.

software companies have started to adopt open-source licensing models as part of their business[15], and many scientists choose to make their work freely available.[16] Open licenses include a series of different licenses and many sub-categories thereof. According to the most acceptable definition on open licenses provided by the Open Source Initiative,[17] an open license contains the following features: (1) free distribution of the software; (2) free access to the source code (just reproduction costs are covered); (3) authorization of modifications and the distribution of derived works; (4) no discrimination between people and fields of endeavor; (5) no restriction on other software; and (6) technological neutrality as well as independence from a specific product. Such licenses vary, depending on the copyrighted work (software, data, content or other) and the rights and powers granted to users, such as the right to use a work, to merge two different works, to relicense a work under different terms, etc.

b) Standardization aspects

In addition, a standardization problem exists, namely, license texts may either form part of the source file or may be missing completely. Even in cases where license information is put at the beginning of a source file, it usually does not follow proper standards.[18] This diversity of open licenses deprived from standardization frequently leads to situations where more than one open license with different or conflicting terms apply at the same time, and that in its turn gives rise to license compatibility concerns. For instance, a license that excludes commercial use cannot be combined with a license that permits so, and they, thus, may be jointly used. Similarly, a license that forbids the distribution of a derivation (remix, transform or build upon) cannot be combined with a license that permits so.

Joint use of different licenses may happen in case of relicensing, dual licensing, sublicensing, or in case of derivative works, either by adding a new material to the existing work and seeking for a new license for the new work, or by combining two works with different licenses. For instance, in open software licenses, the problem that many software vendors often face is how to incorporate third party software in their implementations correctly without causing any license violations, guaranteeing thus legal compliance.[19]

c) Broadening of initial license scope

Another confusing aspect of licenses relates to their scope: most open licenses have been developed for licensing software. They differ from open licenses that have been developed for licensing other material, which is also protected by copyright.[20]

This situation affects the scientific community and all stakeholders wishing to use an open license for their work. It constitutes a major barrier in open access, because it creates legal uncertainty that discourages authors, scientists, and researchers from releasing their work under an open license; the need for sufficient expertise to detect compatibility conflicts between licenses leads to high transaction costs associated with the manual clearance of licensing terms and conditions.[21] Especially for software, the dependency-related license violations are overlooked and misunderstood by the developers for various reasons. Managing dependency-related license violations is difficult and the developers are demanding help.[22] Furthermore, for an individual author who wishes to make his/her publication open access, the procedure used to select

[15] Mikko Valimaki "Rise of Open Source Licensing – A challenge to the use of intellectual property in the software industry", MA thesis, Helsinki University of Technology (2005) http://lib.tkk.fi/Diss/2005/isbn9529187793/isbn9529187793.pdf.

[16] International Science Council "Open Science for the 21st century", Draft ISC Working Paper (2020) https://council.science/wp-content/uploads/2020/06/International-Science-Council_Open-Science-for-the-21st-Century_Working-Paper-2020_compressed.pdf.

[17] Lucie Guibault and Christina Angelopoulos, "Open Content Licensing, From Theory to Practice", Amsterdam University Press (2011).

[18] Georgia M. Kapitsaki and Frederic Kramer and Nikolaos D. Tselikas, *"Automating the license compatibility process in open-source software with SPDX"*, The Journal of Systems and Software (2016).

[19] ibid.

[20] Guibault (n 17).

[21] Giray Havur, Simon Steyskal, Oleksandra Panasiuk, Anna Fensel,Victor Mireles, Tassilo Pellegrini, Thomas Thurner, Axel Polleres, and Sabrina Kirrane, "Automatic License Compatibility Checking", CEUR Workshop Proceedings (2019) ← https://ceur-ws.org/Vol-2451/paper-13.pdf→.

[22] Shi Qiu and Daniel M. German and Katsuro Inoue, "Empirical Study on Dependency-related License Violation in the JavaScript Package Ecosystem" (2021) ←DOI: 10.2197/ipsjjip.29.296→.

[23] OECD (n 3).

[24] Moreau (n 14).

[25] Catherine Doldirina, Anita R. Eisenstadt, Harlan Onsrud and Paul F. Uhlir, "Legal Approaches for Open Access to Research Data" (2018) https://doi.org/10.31228/osf.io/n7gfa; Ignasi Labastida, "Legal requirements, RDM and Open Data" (2017) DOI: https://doi.org/10.14324/000.learn.22.

[26] Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (2016), OJ L 119, p. 1–88.

[27] The Royal Society Science Policy Centre, *"Science as an open enterprise"*, Science Policy Centre report (2012).

[28] Havur (n 21).

[29] Directive (EU) 2019/1024 on open data and the re-use of public sector information (recast) (2019) OJ L 172, p. 56–83, recital 28, art. 1.

[30] Christopher Vendome, Mario Linares-Vasquez, Gabriele Bavota, Massimiliano Di Penta, Daniel German, Denys Poshyvanyk, "License Usage and Changes: A Large-Scale Study on GitHub" (2015) doi: 10.1109/ICPC.2015.32.

the appropriate license for his/her work can be cumbersome; individual negotiations, for example, can be a burden on the author.[23] In case of a combined resource, the selection of the appropriate license is even more challenging, because it involves choosing a license compliant with all the licenses of combined resources as well as analyzing the reusability of the resulting resource through the compatibility of its license.[24]

d) Plethora of applicable legal requirements.

Sharing of knowledge (including texts, methods etc.), data and tools, hereinafter referred to as 'intellectual assets', in the context of EU's Open Science policy presupposes that such assets comply with the applicable EU and local Member State regulations; otherwise, no intellectual asset can be used safely and thus all stakeholders from across academia would be discouraged from using and sharing assets under the Open Science ecosystem. Thus, compliance with the applicable licensing frameworks guarantees the establishment of a trust framework in which open practices can be embraced as the modus operandi for all interested parties. In addition, intellectual assets are usually subject to more than one different legal regime regulating their use.[25] For instance, where open science involves the processing of personal data, it is subject to the applicable rules including the General Data Protection Regulation (GDPR[26]); or, if it includes confidential information, it is subject to contractual limitations (e.g. Non-Disclosure Agreements) and legal limitations (e.g. Trade Secrets legislation). Finally, before any intellectual assets are made available, they will need to be cleared off any other IPR, ranging from Trade Secrets to Patents and Utility Rights, as well as by other contractual or statutory restrictions, e.g. cultural heritage laws, national security provisions or statistical confidentiality provisions.[27] In other words, the key sources of legal transaction costs stemming are: first, issues of rights clearance and compliance with existing legal and contractual regimes; and second, issues of license compatibility when multiple assets under different - and often conflicting - terms are combined.[28] This is reflected in the relevant Open Data European legislation, particularly Open Data Directive (Directive (EU) 2019/1024 on open data and the re-use of public sector information), where it is expressly mentioned that all the aforementioned legal limitations should be taken into consideration and be excluded from the scope of application of the Open Data Directive according to the principle 'as open as possible, as closed as necessary'.[29]

e) Absence of publicly available tools for rights clearance.

Despite the proliferation of assets licensed under open licenses, and the fact that there are tools mostly focusing on the documentation of rights clearance processes as well as tackling license compatibility issues, major problems still exist. Most notably, such problems include: (a) the lack of free to access rights clearance tools; (b) the lack of maintenance of open licenses compatibility or public domain calculator tools, as well as the lack of traceability on license changes[30]; and (c) the absence of linking compatibility and clearance assessment to publicly available in open repositories resources.

3 THE LICENSE CLEARANCE TOOL (LCT)

The License Clearance Tool (LCT), a tool that is consistent with EU's open science policy, comes as a response to the increased demand for holistic technical solutions suitable for promoting the adoption of open science practices and the re-use of existing research and other types of work. In comparison to pre-existing tools dealing merely with a guided choice of open licenses, LCT has at its core the resource, or the digital asset generated either as original or derivative work. It helps addressing issues of copyright, privacy and confidentiality, data protection, limitations of national legislation, as well as any other additional limitation that may further restrict the use of the asset in the Open Science ecosystem. More specifically, LCT enables the proper IPR management through the clearance of open licensing terms and conditions, the indication of any applicable embargo policy and any other limitation that relates to cultural heritage legislation. It aims to facilitate and automate the clearance of rights (copyright) for datasets, media and software that are to be cleared before they are publicly released under an

open license and/or stored at a publicly trusted FAIR repository. The clearance metadata itself will be stored and licensed as an open-source resource. It provides equivalence, similarity and compatibility between licenses if used in combination, which is essential for derivative works. Furthermore, it helps users to take into consideration some of the core GDPR principles and raises awareness about privacy concerns. Identification of a valid legal basis that permits data processing, indication of the appropriate data masking techniques that safeguard the protection of personal data, such as anonymization and pseudonymization the transparency obligations of the data controller and the existence of any confidentiality agreement is an indicative list of the privacy-related issues addressed in LCT. The aforementioned information, together with any other types of rights (national legislation, national security etc.) that should be cleared before the asset is released, is accessible through a checklist. In this way, through a user-friendly and straightforward workflow, LCT assists users in determining the legal boundaries that exist in a specific asset, which contradict with the principles of being Fair, Accessible, Interoperable and Reusable (FAIR) and impede the free or under pre-defined conditions circulation of the asset in the Open Science ecosystem. LCT provides a solution to the challenge of addressing legal aspects in FAIR and in Open Research Data Management (ORDM).[31] It is, thus, intended to support mainly researchers and in general non-legal experts to publish in FAIR/open modes and facilitates the sharing of knowledge among the research communities and the attribution of the creator's work.

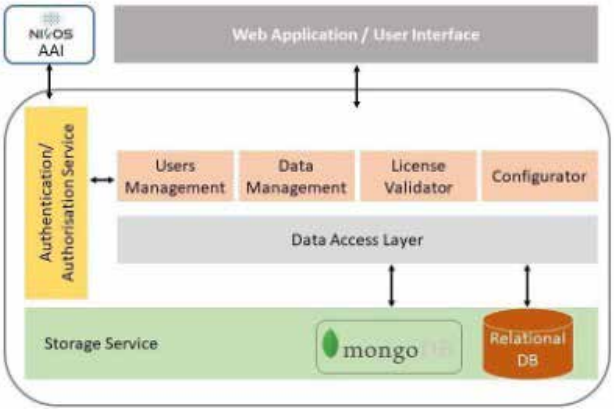
The tool provides guidance for 73 existing standard open-source licenses, as these are the most widely used and may thus accommodate most of the license clearance and IPR needs for non-legal experts for different types of resources. Finally, LCT is designed to be extensible with a plan to include and allow options for crowdsourced clearance in future work, for custom licenses that would otherwise require input from a legal expert.

3.1 Legal insight

LCT's development has been preceded by extended legal search and analysis of most used licensing schemes. The driving force has been to offer to all stakeholders and especially those deprived of a legal background, an easy-to-use tool to support them during the open license clearance process and in parallel address the most frequent case scenarios that restrict the use of an asset in the context of Open Science. In this context, we integrated in LCT the most usual legal boundaries relating to privacy and confidentiality. We attempted to cover core principles of GDPR including the transparency obligation of the data controller, identification of the applicable legal basis for data processing, technical measures for the protection of personal data, that are a *conditio sine qua non* for sharing personal data and using the asset in the Open Science context. Further IPR restrictions that may apply to an asset and national legislation limitations for national security or other reasons were taken into consideration as well when creating the workflow of LCT. It automates the clearance based on the actions or omissions that each standard open-source license provides for. These have been put in a matrix, to allow the comparison 'all with all' and unveil compatible and conflicting licenses. More specifically, we selected 73 most used standard open-source licenses for a wide variety of assets such as software, hardware, font, data etc. We reviewed the legal text of each license and categorized them on the basis of permissions, duties or prohibitions stipulated in each license (e.g. creation of derivative works, commercial use, distribution etc.) and upon categorization, licenses have been compared to each one in pairs. Licenses have been further classified in distinct license elements for each of the three categories. Through this assessment a core element of the application has been created, the license compatibility matrix.

An important concern in our work, has been to increase the legal transparency and awareness of the users. For this reason, the dedicated 'License Information' section is available, and users can navigate through it to understand the main elements of each open license. More specifically, this section provides a short summary as per license that enables users to check their elements with respect to the permissions, prohibitions and obligations, which determine the conditions under which the work is released: indicatively the permission to allow commercial use or not, permission of modification (creation of derivative works), or reciprocity obligation (copyleft or permissive). In this way, a codified version of licenses' summaries has been created, and next to each element an explanatory note has been added for the users' convenience so that they can understand the meaning of each attribute and select the most suitable license that corresponds more closely to their needs. A URL link leading to the entire legal text of each license is available, should users wish to consult it for more details. This section was a key step in LCT's development, as it allows the codification of licensing practices and further contributes to the reduction of transaction costs in the reuse of assets licensed under an open license. All users can easily compare among open licenses and choose the most appropriate one simply by navigating through the tool.

[31] Mark D Wilkinson et al., "The FAIR Guiding Principles for scientific data management and stewardship" (2016) doi: 10.1038/sdata.2016.18.



3.2 LCT approach & methodology

LCT is offering a user friendly and intuitive web user interface enabling its users to efficiently clear their work on a resource basis, receive a clearance report including all the provided information and retrieve detailed information for each supported license.

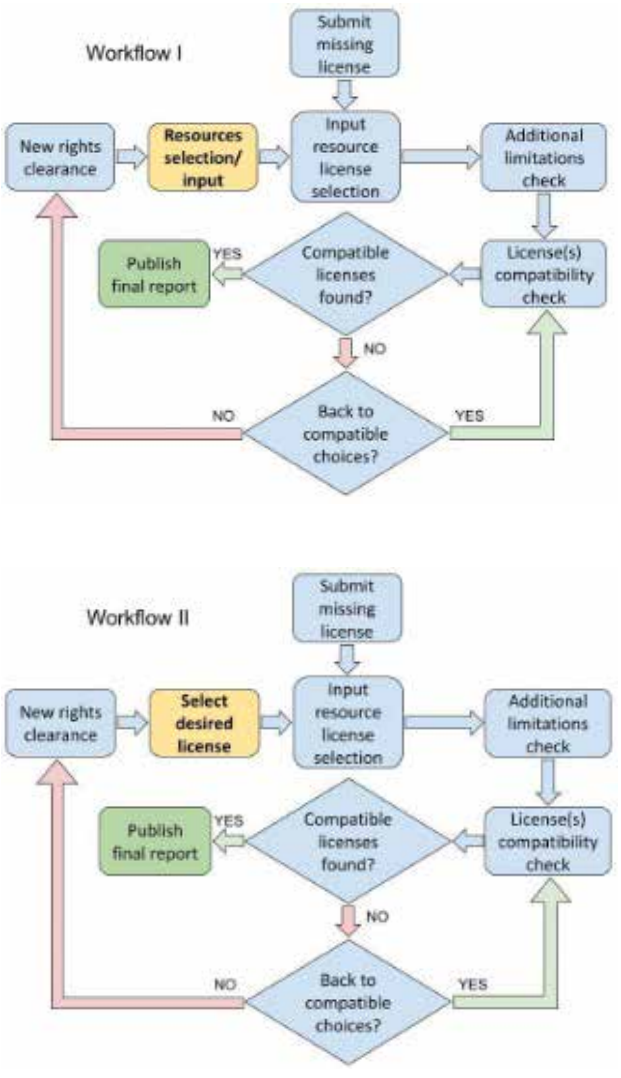
The tool incorporates two main scenarios aiming on supporting the two most common use cases, as these were described by our target users during the design process. The 'resource driven clearance', where the user aims to associate an appropriate open-source license for existing work composed by different elements that are licensed separately, or the 'license driven clearance' for derivative work by combining licenses from the originating licensed or possibly unlicensed content and the reverse rights clearance procedure. Both scenarios incorporate all IPR, personal and other rights related to the resource, aiming on raising awareness on the most common legal aspects that affect the future usage and exploitation of the cleared resource.

To support these scenarios, LCT has developed a compatibility mechanism able to calculate the compatibility among an arbitrary number of given licenses. This mechanism is enhanced with the option to further limit the compatible licenses based on a set of given attributes that should be met. These attributes are a subset of the list of license elements for each of the four categories.

The web application is supported by a web service responsible for the initiation and display of the guided wizards, the license compatibility check, the report generation, and the user's management. LCT's dynamic approach has been reflected in the development of two different schemas one for each scenario, using the JSON notation, that model the different inputs and the structure of each workflow that is dynamically interpreted by the front-end application and is displayed to the end users. Following this design approach our application can dynamically accompany any changes in its wizards, eliminating the need of source code updates and releases. Figure 1 presents an architecture block diagram of the LCT application showcasing the different modules and services and the interaction with external modules for the authentication of the registered users.

3.3 LCT workflows

Two main workflows are supported in LCT. These are following the two possible usage scenarios the application covers. Workflow I in the flowchart below, describes the process designed in the tool to implement the first usage scenario. It starts with a new rights clearance process initiated by the user by selecting the type of the resource under clearance. The process is bound to the resource itself and not the user who performs the clearance, allowing different users to complete the clearance of the same work. It is then followed by the association of each input/used internal resource with a corresponding 'license-in' license and information. After this step is completed, the application invokes the compatibility module and calculates the list of the compatible open-source licenses based on the previous ones and allows the user to select the desired one. In the last steps additional information related to personal data and other rights is collected and the clearance is submitted leading to the generation of a compatibility report for the provided resource. In case no compatible licenses are found, the process can be refined or aborted.



Workflow II allows the user to start a new clearance process by first selecting the license he/she is interested in using for releasing his/her resource. The algorithm works with this target for the additional steps that do not otherwise deviate from workflow I, at least from the user's point of view.

If there is no compatibility among the desired license choice and the given used internal resources' licenses, then the derived work cannot be published with the chosen license-out and a different one should be selected. This workflow provides a user with license options by eliminating incompatible choices considered at zero cost.

3.4. The application

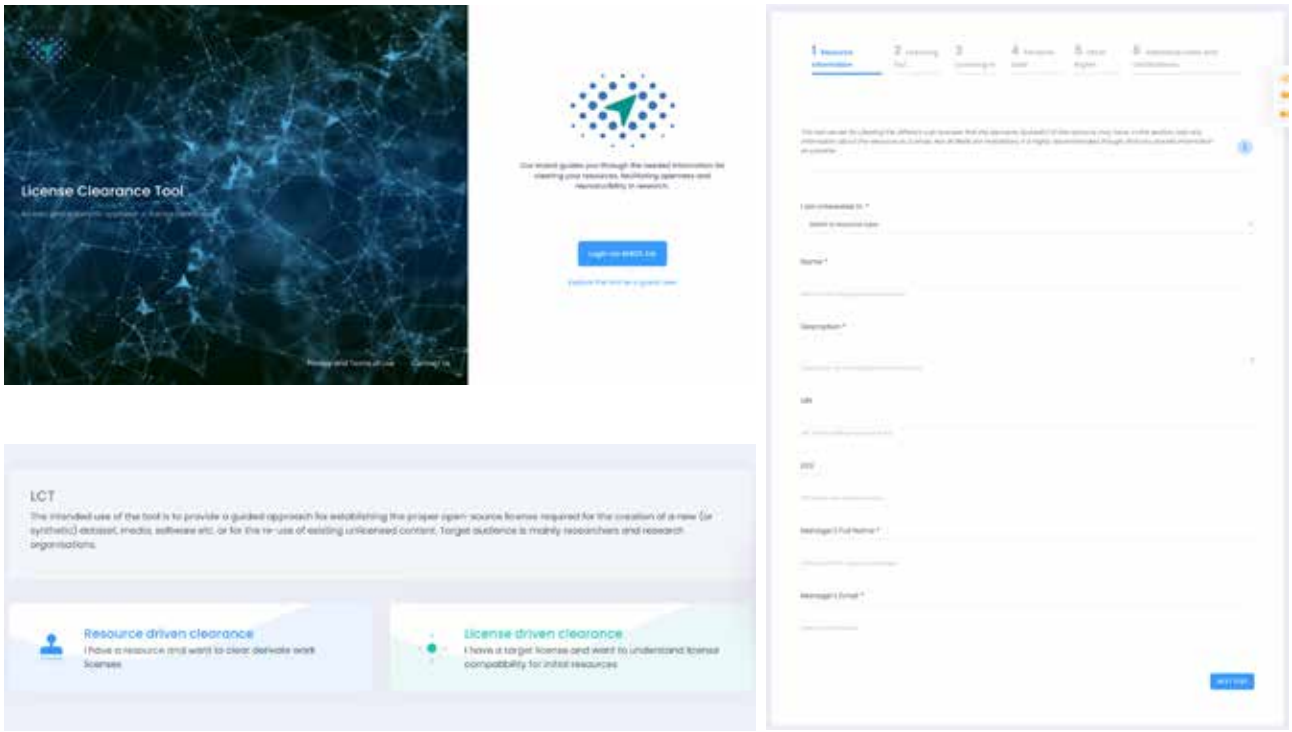
An end user web application has been designed and implemented and can be used by the research communities. The application is available for both guest users and registered ones and is based on a guided form wizard for the two distinct scenarios. It facilitates the clearance process and after the submission of a form, the user can download the assessment result as a custom pdf report. Registered users have the option to access all their past clearances and download their reports at any time. Figure 3 presents a sample of the application pages and the guided form wizard.

4 CONCLUSIONS

Licensing and rights clearance with respect to a broad range of legal aspects in the Open Science ecosystem is a complex issue and requires a great level of legal expertise. The difficulties lie not only in the need to be up-to-date with the current developments in terms of law, policies or other regulations with binding effect, adopted at either international or EU level, but also in the way this legal information is accessed and used. Techno-legal tools such as LCT do provide a possible solution, however further work is needed to support the changing and increased needs of researchers for publishing open and FAIR.

In this frame, LCT development investigates changes in two directions. A first expansion aims at including even custom licenses in order to properly address legal complexity. The potential direction for future work is the comparison of 'standard to custom' and 'custom to custom' licenses that poses a challenge for both directions: on the one hand, it involves a detailed legal analysis on the compatibility of licenses, and on the other hand it requires the technical deployment of the solution, which could be achieved though appropriate means that would make feasible the classification of custom licenses to specific license elements, thus enabling their automatic compatibility assessment with existing standard licenses and allow their usage in research outcomes and other types of work.

A second development direction under consideration for LCT, is the implementation of crowdsourced clearance. This requires a parallel effort at two levels: creating the grounds by setting up the environment, technically, as in a platform, and physically, as in community building, driving awareness, generating motivation. This will allow to provide a complete framework for crowdsourced clearance of custom licenses. The advantages of an open and citizen science-oriented approach are evident: as researchers aim to work in increasingly open and reproducible



ways to address challenges and solve problems, the crowdsourced license clearance can help to identify the best options and increase reproducibility even more.

We are well aware of difficulties and limitations the above processes may have. We consider, however, that they considerably enhance the sharing and use of knowledge in open research environments, without compromising in terms of awareness of the general legal framework as well as of IPRs.

ACKNOWLEDGEMENT

This research was funded by the European Commission, under the Horizon 2020 European research infrastructures, grant number 857645 (NI4OS-Europe project).

Marianna Katrakazi

Marianna Katrakazi is a lawyer-member of the Athens Bar Association specialized in IP and data privacy law. She holds an L.L.M. in Internet and Information Systems law from the University of Strasbourg and has extensive experience advising on GDPR compliance issues, telecommunications and IP law. Ms. Katrakazi has in-house experience in the Hellenic Industrial Property Organization, where she was involved in many European cooperation projects and law drafting activities related to patent law. Currently, she is Legal Counsel, Head of Legal Department at a multinational group of companies and participates as legal expert in the ATHENA RC's team for NI4OS-Europe, where she advises on legal matters related to privacy, IP licensing and Open Data.

Panagiota Koltsida

Panagiota Koltsida is a senior software engineer and research associate, with extensive experience in RTD European funded projects working for the National and Kapodistrian University of Athens (NKUA) and the ATHENA Research Centre. Mrs Koltsida has a degree in Informatics from NKUA and her MSc diploma from NKUA. She has been a technical project manager and software engineer in a series of EU funded projects . She is experienced in managing all stages of large scale distributed software implementation, designing ICT systems and technically coordinating their development. Her research interests include Information Retrieval systems, Data and Metadata Management systems, innovative ICT solutions for cross-disciplinary communities, and facilitating the adoption of open science principles and Open Data.

Eleni Toli

Eleni Toli is a research associate and Senior Project Manager at the Athena RC and at the Department of Informatics and Telecommunications at the University of Athens. She has been involved in the scientific and technical coordination of several EU-funded projects related to Open Science and Open Data, Research Data Infrastructures and Data Repositories. She holds an MSc degree in Quality Assurance, Information and Technology Management from the University of Bochum and a BS degree from the University of Ioannina. She is the NI4OS-Europe Project Director.

Prodromos Tsiavos

Prodromos Tsiavos is the Head of Digital Development and Innovation at the Onassis Group and a Senior Policy Advisor to the European Research Infrastructure for Open Science, OpenAIRE. He has served as the Chair of the Administrative Council of the Greek Industrial Property Organisation (OBI) and has founded the Hellenic Industrial Property Academy. He has over 150 publications and talks on legal and business aspects of e-government, Ethics by Design, open technologies, digital content and IPR and innovation policy and strategy. Prodromos is the director of the Institute of IPR and Innovation at the European Public Law Organisation and has served as the Chair of the Supervisory Board of the European Patent Academy (EPA). He is currently teaching AI Law and Ethics at the Athens University of Economics and Business (AUEB) and IPR Law at University of Athens (UoA).