Automated Rights Clearance Using Semantic Web Technologies: The DALICC Framework

14

Tassilo Pellegrini, Victor Mireles, Simon Steyskal, Oleksandra Panasiuk, Anna Fensel, and Sabrina Kirrane

Key Statements

1. The creation of derivative data works, i.e., for purposes such as content creation, service delivery or process automation, is often accompanied by legal uncertainty about usage rights and compliance issues with applicable law.
2. Challenges associated with clearance issues are: (1) high transaction costs in the manual clearance of licensing terms and conditions, (2) sufficient expertise to detect compatibility conflicts between two or more licenses, and (3) negotiation and resolution of licensing conflicts between involved parties.

T. Pellegrini (✉)
University of Applied Sciences St. Pölten, St. Pölten, Austria
e-mail: tassilo.pellegrini@fhstp.ac.at

V. Mireles
Semantic Web Company GmbH, Wien, Austria
e-mail: victor.mireles-chavez@semantic-web.com

S. Steyskal
Siemens AG Österreich, Wien, Austria
e-mail: simon.steyskal@siemens.com

O. Panasiuk · A. Fensel
University of Innsbruck, Innsbruck, Austria
e-mail: oleksandra.panasiuk@sti2.at; anna.fensel@sti2.at

S. Kirrane
Vienna University of Economics and Business, Wien, Germany
e-mail: sabrina.kirrane@wu.ac.at

3. Semantic processing of license information can ease the process of rights clearance in terms of cost reduction and improvement of the decision quality. However, they are not a substitute for a human expert.
4. Reliable and trustworthy semantic systems are transparent. If the user can't reproduce or retrace a given recommendation – be it from the methodologies applied by the system or the plausibility of the output – the recommendation should be rejected.

14.1 Introduction

Publishing data and reusing it for commercial or non-commercial purposes has become a common practice and a cornerstone of the so called digital economy [33]. IDC & Open Evidence [18] estimated in 2013 that the European Data Economy provides 6.1 million jobs in the EU28 and could almost double by the year 2020 if high-growth is ensured. Similarly, the number of organizations producing and supplying data-related products and services could reach almost 350,000 in 2020, from 257,000 in 2014, and there could be more than 1.3 million data users by 2020.[1]

New data practices stimulated by phenomena such as open data, open innovation and crowdsourcing initiatives as well as the increasing interconnection of services, sensors and (cyberphysical) systems have nurtured an environment in which the effective handling of property rights has become key to innovation, productivity and value creation. According to the OECD, the effective management of intangible assets is the primary driver of innovation in the ICT-enabled service sector and a source of competitive advantage at the macro- and micro-level [21]. This line of argument corresponds with a study conducted by Oxford Economics which argues that "insights derived by linking previously disparate bits of data can become the sparks that ignite rapid innovation" [28]. However, according to the EU Agency for Network and Information Security, the main obstacle in the digital ecosystems of the future is the legal impact of information exchange [3]. This is especially relevant in the context of the European strategy for a data-driven economy which aims to "nurture a coherent European data ecosystem, stimulate research and innovation around data and improve the framework conditions for extracting value out of data" [5]. Accordingly rights clearance to ensure legal compatibility has become a key topic in digital ecosystems as modern IT applications increasingly retrieve, store and process data from a variety of sources [15].

Clearing and negotiating rights issues is a time-consuming, complex and error-prone task. Challenges associated with clearance issues are:

---

[1] Similar trends and figures are reported for the Spanish data broker market by Granickas [13].

(1) High transaction costs in the manual clearance of licensing terms and conditions.
(2) Sufficient expertise to detect compatibility conflicts between two or more licenses.
(3) Negotiation and resolution of licensing conflicts between involved parties.

The following chapters introduce the DALICC system, a software framework that solves some of these problems by applying Semantic Web technologies to the purpose of license clearance.

DALICC stands for **Da**ta **Li**censes **C**learance **C**enter. It supports legal experts, innovation managers and application developers in the legally secure reutilization of third party data.[2] DALICC allows the attaching of licenses in a machine readable format to a specific asset and supports the clearance of rights by providing the user with information about similarity and compatibility between licenses if used in combination in a derivative work. Thus, DALICC helps to detect licensing conflicts and significantly reduces the costs of rights clearance in the creation of derivative works. Figure 14.1 gives an overview over the functional spectrum of the DALICC framework.
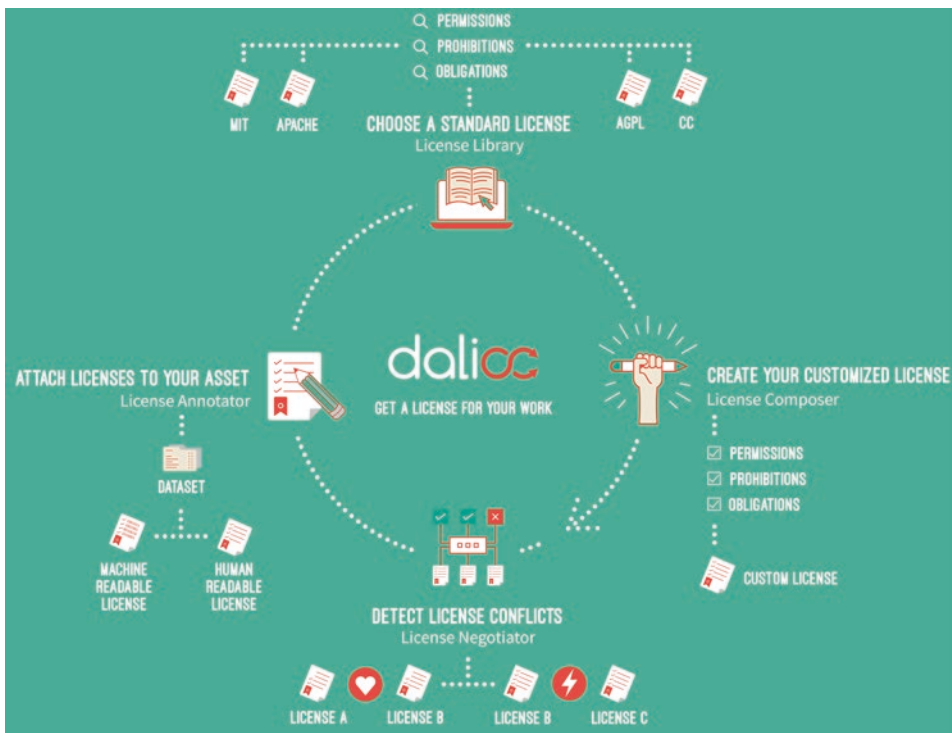

Fig. 14.1   The functional spectrum of the DALICC framework

---

[2]At the time of publication the DALICC service was in closed beta mode and not available to the public. A public version will be available at www.dalicc.net.

The following sections will discuss several challenges in automated license clearance and illustrate how Semantic Web technologies can be applied to solve these issues as exemplified with the DALICC framework.

14.2   Challenges in Automated License Clearance

14.2.1   License Heterogeneity

Licenses express permissions, obligations and prohibitions associated with a protectable asset as defined by copyright law or competition law. Licenses control access to, usage of, and transactions on top of digital assets, be it under conditions of property rights (all rights reserved) or public domain (no rights reserved) [34]. Figure 14.2 depicts the spectrum of available licensing models.

The growing popularity of protective and permissive licenses (some rights reserved) has added to the complexity of rights clearance in the commercial exploitation of derivative works. As a consequence, a wide array of data publishing guidelines were recommended [7, 14, 17] giving expression to the fact that licensing of data is a fairly new kind of economic practice and still subject to debate concerning the adequate design of licensing policies [1, 22, 29, 30]. This is supported by a recent survey conducted by Ermilov and Pellegrini [4] on 441,315 publicly accessible datasets. The situation is characterized by (1) insufficient documentation of licensing information (64% of all datasets had no licenses at all), (2) a high degree of license heterogeneity (more than 60 different license types), and (3) the absence of machine-readable licenses as a foundation for the automated clearance of compatibility issues.[3] Hence, the creation of derivative data works, e.g., for purposes such as content creation, service delivery or process automation, is often accompanied by legal uncertainty about usage rights and high costs in the clearance of rights issues [16]. This situation is further complicated as the efforts of license clearance increase with each additional source added to a system [f(n) = n*(n−1)/2]. According to Frangos [6] these efforts can be a serious obstacle for a company to create new products and services. Large companies usually operate rights clearance centres that manually evaluate legal issues in the repurposing of existing works (e.g., open source software). Such undertakings are
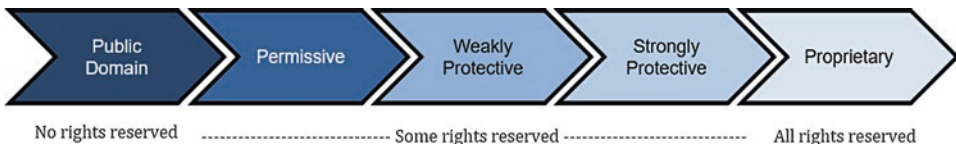

Fig. 14.2 Spectrum of licensing models

---

[3] Similar results are reported by Jain et al. [19] especially with respect to the large variety of licenses in European data clouds.

costly in terms of time and expert knowledge needed and are often out of scope, especially for small and medium sized enterprises. This is not just an obstacle to the emergence of new business models associated with data, but also slows down the rate of adoption of new data management practices, especially in the context of "a coherent European data ecosystem" as envisioned by the European Commission [5].

14.2.2   Rights Expression Languages

Rights Expression Languages (RELs) are a subset of Digital Rights Management technologies that are used to explicate machine-readable rights for the purposes of digital asset and access management. RELs are used to control access, explicate usage rights and govern behavioural aspects of a transaction process. Among the most prominent REL-vocabularies are MPEG-21, ODRL-2.0 (and derivatives such as RightsML), ccREL, XACML and WAC to name but a few [20]. Some RELs have a highly specific application focus, while others serve a general purpose. For example, MPEG-21 is optimized for rights management in the area of multimedia (especially digital television). On the contrary ccREL (Creative Commons Rights Expression Language) and ODRL (Open Digital Rights Language) are designed for broader application areas and have gained popularity especially in the area of content and data licensing.

Although the primary purpose of RELs is to explicate usage rights in a machine-readable form, simple tools to create, compare and process licenses attached to data assets are slowly emerging but they all have limitations.

A rudimentary version of a policy composition tool based on ODRL has been provided by the Ontology Engineering Group[4] of the University of Madrid, but this tool should be considered a proof of concept and has not been tested against real world circumstances. The same holds true for Licentia (http://licentia.inria.fr/), a license comparison tool developed by the French research institute INRIA. Also, the International Press Telecommunications Council (IPTC) working group on RightsML[5] has started to provide experimental libraries for generating RightsML and ODRL licenses in Python and JavaScript, but again, these serializations are just proof of concepts and lack a sufficient level of usability and legal validation to be suitable for commercial purposes.

Recently we have seen developments in the area of open source software that address the problem of license compatibility in the compilation of software from multiple open source libraries. For example, the Free Software Foundation provides a rich textual guide on potential licensing conflicts between open source standard licenses. This information however is not provided in a machine-readable format. Complementary to this, US-based auditing firms such as TLDR Legal (https://tldrlegal.com/) or TripleCheck

---

[4] See also http://oeg-upm.github.io/odrlapi/ & http://conditional.linkeddata.es/ldr/manageren (accessed January 4, 2016).

[5] See also http://dev.iptc.org/RightsML-Implementation-Examples (accessed January 4, 2016).

(http://triplecheck.net/index.html) have started to provide commercial services that help detecting open source licensing conflicts. What these initiatives have in common is that (1) they are specialized on software licensing, (2) compliance checking is provided as a commercial service conducted by auditing experts on top of proprietary tools, (3) none of these tools /services allows the creation of custom licenses, thus limiting the compliance check to standard licenses, and (4) no machine-readable representations of the licenses are provided to the public for advanced analytics and further reuse.

14.2.3 Machine-Processing of Licensing Information

Most of the work done in this research area is situated in the context of digital rights management systems and often associated with contracting issues [24, 26, 27]. Little attention has so far been paid to the issue of license compatibility and reasoning over machine-readable licensing information.

An interesting proposal for a generic logic for reasoning over licenses is provided by Pucella and Weissman [25], but it has not been implemented with existing RELs such as ODRL or MPEG-21, nor has it been evaluated in practice.

García et al. [8–10] propose an OWL ontology to describe copyright issues in closed datasets for rights clearance purposes. Their approach is based on an old version of the ODRL vocabulary and constitutes a proof of concept that has not been implemented or tested against issues arising from contemporary open data licensing.

Villata and Gandon [32] and Governatori et al. [12] describe the formalisation of a license composition tool for derivative works. They extend their research by introducing semantics based on a deontic logic [35–37] for the comparison of the permissions, prohibitions and duties stated in a given license. They also provide a demo called Licentia (http://licentia.inria.fr/) that exemplifies the practical value of such a service. This line of work is an interesting approach to detect and potentially solve licensing conflicts, e.g., by composing a new license that resolves the conflict. The pitfall of their approach is that an automatically composed license that resolves a given conflict might be logically correct but practically useless, because its conditions are either too strict or the machine-readable representation does not conform to human-readable deeds.

14.3 The DALICC Framework

14.3.1 System Requirements

According to Sect. 14.2, the following requirements can be derived for the DALICC system: (1) the output has to comply with applicable laws, (2) it needs to correctly interpret permissions, obligations and prohibitions from given licenses, (3) it must preserve abstractness and

technological neutrality of the rules, and (4) it needs to support the dynamics of the rules under conditions of real world applications and usage. To achieve these goals the following problems need to be addressed:

14.3.1.1 Tackling REL Heterogeneity

Combining licences is simpler if all of the licences involved are expressed through the same REL. But as we have seen, various RELs have emerged for various purposes, each providing their own vocabulary and level of expressivity. Hence, it is difficult to compare licenses that have been represented by different RELs. Additionally, it can sometimes be reasonable to extend the semantic expressivity of a given REL by adding deontic expressions from other RELs to cover the requirements of a real world scenario.

DALICC solves this problem by linking vocabularies from various RELs utilising W3C-approved standards (such as RDF and OWL), thus allowing mappings between various RELs to be created. This approach allows vocabulary terms from various RELs to be combined and their expressivity to be extended beyond their original scope. Figure 14.3 illustrates a RDF graph of the standard license CC-BY utilizing vocabulary expressions from ccREL, ODRL and the DALICC vocabulary.

This RDF graph is used as input to the reasoning engine described in Sect. 14.3.1.3.

```
@prefix odrl:<http://www.w3.org/ns/odrl/2/> .
@prefix : <https://dalicc.net/license-finder> .
@prefix dalicc: <https://dalicc.poolparty.biz/DALICCVocabulary>.
@prefix rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix cc:<http://creativecommons.org/ns#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .

:CC-BY_4.0 a odrl:Policy;
 odrl:permission [
            a odrl:Permission;
            odrl:target
            <http://purl.org/dc/dcmitype/Dataset>,<http://purl.org/dc/dcmitype/Sound>,<http://purl.org/dc/dcmitype/Text>,
            <http://purl.org/dc/dcmitype/Image>,<http://purl.org/dc/dcmitype/MovingImage>;
            odrl:action odrl:distribute, odrl:reproduce,odrl:extract, odrl:derive, odrl:present;
            odrl:duty [
            a odrl:Duty;
            odrl:action cc:SourceCode, dalicc:royaltyFree, dalicc:irrevocable, dalicc:worldwide, cc:Notice,
            dalicc:noWarrantyNotice, dalicc:modificationNotice, cc:attributionName
            ]
  ];
 odrl:prohibition [
            a odrl:Prohibition;
            odrl:target
            <http://purl.org/dc/dcmitype/Dataset>,<http://purl.org/dc/dcmitype/Sound>,<http://purl.org/dc/dcmitype/Text>,
            <http://purl.org/dc/dcmitype/Image>,<http://purl.org/dc/dcmitype/MovingImage>;
            odrl:action odrl:ensureExclusivity, dalicc:sublicense, dalicc:charge
  ];
  dct:title "Attribution 4.0 International"@en ;
  dct:alternative "CC BY 4.0";
  dct:publisher "Creative Commons";
  foaf:logo <http://i.creativecommons.org/l/by/4.0/88x31.png> ;
  dct:source <http://creativecommons.org/licenses/by/4.0/> ;
  cc:legalcode """https://creativecommons.org/licenses/by/4.0/legalcode"""@en.
```

Fig. 14.3 RDF-representation of CC-BY using ccREL, ODRL & DALICC vocabulary extensions

14.3.1.2   Tackling License Heterogeneity

Is it possible to combine a GPL Documentation License (as used by Wikipedia) with the Italian Open Government Data License v.1? Is CC-BY-ND compatible with UK-CROWN? In the creation of derivative works, the simplest approach is to only combine content under the same well-known licence. But this approach is over-restrictive, as many licences permit the licensed content to be combined. It is, however, difficult to judge whether it is permitted and how the resultant content should be licensed. There may still be subtleties arising from unclear definitions of terms (e.g., "open" or "commercial use"), special clauses (e.g., share-alike) or implicit preconditions (e.g., "everything not permitted is forbidden" or "CC0 apart from images – see restrictions in further links").

DALICC resolves these issues by producing an audited set of machine-readable licenses utilizing a given set of permissions, obligations and prohibitions – also called deontic expressions. Thus, DALICC is able to compose the crucial actions of existing standard licenses (like Creative Commons, Apache, BSD or GPL) but also allows the creation of customized licenses if none of the standard licenses suit the user's demand. To achieve the necessary level of semantic expressivity, an indepth analysis of deontic expressions in existing licenses has been conducted, matched against the vocabulary of existing RELs and complemented with additional properties, so that a sufficient level of expressivity could be reached.[6] Figure 14.4 illustrates some of the new properties DALICC utilizes to represent an arbitrary license in a machine-readable format.

By providing a sufficiently expressive set of deontic expressions, DALICC is able to represent licensing terms at a highly granular level, identify equivalent licenses and point

Properties Affecting The Asset

- » dalicc:charge
- » dalicc:sublicense
- » dalicc:promote
- » dalicc:publish

Properties Affecting The License

- » dalicc:addStatement
- » dalicc:attributionNotice
- » dalicc:attachOffer
- » dalicc:chargeOffer
- » dalicc:irrevocable
- » dalicc:modificationNotice
- » dalicc:noWarrantyNotice
- » dalicc:patentFree
- » dalicc:patentNotice
- » dalicc:perpetual
- » dalicc:royaltyFree
- » dalicc:worldwide

Fig. 14.4 DALICC vocabulary extensions for expressing actions associated with assets or licenses

---

[6] The following 14 commonly used licenses have been semantically represented: CC-BY, CC BY-SA, CC BY-ND, CC BY-NC, CC BY-NC-SA, CC BY-NC-ND, BSD (2 clause), BSD (3 clause), MIT, APACHE, GPLv3, GPLv2, AGPL and LGPL.

the user to potential conflicts if licenses with contradicting conditions are to be combined in a derivative work. These licenses lay the foundations of the DALICC Framework, and provide the grounding for its functional components:

- The License Library lets the user select either from a set of standard licenses or customized licenses provided to the library by users.
- The License Composer employs the license models to create customized licenses.
- The License Negotiator processes and interprets the semantics encoded in the licences and checks compatibility, detects conflicts and supports conflict resolution.
- The License Annotator provides a machine-readable and human-readable version of the license that can be attached to an asset.

The mechanisms for these functional components are described in the next section.

14.3.1.3   Compatibility Check, Conflict Detection and Neutrality of the Rules

A common problem with semantic translation between schemas (such as RELs) is in making sure that the meaning of different terms are aligned. However, it is difficult to demonstrate the equivalence of classes, properties and instances. For RELs, the major problem arises for the instances, e.g., the precise definitions of "non-commercial", "distribution", "share-alike" etc. The classes and properties are usually simple concepts and very similar. Not all RELs support all classes though: some ignore "Jurisdiction" or even "End-user" according to the needs of the market they were developed for. To a certain degree, this will be resolved by applying Semantic Web standards, but mapping alone cannot solve the issue. More elaborated techniques, such as reasoning and inference mechanisms, are necessary to improve the accuracy of conflict detection.

To solve these issues, DALICC applies a knowledge graph that is comprised of three components: (1) a set of defined actions representing permissions, obligations and prohibitions, (2) the RDF representation of these actions, and (3) a dependency graph representing the semantic relationship between the defined actions. The core function of the knowledge graph is to encode the expert knowledge about the implicit and explicit semantic dependencies between actions. Following the work of Steyskal and Polleres [31], the corresponding dependency graph represents hierarchical relationships (e.g., use includes reproduce), implications derived from a specific action (e.g., sell implies charge) and contradictions between specific actions (e.g., non-derivative contradicts derivative). Figure 14.5 illustrates the interplay of the various components within the DALICC knowledge graph.

The DALICC reasoning engine is based on the POTASSCO suite of grounders and solvers for Answer Set Programs [11] and uses ODRL policies to detect potential conflicts in licensing terms.[7] Policies should be understood as a set of rules derived from the RDF graphs of the licenses. Herein, a rule that permits or prohibits the execution of an

---

[7]Answer Set Programming has sparked theoretical interest from the Semantic Web Community before [23] because of its relationships to the well-known SPARQL querying language.
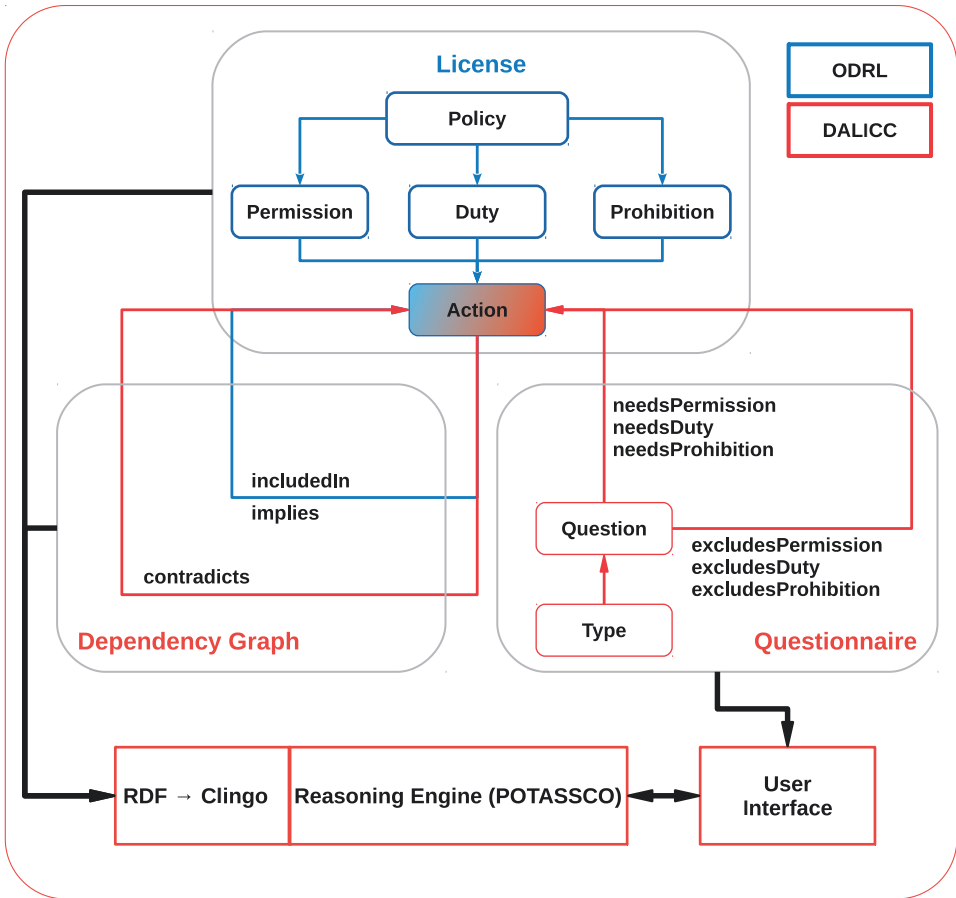
Fig. 14.5 DALICC knowledge graph overview

action on certain assets does not only affect other rules that govern the execution of the same action on the same asset(s) but also those permitting or prohibiting related actions on the same asset(s). DALICC utilizes an RDF-to-CLINGO[8] translator to translate the given rules to a processable format and wrap it into a web service that also allows SPARQL queries. In this sense, CLINGO is not only an alternative to extensive materialization, which in this case is essential for search, but also enables listing sets of compatible statements. This latter possibility is necessary for effective computation of conflicts between licences, in particular for identifying the conflicting and non-conflicting parts of a license.

---

[8] Clingo is an ASP system to ground and solve logic programs. See also https://potassco.org/ (accessed December 12, 2017).

14.3.1.4 Legal Validity of Representations and Machine Recommendations

The semantic complexity of licensing issues means that the semantics of RELs must be clearly aligned within the specific application scenario. This includes a correct interpretation of the various national legislations according to the country of origin of a jurisdiction (e.g., German Urheberrecht vs. US copyright), the resolution of problems that are derived from multilinguality (e.g., multiple connotations of "royalties" within German jurisdiction as "Lizenzgebühr", "Honorar", "Tantiemen", "Abgabe", etc.) and the consideration of existing case law in the resolution of licensing conflicts (e.g., Versata vs. Ameriprise)[9].

To tackle these issues, legal experts from inside and outside the DALICC consortium checked the legal validity of machine-readable licenses and the output of the reasoning engine for compatibility with applicable laws. In several iteration cycles the DALICC output has been tested against laws and jurisdictions, checked for its semantic accuracy and adjusted accordingly.

14.3.2 DALICC Implementation and Services

The DALICC user interface is based on the widespread content management system Drupal (https://www.drupal.org). Back office services are implemented with the PoolParty Semantic Suite (https://www.poolparty.biz/), which (1) manages a set of questions that guide the user in selecting a licence according to his/her needs and (2) maintains the knowledge graph which incorporates legal-expert knowledge about the licence clearing domain. The DALICC framework additionally provides a SPARQL endpoint which enables quick communication with the VIRTUOSO triple store (https://virtuoso.open-linksw.com/) containing the RDF data, the reasoning engine and the user interface. The architecture is designed to provide four basic services to the user as depicted in Fig. 14.6:

(1) *The License Library* contains machine-readable and human-readable representations of licenses. It can be accessed either via a full text search – best suited if the user already knows which license they needs – or by a faceted search that allows the user to filter licenses according to specific criteria such as asset type (e.g., data set, content or software), permissions, obligations or prohibitions. By default, the License Library is populated with the most important software and data licenses currently available. According to a recent study by Ermilov and Pellegrini [4],these include CC0, CC-BY, CC-NC-SA, UK-OGL, DL-DE-BY-1.0, IODLv2, APACHE, BSD, GPL and MIT to name but a few. Over the course of time, the library will be extended with additional licenses that frequently appear in the data domain or which are of specific importance for future applications (e.g., national open data licenses). The DALICC system also allows users to provide their customized license to the library for further reuse.

---

[9] See also Versata, Trilogy Software, Inc. and Trilogy Development Group v. Ameriprise, Ameriprise Financial Services, Inc. and American Enterprise Investment Services, Inc., Case No. D-1-GN-12-003588; 53rd Judicial District Court of Travis County, Texas.
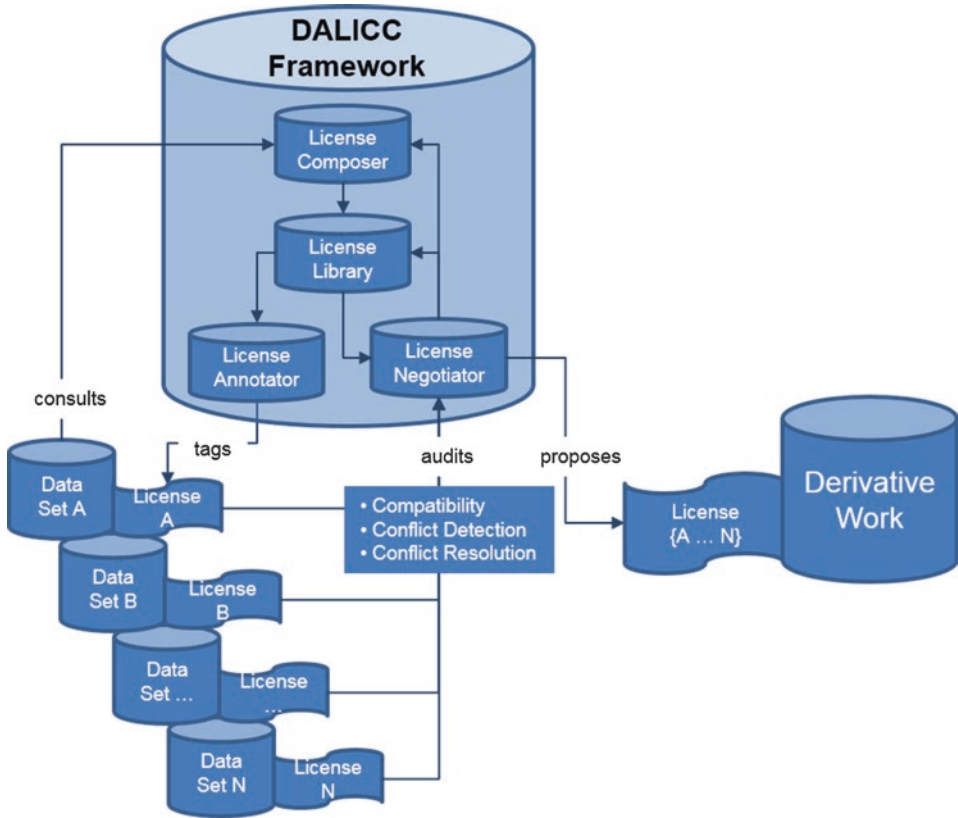
Fig. 14.6 DALICC service architecture

(2) **_The License Composer_** provides the user with a simple service that allows the declaration of necessary provenance information about the asset and guides the user through a relevant set of questions that need to answered to compile a legally valid license. The composer uses ODRL, ccREL and the DALICC vocabulary as a baseline vocabulary for the specification of licensing terms. The user is additionally provided with comprehensive explanations about specific terms so that non-experts are able to understand the legal impact of their decisions and acquire the needed literacy to compile a license that suits their purposes in the wide spectrum between open and closed licensing.

(3) **_The License Negotiator_** is DALICC's core component. It caters for reasoning over licenses taking into account the specific context of the application provider. The negotiator checks the logical coherence of the created license, provides information on equivalence, similarity and compatibility with other licenses and supports conflict resolution between licenses. Identified resolution strategies, i.e. for re-establishing compatibility among a set of licenses, do not solely refer to choosing the most restrictive license at hand and thus potentially reducing the usefulness of the resulting (combined) license. Instead it proposes a semantically equivalent and legally sound alternative license that might resolve the detected conflict.

(4) *The License Annotator* finally allows exporting and/or attaching a machine-readable and human-readable license to an asset. This can be done for either standard licenses (e.g., CC-BY) already available in the License Library or for customized licenses created with the License Composer. Each newly created license can also be added to the License Library, thus allowing incremental growth of the repository and the associated knowledge base. The licenses are also be available in various formats and provided as open data to foster maximum reuse.

14.4   Recommendations

The DALICC framework should be understood as a supporting infrastructure for the cost-effective clearance of rights issues, thus contributing to a significant reduction of transaction costs in the commercial exploitation of derivative works. Nevertheless, it is not intended to and never should replace the knowledgeable and critical human expert on the subject matter. Users of the DALICC system or similar services utilizing semantic technologies to support critical decisions, should be aware of the following things:

- Whenever you publish a derivative work and attach a license to it, you will be held accountable. Even if you have the best intentions, make sure that the assets you built your work upon have not violated other's intellectual property. Just having a license in place does not mean that prior clearance has taken place.
- Machine-readable representations of complex intellectual constructs will never capture and resemble the semantic accuracy given in a natural language text. Hence, machine-recommendations always come with a scope open to interpretation. Recommenders should be understood as decision support mechanisms but never be taken for granted.
- Reliable and trustworthy semantic systems are transparent. If you can't reproduce or retrace a given recommendation – be it from the methodologies applied by the system or the plausibility of the output – reject it.

14.5   Conclusion

Licensing in general and rights clearance in particular are complex topics that require a high level of problem awareness and legal expertise. Due to the abstractness and complexity of the topic, non-legal professionals need to invest a lot of time and/or money to acquire this knowledge and search for viable solutions. Semantic Web technologies are a viable means to create systems that reduce the complexity of the subject matter and provide services that can support stakeholders at various levels of expertise to engage in and contribute to emerging digital ecosystems.

Despite the new and exciting technological opportunities semantic technologies offer to us, it is still important to stress that technology should never replace the human expert. Hence, DALICC should be understood as a supporting service in the accountable and

ethical usage of property rights, to provide people with recommendations on how to protect their assets from misappropriation – be it for purposes of copyright or copyleft or something in between – and also to avoid unintentional misuse of other people's assets, that could undermine derived work.

To do so, DALICC will provide an open documentation of the system and provide its output as (linked) open data once it is fully operational. Additionally it is planned to make the DALICC framework available under a dual license, thus allowing various forms of collaborative exploitation. The framework closes the existing gap between the technological capabilities to create and publish data, and the legal infrastructure necessary to provide them on a legally secure basis for reuse. Hence, DALICC is a tool that puts data policies into practice and thus facilitates data governance. Hence, according to the data value chain provided by Deloitte [2], the DALICC framework should be understood as an enabling service for the emerging data economy.

References

1. Archer P, Dekkers M, Goedertier S, Loutas N (2013). Study on business models for linked open government data. ISA programme by PwC EU Services. European Commission. See also https://www.w3.org/2013/share-psi/workshop/krems/papers/LinkedOpenGovernmentDataBusinessModel. Accessed 29 Nov 2017

2. Deloitte (2012) Open growth. Stimulating demand for open data in the UK. See also http://www2.deloitte.com/content/dam/Deloitte/uk/Documents/deloitte-analytics/open-growth.pdf. Accessed 29 Nov 2017

3. ENISA (2013) Detect, SHARE, protect solutions for improving threat data exchange among CERTs, Oct 2013

4. Ermilov I, Pellegrini T (2015) Data licensing on the cloud: empirical insights and implications for linked data. ACM Press, pp 153–156. https://doi.org/10.1145/2814864.2814878

5. European Commission (2014) Towards a thriving data-driven economy. Brussels, 2.7.2014, COM(2014) 442 final

6. Frangos J (2015) New transparency in licensing: overview of the licensing facilitation act. Informed Couns 6(1):2

7. Frosterus M, Hyvönen E, Laitio J (2011) Creating and publishing semantic metadata about linked and open datasets. In: Wood D (ed) Linking government data. Springer, New York, pp 95–112

8. García R, Gil R (2009) Copyright licenses reasoning an OWL-DL ontology. In: Proceedings of the 2009 conference on law, ontologies and the semantic web: channelling the legal information flood. IOS Press, Amsterdam, pp 145–162

9. García R, Gil R, Delgado J (2004) Intellectual property rights management using a semantic web information system. In: Meersman R, Tari Z (eds) On the move to meaningful Internet systems 2004: CoopIS, DOA, and ODBASE, vol 3290. Springer, Berlin/Heidelberg, pp 689–704

10. García R, Gil R, Delgado J (2007) A web ontologies framework for digital rights management. Artif Intell Law 15(2):137–154. https://doi.org/10.1007/s10506-007-9032-6

11. Gebser M, Kaufmann B, Kaminski R, Ostrowski M, Schaub T, Schneider M (2011) Potassco: the potsdam answer set solving collection. AI Commun 24(2):107–124
12. Governatori G, Lam H-P, Rotolo A, Villata S, Auguste Atemezing G, Gandon F (2014) LIVE: a tool for checking licenses compatibility between vocabularies and data, vol 1272. See also https://hal.inria.fr/hal-01076619. Accessed 29 Nov 2017
13. Granickas K (2013) Understanding the impact of releasing and re-using open government data. In: European Public Sector Information Platform. Topic report no. 2013/08. See also http://www.epsiplatform.eu/sites/default/files/2013-08-Open_Data_Impact.pdf. Accessed 29 Nov 2017
14. Guibault LM (2011) Open content licensing: from theory to practice. Amsterdam University Press, Amsterdam
15. Hoffmann A, Schulz T, Zirfas J, Hoffmann H, Roßnagel A, Leimeister JM (2015) Legal compatibility as a characteristic of sociotechnical systems. Bus Inf Syst Eng 57(2):103–113. https://doi.org/10.1007/s12599-015-0373-5
16. Houghton J (2011) The costs and benefits of data provision. Report to the Australian National Data Service. Centre for Strategic Economic Studies, Victoria University
17. Hyland B, Wood D (2011) The joy of data – a cookbook for publishing linked government data on the web. In: Wood D (ed) Linking government data. Springer, New York, pp 3–26
18. IDC & Open Evidence (2013) European Data Market. SMART 2013/0063. See also https://drive.google.com/a/open-evidence.com/file/d/0B5Co3wBffnzhUTBQUklCS0VoRTg/view?pref=2&pli=1. Accessed 29 Nov 2017
19. Jain P, Hitzler P, Janowicz K, Venkatramani C (2013) There's no money in linked data. http://knoesis.wright.edu/pascal/pub/nomoneylod.pdf. Accessed 29 Nov 2017
20. Kirrane S, Mileo A, Decker S (2015) Access control and the resource description framework: a survey. Semant Web J. See also http://www.semantic-web-journal.net/content/access-control-and-resource-description-framework-survey. Accessed 29 Nov 2017
21. OECD (2008) Intellectual assets and value creation. See also http://www.oecd.org/sti/inno/40637101.pdf. Accessed 29 Nov 2017
22. Pellegrini T (2014) Linked Data Licensing – Datenlizenzierung unter netzökonomischen Bedingungen. In: Schweighöfer E et al (eds) Transparenz. 17. Int. Rechtsinformatik Symposium IRIS 2014. OCG Verlag, Wien
23. Polleres A, Wallner JP (2013) On the relation between SPARQL1. 1 and answer set programming. J Appl Non-Class Log 23(1–2):159–212
24. Prenafeta J (2010) Protecting copyright through semantic technology. Publ Res Q 26(4):249–254
25. Pucella R, Weissman V (2002) A logic for reasoning about digital rights. In: Proceedings of the 15th IEEE workshop on Computer Security Foundations. IEEE Computer Society, Washington, DC, pp 282–294
26. Rodriguez E, Delgado J, Boch L, Rodriguez-Doncel V (2015) Media contract formalization using a standardized contract expression language. IEEE Multimed 22(2):64–74. https://doi.org/10.1109/MMUL.2014.22
27. Rodriguez-Doncel V, Delgado J (2009) A media value chain ontology for MPEG-21. IEEE Multimed 16(4):44–51. https://doi.org/10.1109/MMUL.2009.78
28. Roehring P, Pring B (2013) The value of signal and the cost of noise. Oxford Economics, London
29. Sande MS, Portier M, Mannens E, Van de Walle R (2012) Challenges for open data usage: open derivatives and licensing. https://www.w3.org/2012/06/pmod/pmod2012_submission_4.pdf. Accessed 12 Feb 2016
30. Sonntag M (2006) Rechtsschutz für Ontologien. In: e-Staat und e-Wirtschaft aus rechtlicher Sicht. Richard Boorberg Verlag, Stuttgart

31. Steyskal S, Polleres A (2015) Towards formal semantics for ODRL policies. In: Rule technologies: foundations, tools, and applications – 9th international symposium, RuleML 2015, Berlin, Germany, August 2–5, 2015, Proceedings, pp 360–375. https://doi.org/10.1007/978-3-319-21542-6_23

32. Villata S, Gandon F (2012) Licenses compatibility and composition in the web of data. In: Third international workshop on consuming linked data (COLD 2012), Boston, Nov 2012, https://km.aifb.kit.edu/ws/cold2012/. hal-01171125

33. World Bank (2014) Open data for economic growth. See also http://www.worldbank.org/content/dam/Worldbank/document/Open-Data-for-Economic-Growth.pdf. Accessed 29 Nov 2017

34. Ball A (2014) How to license research data. A Digital Curation Centre and JISC Legal 'working level' guide. http://www.dcc.ac.uk/resources/how-guides/license-research-data. Accessed 29 Nov 2017.

35. Rotolo A, Villata S, Gandon F (2013) A deontic logic semantics for licenses composition in the web of data. In: Proceedings of the fourteenth international conference on artificial intelligence and law, ICAIL '13. ACM, New York, pp 111–120. https://doi.org/10.1145/2514601.2514614

36. Guido G, Ho-Pun L, Antonino R, Serena V, Fabien G (2013) Heuristics for licenses composition. Frontiers in artificial intelligence and applications. pp 77–86. https://doi.org/10.3233/978-1-61499-359-9-77

37. Cabrio E, Palmero Aprosio A, Villata S (2014) These are your rights. In: Presutti V, d'Amato C, Gandon F, d'Aquin M, Staab S, Tordai A (eds) The semantic web: trends and challenges. Springer International Publishing, Cham, pp 255–269