BOLD Vision 2020

Designing a vision for the future of Big Open Legal Data

Wass, C.; Sageder, C.; Winkels, R.; van Eechoud, M.; Marsden, C.; Guadamuz, A.; Dini, P.; van der Graaf, S.; Passani, A.; Heistracher, T.; Hirsch, E.; Marcon, G.; Zanioli, M.

DOI
10.5281/zenodo.158999

Publication date
2016

Document Version
Final published version

openlaws

BOLD Vision 2020

Designing a Vision for the Future of Big Open Legal Data

openlaws.eu is co-funded by the European Union
under Civil Justice Action Grant JUST/2013/JCIV/AG

BIG OPEN LEGAL DATA

Clemens Wass, Christian Sageder, Radboud Winkels, Mireille van Eechoud, Chris Marsden, Andres Guadamuz, Paolo Dini, Shenja van der Graaf, Antonella Passani, Thomas Heistracher, Eduard Hirsch, Giulio Marcon, Matteo Zanioli

BOLD Vision 2020

Designing a Vision for the Future of Big Open Legal Data

University of Amsterdam | University of Sussex | Salzburg University of Applied Sciences| London School of Economics | Alpenite | BY WASS

Deliverable 4.4.d3 - BOLD Vision

openlaws.eu is co-funded by the European Union / DG Justice
under Civil Justice Action Grant JUST/2013/JCIV/AG

Project start date: April 1, 2014

Project end date: March 31, 2016

Participants: Universiteit van Amsterdam (UVA), University of Sussex (SUSS), London School of Economics and Political Science (LSE), Fachhochschule Salzburg GmbH (SUAS), Alpenite srl (ALP), BY WASS GmbH (BYW)

Disclaimer:

www.openlaws.eu

Table of Contents

Building better access to justice based on legal technology and community collaboration.

openlaws.eu

Executive Summary

The vision of openlaws.eu is to make access to justice easier for citizens, businesses and legal experts. For this purpose, an innovative legal information platform has been designed by the openlaws.eu project, considering the needs of various stakeholder groups as well as the latest developments in technology and our information society.

Access to justice is a fundamental problem in the European Union. There are over 500 million citizens and over 21 million businesses who live, work and operate in 28 jurisdictions, written in 24 official languages. A common market cannot work without a legal system as a basis. Legal information is a public good and it is the duty of governments and the EU to inform citizens and business about the law. In a democracy and under the rule of law everybody should know legislation and case law – or at least have access to it.

Legal tech is a new terms for new technology that can be applied to legal information in order to create better access and better understanding of the law. However, just because things can be done, does not mean automatically that they are done. Financial and organisational restrictions and the lack of competency can be a deal-breaker for innovation. Open data, open innovation and open source software can be a potential solution to this problem, especially when combined to one coherent ecosystem.

openlaws.eu has developed a prototype platform upon these new open concepts. The application and implementation of some of the features of this innovative legal cloud service indicate where the road of “Big Open Legal Data” can lead us in the upcoming years. The project team envisages an environment, where a “social layer” is put on top of the existing “institutional layer”. Citizens, businesses and legal experts can actively collaborate on the basis of primary legislation and case law. Linked and aggregated legal data provide a solid basis. Such information can then be represented in traditional and more innovative ways. Text and data mining as well as legal intelligence help to process large amounts of legal information automatically, so that experts can focus on the more complicated questions.

In the next five years more and more legal data will be opened up, not only because of the PSI Directive, but also because it is in the best interest of governments. As a result, we anticipate that more legal tech start-ups will emerge, as already happened during the past two years. They will apply innovative concepts and new technology on existing legal information and create better access to justice in the EU, in Member States and in the world.

Data is a raw material for information businesses, just as oil is a raw material for fuel and plastic businesses.

Neelie Kroes

Big Open Legal Data

More and more sources of law are (freely) available online in Europe and the rest of the world. It concerns both legislation and case law and possibly others like legal commentaries. Professionals may already get lost in the multitude of information, let alone ordinary citizens and (small and medium sized) enterprises. Traditionally commercial publishers provide this information, plus support for access. Legal experts write commentaries for them, editors provide links between different (types of) sources and warn subscribers for interesting new developments and case law. Now that large amount of sources of law become electronically available online, the question is whether new ways for supporting access can be developed. One stream of research may be directed towards using the 'wisdom of the crowd', have users of legal information share their collections of material, the links they see between different sources, their commentaries, etc. Another stream of research is directed at (semi-) automated linking and clustering of sources of law, analysis of the network of law to find authoritative sources or predict the change of opinion of higher courts, etc. All these approaches can be useful to build Legal Tech applications.

The legal data out there comes in many tastes and varieties. Not only the licenses under which they can (or cannot) be used differ, but also the formats in which they are published. Everything from scanned PDF documents to full machine readable linked data can be found.

Big Data

Legal data is not big data. At least not according to common definition. However, legal data is still ‚big enough' data, which makes it hard for people to comprehend it as a whole. Unlike other data, legal data is hardly structured, making it difficult to apply big data and statistical analyses. Still, big data is supposed to be the future and will change society.

But what is 'big data' anyway? Without context one could argue that 'big data' is a meaningless buzz-phrase. Big data is taken to refer to data which is machine-readable,

Legal Tech

Legal technology refers to the use of technology and software to provide legal services, similar to the use of Fin Tech for financial services. Often the term is used in combination with disruptive solutions.

interoperable and often non-proprietary (shared), as well as having some minimum size (its 'bigness') in terms of objects or files. Laney in 2001 defined data growth challenges and opportunities as being three-dimensional, increasing

- volume (amount of data),
- velocity (speed of data in and out), and
- variety (range of data types and sources).

His company continue to use this '3Vs' model for describing big data, and updated its definition as follows:

> *"Big data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization."*[1]

The European Commission has taken this definition and added adaptations with further 'V's for Veracity and Value (McKinsey having estimated a potential European market of €250 billion).[2] The dynamic development of Information and Communications Technologies (ICTs) has the potential to lower information, and hence transaction, costs. Couldry and Powell state:

> *"It is digital infrastructures of collection, transmission, analysis and presentation that have made possible continuous data-mining. Compared to representative sampling, such new approaches to data collection are totalising; they are also characterised by the aggregation of multiple data sets through the use of calculation algorithms."*[3]

The key to the transformative effect of these productivity gains is that networks increase the productivity effect with each new addition to the network (this is known as Metcalfe's Law), thus creating a 'bandwaggoning' growth in data transfer and processing.

Several 'laws of the network and device' have 'network effects' on the others – high processing speed (Moore's Law) and storage (Disc Law[4]) are needed to process and store the highly compressed (Compression) data files sent via switchers (Gilder's

1 *Laney, Douglas* (2001) File 949 "3D Data Management: Controlling Data Volume, Velocity and Variety". Gartner 6 February at http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf *Laney, Douglas* (2012) "The Importance of 'Big Data': A Definition" Gartner at http://www.gartner.com/document/2057415

2 See European Commission (2014) Big Data: Futurium at http://ec.europa.eu/digital-agenda/futurium/en/content/big-data See further COM/2014/0442 final Towards a thriving data-driven economy.

3 Couldry, Nick and Alison Powell (July 1, 2014) Big Data from the bottom up, Big Data & Society July-December 2014 vol. 1 no. 2 2053951714539277 doi: 10.1177/2053951714539277

4 See further, *Seeley Brown, John* (2002) The Social Life of Innovation in the Digital Age, 15 July presentation at http://www.ruschlikon.net/INTERNET/rschwebp.nsf/(ID)/6C5A73B4FEBA95A9C1256C13002820A2/$-FILE/JSB-speech-.pdf

Technical Process	Component	Cost Effect
Moore's Law	Microprocessor	Doubles every 18 months e.g. from 2GHz to 4GHz
Metcalfe's Law	Network	Increases potential value of network by square of number of nodes – any new user is both receiver and sender of information e.g. e-mail
Disc Law	Storage - hard disk	Doubles storage cost-efficiency each year
Data Packet Transfer	Data compression	Increases: boosted by improved codecs e.g. DivX, H.260, MPEG4
Gilder's Law	Transmission equipment	Potential bandwidth increases three times faster than microprocessor power – Moore's Law x3 – every 6 months
Fibre Law	Transmission network	Total capacity doubles every 9 months
Linus' Law	Problem solving	Given enough eyeballs, all bugs are shallow

Table 1: Laws of networks and devices

Law[5]) and optical fibre (Fibre Law) over the network (Metcalfe's Law).[6] The open source software community also relies heavily on Linus' Law on collaboration: Given a large enough beta-tester and co-developer base, almost every problem will be characterized quickly and the fix will be obvious to someone.

This combination of ultra-powerful ubiquitous computing and even stronger network effects creates the dynamic for an extraordinary growth in 'bandwidth' transfer and data storage: the capacity of the Internet community to communicate. The move from grid computing to network computing to cloud computing in the 21st century has seen the deployment of larger shared programming tasks between federated and even semi-autonomous machines.[7] While a single super-computer might have been the necessary basis for what was then considered 'big data' in the 1980s, by the 21st century it was clear that federated computing structures are necessary for larger tasks.

An example of big data processing from the early consumer broadband Internet experience would be the distributed computing shared-resource programme for SETI@home (Search for Extra Terrestrial Intelligence) launched in 1999, which had

5 See *Gilder, G.* (1993) Metcalfe's Law and Legacy, Forbes ASAP 13 September at http://www.seas.upenn.edu/~gaj1/metgg.html

6 See further, *Seeley Brown, John* (2002) The Social Life of Innovation in the Digital Age, 15 July presentation at http://www.ruschlikon.net/INTERNET/rschwebp.nsf/(ID)/6C5A73B4FEBA95A9C1256C13002820A2/$-FILE/JSB-speech-.pdf

7 See Winshuttle (2014 undated) Big Data Timeline for an excellent visual summary: http://www.winshuttle.com/big-data-timeline/

145,000 active computers in the system (1.4 million total) in 233 countries, as of 23 June 2013, with the ability to compute over 722 teraFLOPS (approximately 60 times greater than the world's most powerful super-computer).[8]

Based on earlier analyses of use of Big Data, we define data sharing and re-use as: "The active cooperation of two or more bodies to exchange or compare data." We define data federation as: "The merging of that data to produce new forms, services or applications of data, whether for private (controlled) or public (open) use." Finally, we define 'data mashing' (whether for Public Sector Information or other data) as: "A particular type of data sharing based on common use of published and accepted Asynchronous JavaScript and XML (AJAX) software family data standards." Data mashing has become associated with overblown claims as to its potential and current value through the use by proponents of 'Web 2.0' services and applications. O'Reilly states:

> *"The potential of the web to deliver full scale applications didn't hit the mainstream till Google introduced Gmail, quickly followed by Google Maps, web based applications with rich user interfaces and PC-equivalent interactivity."*[9]

More prominently, Apple advertised its iTunes service as offering consumers the ability to "Rip. Mix. Burn" in 2006.[10] This report uses the term 'data mashing' to describe any Internet-based federation of two or more data types using existing tools to remove technical standardisation as a barrier to service delivery. The public are important re-users or 'prosumers' as well as consumers of data. The user is able to 'pull' content and adapt and mix content into a user's own 'mash-up'. A mash-up is a combination of existing media reworked into a potentially innovative type.

It is a practical application of big data capability[11] to create the public good of access to law. There is a tension between the possibilities offered by legal information re-use and the barriers to implementing (or even conceiving) this via open data.

8 Tianhe-2, the world's fastest supercomputer, was able to compute 33.86 petaFLOPS in 2013, see http://www.top500.org/lists/2014/06/ See http://boincstats.com/en/stats/0/project/detail for latest SETI@home data.

9 *O'Reilly, Tim* (2007) What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. Communications & Strategies, No. 1, p. 17. Available at SSRN: http://ssrn.com/abstract=1008839

10 See http://www.criticalcommons.org/Members/ccManager/clips/apple-ad-rip.-mix.-burn

11 'Big Data' is a cliché capable of many definitions – see for instance leading scientists' Challenges and Opportunities with Big Data (undated 2012) building on *Lazer, et al* (2009). For provocations based on the role of inter-disciplinary analysis of Big Data, see *Haddadi, et al* (2013).

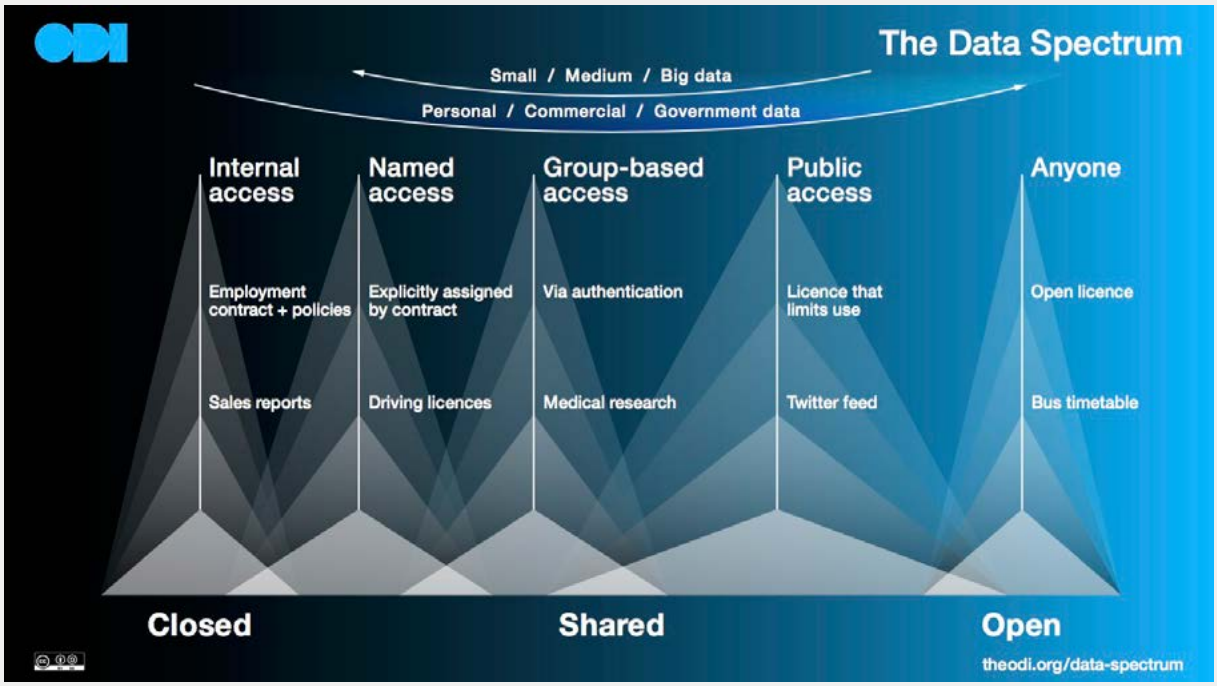


Figure 1: The Data Spectrum (ODI)

Open Data

What is open data and what can we do with it? The European Commission generally states "Open data refers to the idea that certain data should be freely available for use and re-use".[12] The UK government defines Open Data as

> *"Data which can be used, re-used and re-distributed freely by anyone – subject only at most to the requirement to attribute and share-alike. There may be some charge, usually no more than the cost of reproduction".*[13]

Open data has recently been defined by the Open Data Advisory Council as "principles that define openness in relation to data and content," with a license which permits anyone to "freely access, use, modify, and share that content, for any purpose, subject, at most, to requirements that preserve provenance and openness."[14]

12 https://ec.europa.eu/digital-agenda/en/open-data-0
13 Source: APPSI Glossary http://www.nationalarchives.gov.uk/appsi/appsi-glossary-a-z.htm
14 Open Definition Version 2.0 which allows for identification of rights in works as well as licences: see http://opendefinition.org/ and *Vollmer, Timothy* (2014) Open Definition 2.0 released, Creative Commons Public Policy October 7th, at http://creativecommons.org/weblog/entry/43812

Open Data Institute

Founded by Sirs Tim Berners-Lee and Nigel Shadbolt, the ODI is an independent, non-profit, non-partisan, limited by guarantee company. The ODI has secured £10m over five years from the UK Government (via the UK innovation agency, Innovate UK), and $4.75m from Omidyar Network, and is working towards long-term sustainability through match funding and direct revenue.

The ODI has identified open data-driven UK companies with a combined annual turnover of over £92bn employing over 500,000 people. Transport for London alone has identified a 58:1 return on investment by releasing transport data, in the process helping create global leaders such as Citymapper. Denmark has seen a 70:1 return on investment by choosing to publish address data openly. In the US, an open data company has sold for $930m and Landsat data create savings of $350m to $436m per year, while at least 84% of American smartphone owners use an application powered by open data every single day.

http://theodi.org

Open Knowledge International

Open Knowledge International is a worldwide non-profit network of people passionate about openness, using advocacy, technology and training to unlock information and enable people to work with it to create and share knowledge.

Open Knowledge wants to see enlightened societies around the world, where everyone has access to key information and the ability to use it to understand and shape their lives; where powerful institutions are comprehensible and accountable; and where vital research information that can help us tackle challenges such as poverty and climate change is available to all.

Open Knowledge International envisions a world where:

- knowledge creates power for the many, not the few.
- data frees us to make informed choices about how we live, what we buy and who gets our vote.
- information and insights are accessible – and apparent – to everyone.

https://okfn.org

Open data is meaningful in terms of openlaws.eu only with the creation of machine-readable 'mashable' data.[15] Without the Internet and WWW, there would be much less open digital data sharing possible of any scale, which makes open data young in this sense. The World Wide Web was designed for open data standards, and its creator Sir Tim Berners Lee is co-founder of the UK Open Data Institute. The European Commission has funded the ODI and Berners Lee to set up the Open Data Incubator for Europe (http://opendataincubator.eu/) as part of its open data initiatives.[16] The WWW was created at CERN (Centre Europeene pour la Recherche Nucleaire), home to the Large Hadron Collider, whose sub-atomic experiments gather 500 exabytes of data per day in raw form (a multiple of 200 of all the data actually created and stored in the entire world each day).[17]

Access to knowledge is now recognized as a key driver of social, cultural and economic development, with tangible economic advantages to be gained by sharing.[18] The 'all rights reserved' model of traditional copyright law, with its legal concepts and requirement for permission for all uses, does not fit well with an environment which enables sharing and reuse of content by users. In the analogue environment the ability to produce, reproduce, distribute, share and promote creative works was relatively restricted, due primarily to geographic, economic and technological limitations. The emergence of consumer digital technologies such as CDs and the internet in the 1990s allowed for increasing levels of functionality, particularly in relation to interactivity as Benkler indicated in 2006. More recent production and communication technologies – mobile phone cameras, mp3 encoding for music, rich media applications, video streaming and peer-to-peer networking provide simple ways for users to collaborate, communicate and create material, including 'mashing up' existing material into new and innovative media.

15 *Kimpton, P.* (2013) Obama to Berners-Lee, Snow to Domesday: a history of open data, Guardian 25 October at http://www.theguardian.com/news/datablog/2013/oct/25/barack-obama-tim-berners-lee-open-data

16 Also funding the European Data Science Academy (EDSA): see *Gibbs, Samuel* (2014) EU commits €14.4m to support open data across Europe, The Guardian 4 November, at http://www.theguardian.com/technology/2014/nov/04/eu-commits-144m-to-support-open-data-across-europe

17 *Brumfiel, Geoff* (2011) "High-energy physics: Down the petabyte highway" Nature 19 January v.469. pp. 282–83. doi:10.1038/469282a

18 See a robustly optimistic assessment and for policy makers of great importance, *Weiss, Peter N.* (2002) Borders in Cyberspace: Conflicting Public Sector Information Policies and their Economic Impacts, US Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, http://www.weather.gov/sp/Borders_report.pdf

Pam states in the publication "Hyperdistribution":

> *"With the advent of the Internet, a global network providing the capability to the general public for peer-to-peer transfer of digital media, it no longer makes sense for the media industry to use the existing producer/publisher/distributor/consumer one-way pipeline business model since a larger proportion of the public are capable, willing and interested to act as producers, publishers and distributors."*[19]

The risk in such an environment is that copyright law will become a barrier to the realization of the full potential of these technologies. A significant legal response to such a challenge has been the development of new licensing systems to open up access to and use of protected material. Access to and re-use of materials produced by government and other publicly funded bodies has also emerged as an important issue in recent years. Historically it has been cumbersome and expensive to provide access to government information. However digital technologies have now removed many traditional barriers to widespread distribution of material to the public. As a result, consumer demand for access to, and reuse of, government information has risen exponentially, driven in part by the emergence of Web 2.0 functionality.

G8 nations signed an Open Data Charter in 2013 which contains five principles for re-use of Public Sector Information.[20] Open data principles are normative claims of activists presented as descriptors; the principles are about what should be, and to some degree governments agree. The Sebastopol Principles of December 2007 [21]are particularly designed for government data, and are not fully transferrable to all legal data, for instance licensing (particularly copyright) is favoured by many official as well as private legal sources, breaching Principle 8.[22] Note also the 'Five-star' approach to open data used by the World Wide Web Foundation, designed by Sir Tim Berners Lee "in order to encourage people - especially government data owners - along the road to good linked data".[23]

In the broadest sense, openness in the context of access to information is defined by the Open Knowledge Foundation (OKFN, now Open Knowledge International) as being able to freely access, use, modify, and share such information for any

19 *Pam, A.* (2002) "Hyperdistribution" Serious Cybernetics at http://www.sericyb.com.au/hyperdistribution.html

20 G8 (2013) Open Data Charter, 18 June, at https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex

21 https://public.resource.org/8_principles.html

22 *Chignard, Simon* (2013) A brief history of Open Data, Paris Tech Review March 29, at http://www.paristechreview.com/2013/03/29/brief-history-open-data/

23 *Berners Lee, Tim* (2006 amended 2010) Linked Open Data, at http://www.w3.org/DesignIssues/LinkedData.html See further for its implementation http://5stardata.info/en/

1	Complete	All public data is made available. Public data is data that is not subject to valid privacy, security or privilege limitations.
2	Primary	Data is as collected at the source, with the highest possible level of granularity, not in aggregate or modified forms
3	Timely	Data is made available as quickly as necessary to preserve the value of the data.
4	Accessible	Data is available to widest range of users for the widest range of purposes.
5	Machine processable	Data is reasonably structured to allow automated processing
6	Non-discriminatory	Data is available to anyone, with no requirement of registration
7	Non-proprietary	Data is available in a format over which no entity has exclusive control.
8	License-free	Data is not subject to any copyright, patent, trademark or trade secret regulation. Reasonable privacy, security and privilege restrictions may be allowed.

Table 2: Sebastopol Principles 2007

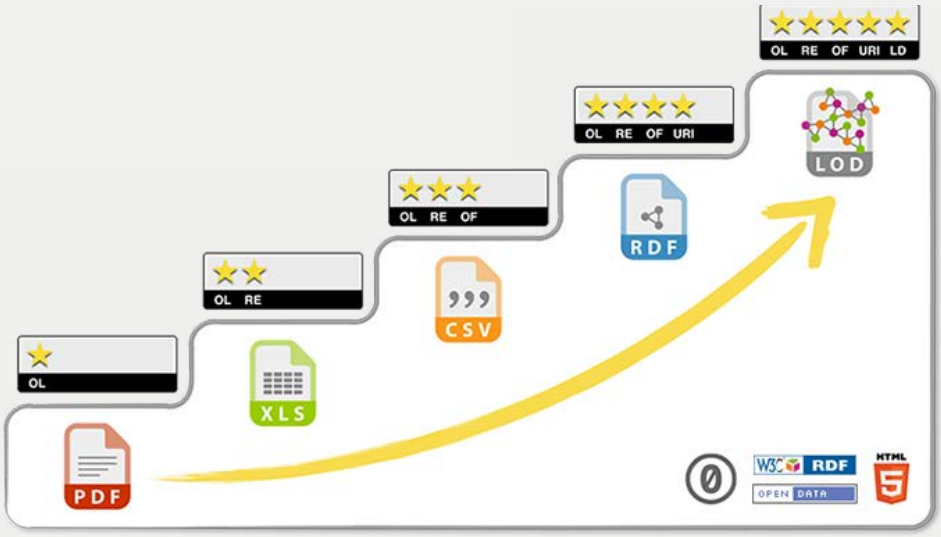


Figure 2: 5-star deployment scheme for open data by Sir Tim Berners-Lee

1. Make your stuff available on the Web (whatever format) under an open license
2. Make it available as structured data (e.g., Excel instead of image scan of a table)
3. Make it available in a non-proprietary open format (e.g., CSV as well as of Excel)
4. Use URIs to denote things, so that people can point at your stuff
5. Link your data to other data to provide context

purpose. In the specific case of open data, this would translate into data that is released using an open licence, which is a document that has to fulfil several requirements to meet the definition.

Open licences must (amongst other requirements):

- allow free use of the licensed work;
- allow redistribution of the licensed work, including sale, whether on its own or as part of a collection made from works from different sources;
- allow the creation of derivatives of the licensed work;
- allow any part of the work to be freely used, distributed, or modified separately from any other part of the work or from any collection of works in which it was originally distributed;
- allow the licensed work to be distributed along with other distinct works without placing restrictions on these other works;
- must not impose any fee arrangement, royalty, or other compensation or monetary remuneration as part of its conditions.[24]

Open data is important not only because it offers the possibility for citizens to read public sector data (whether research or not) but also because it affords the citizen the ability to 'prosume' – to recombine the data with other data sets into new innovative uses for that data.[25] However, the funding argument is controversial for a number of domains. It assumes all information production by public sector is funded through general taxation rather than also via user pays / specific levies/ taxes but such models (e.g. for companies registries, land ownership registries) were introduced to make the public service producer work more efficiently (price as mechanism to determine demand/market need & to allocate costs where benefit arises). There is substantial tension with the "open data for free" model because they cannot be maintained simultaneously.

Open access does not necessarily mean open data. It is in essence a funding model whereby the producer/author of information pays to have it published and the reader/consumer gets free access. The European Commission has recently

24 Open Knowledge Foundation. (2014) Open Definition, http://opendefinition.org/od/.

25 *Coughlan, S. ed.* (2014) Research On Open Innovation, Openforum Academy Publication at http://openforumacademy.org/research/research-on-open-innovation especially *Adams, A.* (2014) The Open versus Closed Debate, Chapter 10 pp.161-198 and *Katz, A.* (2014) Blurring the Line between Creator and Consumer pp.199-217 in *Coughlan (ed)*.

published several papers, a Communication,[26] Recommendation and Expert Report encouraging the use of open access[27] in its funded science research, in government data and procurement,[28] as well as funding pilot Open Access projects in its FP7 and Horizon 2020 research programmes.[29] Alongside law-specific developments is the wider development of the 'science commons', led by reform of access terms to scientific data. Dulong de Rosnay and Martin document this development in the European context.[30]

Discussion of PSI comes from EU policy in 1989 to stimulate re-use of public sector data resources by private sector[31], leading eventually to the 2003 PSI Directive.[32] The link between Big Data (as described above mainly in terms of technological drivers) and PSI is obvious in that the largest publicly accessible datasets have been made available by government. Re-use policy is about extracting more value from public sector data beyond public task. But data sharing, federating and mashing are valuable to increase public sector task efficiency. The evolution of new PSI re-uses faces specific technological, socio-institutional, economic and legal hurdles. The barriers are gradually being tested and – in some cases – overcome through public and private initiatives.

PSI Directive

Directive 2003/98/EC on the re-use of public sector information (amended by Directive 2013/37/EU) provides a common legislative framework for making public sector information available for re-use.

Legislation and case law are good examples for such public sector information, as they are essential information for a working democracy.

26 COM(2012) 401 final, Towards better access to scientific information: Boosting the benefits of public investments in research. See for follow-up actions http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1301&lang=1

27 IP/12/790: Scientific data: open access to research results will boost Europe's innovation capacity of 17/07/2012 at http://europa.eu/rapid/press-release_IP-12-790_en.htm

28 *Buhr, Carl-Christian* (2012) Open Access in Europe. On the Road to 2020, Nov 23, 2012 at http://www.slideshare.net/ccbuhr/open-access-in-europe-on-the-road-to-2020?related=1

29 EC (undated) Open Access in FP7, at http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1300&lang=1

30 *Dulong De Rosnay, M., & De Martin, J. C.* (2012). The Digital Public Domain: Foundations for an Open Culture, at p135, 'Open Access Content'. Available at http://www.communia-association.org/wp-content/uploads/the_digital_public_domain.pdf

31 Commission of the European Communities (1989) Guidelines for Improving the Synergy Between the Public and Private Sectors in the Information Market, ISBN 92-825-9237-5 Catalogue number: CD-54-88-126-EN-C at http://ec.europa.eu/digital-agenda/en/news/guidelines-improving-synergy-between-public-and-private-sectors-information-market

32 See *Janssen, Kathleen* and *Dumortier, Jos* (2003) Towards a European Framework for the Re-use of Public Sector Information: A long and winding road, International Journal of Law and Information Technology Vol. 11 No. 2. See more at: http://www.epsiplatform.eu/content/eu-psi-directive-200398ec#sthash.vyxI-6fiY.dpuf

Legal Data

The justice systems keep the European Union together. As Jean-Claude Juncker, President of the European Commission has put in hi political guidelines:[33]

> *As citizens increasingly study, work, do business, get married and have children across the Union, judicial cooperation among EU Member States must be improved step by step: by building bridges between the different justice systems, by strengthening common tools [...]*

Legal data encompasses a three-part categorisation: legislation, case law and literature. There are several sub-fields which encompass the various regulatory and soft law documents that occupy the gaps between these three main categorisations. Legislation is collated by the government, case law has a less structured pattern of publication, and literature or commentary is found in learned journals and books, speeches by judges, online resources and the guidance issued by various bodies, notably Law Commissions, Law Societies/Bar Associations (LS/BAs), ministries and prosecutors, and other authoritative sources (authority claims combine expertise, organisation and venue of publication).

Legal texts are basic information of all democratic states. The Aristotlean argument is that everyone is presumed to know the law: "Ignorantia juris non excusat (ignorance of the law is no excuse)". Legal information must be accessible to all members of society to the widest possible extent, to aid inclusiveness and enable participation in public decision-making. A section of this report explains the European situation with regard to reuse of public sector materials more generally.

In recognition of the public good in access to legal information, the EU and its Member States work to make laws, court decisions, etc. publicly available on line. Much has been achieved locally already. However, the sheer mass of legal norms, instruments and interpretations in courts decisions, commentaries and other sources makes it increasingly difficult for citizens, civil society, businesses and all involved in legal practices to locate the relevant law.

The challenge for the future is to link local legal information and have in place structures to enrich it through aggregation and mass customization. The technological possibilities to achieve this are there. This contributes to better access to legal information and ultimately to better governance, both of which support higher social welfare goals.

33 A new start for Europe - Jean-Claude Juncker's Political Guidelines - 15/07/2014, http://ec.europa.eu/news/eu_explained/140715_en.htm

Much legal information remains published and administered by a limited number of organizations, typically in closed structures in public authorities and public private partnerships. This includes the management of legal meta data, which is the basis for automated processing. Legal scholars and practitioners publish mainly through traditional highly specialized commercial publishing or isolated websites. Return channels and interactivity with users are limited, and there is little space for contributions from wider communities. Fully automated processing of legal data is not yet possible. Strikingly, whereas in many domains such as spatial information (see INSPIRE[34]) and life sciences research data, open information infrastructures are rapidly developing, this is not the case for legal information.

FALM & LIIs

The Free Access to Law Movement (FALM) is an international movement and organization devoted to providing free online to legal information such as case law, legislation, treaties, law reform proposals and legal scholarship. Some FALM projects incorporate "Legal Information Institute" or "LII" in their names, usually prefixed by a national or regional identifier.

http://www.fatlm.org/

Open legal data is of relatively recent vintage. The Free Access to Law Movement (FALM) dates to pioneering efforts by Legal Information Institutes in the United States at Cornell[35], in Australia[36], in the United Kingdom[37], and elsewhere, documented on a continual basis in the Law via the Internet (LVI) annual conferences[38], whose twentieth anniversary will occur in 2015.

34 Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), OJ L 108, 25.4.2007, p. 1–14, entered into force 15 May 2007

35 See from a US perspective *Danner, Richard A.* (2012) 'Open Access to Legal Scholarship: Dropping the Barriers to Discourse and Dialogue' 7(1) JICLT 65, 65. Cornell University Law School Legal Information Institute <http://www.law.cornell.edu> accessed 1 June 2013. *Carroll, Michael W.* (2006) 'The Movement for Open Access Law' 10 *Lewis & Clark L.* Rev. 74; *Brady Kevin P.* and *Bathon, Justin M.* (2012) 'Education Law in a Digital Age: The Growing Impact of the Open Access legal Movement' 277 Ed Law Rep 589.

36 *Greenleaf, G.* (1995) ""Public legal information" via Internet: AustLII's first six months." Law Technology Journal. 4(2), 510 West *Knights, L. J.* (1997) "The AustLII paradigm." Journal of Information, Law & Technology. 3. *Kendrick, R.* (1999) "Australia Free(law) for all. Solicitors Journal. 143(48), 1204 *Allen, R.* (2000) "With a wysh and a prayer: an experiment in cooperative development of legal knowledge bases." Journal of Information, Law & Technology. 2, Internet Australian Legal Information Institute (AustLII) experiment in collaborative construction of legal knowledge bases over Web *Baski, C.* (2005) "News Infotech: Commonwealth legal breakthrough" The Law Society Gazette. 29 Sept, 10 (1) briefly outlines the establishment and ethos of CommonLII.

37 See *Brooke, Sir Henry* and *Nick Holmes* (2011) Judgment Day for BAILII, interview with retiring chairman of the BAILII trust, at h ttp://www.scl.org/site.aspx?i=ed22972. See *Fellows, Cynthia, Philip Leith* and *Joe Ury* (2012) 'Assessing BAILII 2012' 12(3) LIM 148. See also *Philip Leith* and *Cynthia Fellows*, 'Enabling Free Online Access To UK Law Reports: The Copyright Problem' (2010) 18 IJLIT 72. *Whittle, Steven* (2012) 'Amicus Curiae Pro Bono Publico: Open Access Online Publication at the Institute of Advanced Legal Studies' 12(3) LIM 189.

38 See for instance Jersey edition in 2013: http://www.jerseylvi2013.org/

FALM: Declaration on Free Access to Law

Legal information institutes of the world, meeting in Montreal, declare that:

- Public legal information from all countries and international institutions is part of the common heritage of humanity. Maximising access to this information promotes justice and the rule of law;
- Public legal information is digital common property and should be accessible to all on a non-profit basis and free of charge;
- Organisations such as legal information institutes have the right to publish public legal information and the government bodies that create or control that information should provide access to it so that it can be published by other parties.

Public legal information means legal information produced by public bodies that have a duty to produce law and make it public. It includes primary sources of law, such as legislation, case law and treaties, as well as various secondary (interpretative) public sources, such as reports on preparatory work and law reform, and resulting from boards of inquiry. It also includes legal documents created as a result of public funding.

Publicly funded secondary (interpretative) legal materials should be accessible for free but permission to republish is not always appropriate or possible. In particular free access to legal scholarship may be provided by legal scholarship repositories, legal information institutes or other means.

Greenleaf et al. identify six historic attempts to achieve FALM:

- the example set by the LII (Cornell) and LexuM in the early 90s;
- AustLII's 1995 formulation of the obligations of official publishers;
- the 2002 Declaration on Free Access to Law;
- the 'Guiding Principles' for States formulated by a 2008 expert meeting convened by the Hague Conference on Private International Law;
- the 'Law.Gov principles' developed by Public Resources.org in 2010; and
- the draft Uniform Electronic Legal Materials Act recommended in 2011 by the US National Conference of Commissioners of Uniform State Laws.[39]

39 *Greenleaf et al* (2012) supra.

Legal information institutes:

- Publish via the internet public legal information originating from more than one public body;
- Provide free and anonymous public access to that information;
- Do not impede others from obtaining public legal information from its sources and publishing it; and
- Support the objectives set out in this Declaration.
- All legal information institutes are encouraged to participate in regional or global free access to law networks.

Therefore, the legal information institutes agree:

- To promote and support free access to public legal information throughout the world, principally via the Internet;
- To recognise the primary role of local initiatives in free access publishing of their own national legal information;
- To cooperate in order to achieve these goals and, in particular, to assist organisations in developing countries to achieve these goals, recognising the reciprocal advantages that all obtain from access to each other's law;
- To help each other and to support, within their means, other organisations that share these goals with respect to:
    - Promotion, to governments and other organisations, of public policy conducive to the accessibility of public legal information;
    - Technical assistance, advice and training;
    - Development of open technical standards;
    - Academic exchange of research results.
- To meet at least annually, and to invite other organisations who are legal information institutes to subscribe to this declaration and join those meetings, according to procedures to be established by the parties to this Declaration;
- To provide to the end users of public legal information clear information concerning any conditions of re-use of that information, where this is feasible;

The parties to this Declaration also support the principles stated in the 'Guiding Principles' on State obligations concerning free access to legal information developed by an expert group convened by the Hague Conference on Private International Law in October 2008, and the 'Law.Gov principles' for 'the dissemination of primary legal materials in the United States' developed in 2010 by Public Resources.org.

This declaration was made by legal information institutes meeting in Montreal in 2002, as amended at meetings in Sydney (2003), Paris (2004), Montreal (2007) and Ithaca (2012).

http://www.fatlm.org/declaration/

openlaws.eu documented individual FALM efforts and LIIs in several country case studies. A brief summary is that the effect of LIIs has been strongest in common law jurisdictions notably Australia led by pioneer Graham Greenleaf[40], most profound in Canada[41] where CANLII and LexuM is supported by the professional bar,[42] and of variable impact elsewhere[43]. Sustainability and professional adoption remain challenges.[44]

Big data for legal informatics pre-dates the consumer Internet, with the Free Access to Law Movement (FALM) dating to the early 1990s when only corporations, governments and universities had high bandwidth networks capable of sharing such larger data sets. Much of the historic pre-Internet discussion of legal informatics relates to the effect of digital information retrieval on the work of lawyers and courts.

Bing explained that the origins of legal informatics effects in access to law date from the 1970s in pioneering academic-professional collaborations.[45] Biegel lays out the effects of the Internet on the usefulness of traditional enforcement techniques across several branches of the law, following the pioneering work of Berring.[46] Katsch's pre-Internet but very Internet-aware critique of print media and transformative effect of digital information on the law, states:

---

40 See *Greenleaf, Graham* (2010) 'The Global Development of Free Access to Legal Information' pp5382 in *Abdul Paliwala (ed)*, A History Of Legal Informatics (Prensas Universitarias de Zaragoza) *Greenleaf, Graham, Philip Chung and Mowbray, Andrew* (2005) 'Emerging Global Networks for Free Access to Law' WorldLII's Strategies' 1(1) J of Electronic Resources in L Lib <http://ssrn.com/abstract= 975614> *Graham Greenleaf, Andrew Mowbray and Philip Chung* (2011), 'AustLII: Thinking Locally, Acting Globally', (2011) 19(2) Australian Law Librarian 101, 101115.

41 *Poullin, D.* (2004) "CanLII: how law societies and academia can make free access to the law a reality." Journal of Information, Law & Technology. 1, *Miller, J.* (2008) "The Canadian Legal Information Institute a model for success" Legal Information Management

42 *Daniel Poulin*, 'Free Access to Law in Canada' (2012) 12(3) LIM 165172 sketching the principles supporting free access and also trying to make the business case for establishing it; discussing also the creation of CanLII. On academic interest see *Wilner, Josh* (2008) 'Editor's Note – Open Access to Legal Publishing' 2 McGill J L & Health 1

43 See *Greenleaf, G., Mowbray, A. and Chung, P* (2004) "A new home online for Commonwealth law: a proposal for CommonLII." Journal of Information, Law & Technology. 2, Internet. *Greenleaf, G., Mowbray, A. and Chung, P.* (2010) "Building a commons for the common law: the Commonwealth Legal Information Institute (CommonLII) four years on." Commonwealth Law Bulletin. 36(1), 127134. *Greenleaf, G., Vivekanandan, V. C., Chung, P., Singh, R. and Mowbray, A.* (2011) "Challenges for free access to law in a multijurisdictional developing country: building the Legal Information Institute of India." SCRIPTed. 8(3), Internet. *Greenleaf, G., Chung, P. and Mowbray, A.* (2007) "Emerging global networks for free access to law: WorldLII's strategies 20022005" SCRIPTed. 4(4), Internet. *Greenleaf, G.* (2005) "Global legal research: WorldLII and the future." Internet Newsletter for Lawyers. 2005 Jan/Feb, 13. *Greenleaf, G* (2007) "Networking LIIs: how free access to law fits together". Internet Newsletter for Lawyers. Mar/Apr 2007, 35. *Onwonga, D. A.* (2003) "LawNet initiative: a case for the East African Legal Information Institute." Journal of Information, Law & Technology. Volume 1.

44 See *Marsden et al* (1998) IJCLP Editors' Statement: http://ijclp.net/old_website/1_1998/editors_statement.html

45 See *Bing, J. in Paliwala, Abdul* (2010) [ed] A History of Legal Informatics, Prensas Universitarias de Zaragosa, Spain

46 See *Biegel, Stuart*, Beyond Our Control? Confronting the Limits of Our Legal System in the Age of Cyberspace, MIT Press (2001) building on *Berring, Robert C.* Legal Research and Legal Concepts: Where Form Molds Substance, 75 Cal. L. Rev. 15, 17 (1987). Berring, R. C. (1995) "The current state of networked infor-

"The process of legal readjustment that will be necessary in the future may prove painful to those who idealize the current model of law, who mistakenly associate the rules of law with the rule of law, or who do not understand that what we have now is not perfect and has never been static."[47]

Susskind provides an updated provocation on the possible future effect of informatics on lawyers, including legal publishing.[48] Many have previously reflected on legal information in the wider setting of copyrighted public sector information (PSI) and the challenges of freeing such information (Ubaldi 2013).[49] Many of the early tested ideas have flowered into the wider government #OpenData movement[50] and work pioneered by the European Union 1989 Guidelines (Eechoud 2013), OECD[51] and Gore-Clinton 'National Partnership for Reinventing Government' (1993).

The latest financial estimates for legal information and services are most accurate for the United States market, unsurprisingly given that the US is a single market that is still valued more highly than the entire rest of the world including the EU.[52] According to Reed Elsevier, the second largest private legal information provider, the US market is 53% of the entire global legal services market of $625billion, with the whole of Europe at 30%. The market is growing at about 5% per annum. The same source (slide 10) suggests that the US market accounts for 57% of the entire legal information solutions market of US$18billion, with Europe 30% or $5.4billion in 2011. This market typically accounts for ca. 3% of the global legal services market, which includes law firm revenue and internal corporation and government spend.

mation in the United States and why you should care about it." Law Librarian. 26 (1), 246248. *Bruce, T. R.* (2000) "Tears shed over Peer Gynt's onion: some thoughts on the constitution of public legal information providers." Journal of Information, Law & Technology. 2, Internet

47 *Katsh, Ethan*, The Electronic Media and the Transformation of the Law, Oxford University Press (1989)

48 *Susskind, Richard* (2008) The End of Lawyers? Rethinking the Role of Legal Services, Oxford University Press continues his thesis that lawyers need to focus on added value advocacy skills as much lower value work is automated and can be outsourced to lower cost locations.

49 *Marsden, et al* (2006). For the UK Cabinet Office (2006), the Government Data Mashing project was located in the field of public sector reform and deals with one specific instrument that was high on the agenda for the Cabinet Office "Data Grand Challenge" in the UK: data mashing. The policy analysis was based on law and economics literature (especially on new institutional economics) to analyse barriers to adoption of data mashing. The analysis included that of the type of new institution necessary to overcome existing data sharing barriers. The analysis concluded by identifying outstanding issue areas and need for exploration of further research possibilities, especially including pricing (and 'payorplay') models and legal (e.g. copyright) reforms necessary to ensure further data sharing capability.

50 See for instance http://data.gov.uk/aboutus

51 See http://www.oecd.org/sti/ieconomy/workshoponaccesstopublicsectorinformationandcontent.htm

52 Reed Elsevier (2012) Investor Seminar, 11 October 2012, LexisNexis Legal & Professional

Glassmeyer and Smith state that:

> *"it is nearly impossible to find, cite or read the law in the United States without someone paying a for-profit corporation for the ability to do so".*[53]

At an exchange rate in 2012 of approximately €1=$1.30, that values the European market as a whole at € 4.15 billion in 2012.

Note the dominant use of search engines to find publicly indexed law sources, notably Google. The openlaws.eu 2014 user survey showed that Google remains more used than Lexis-Nexis or any other database. While this is a generic search engine rather than a legal-specific database, its powerful search ability means that it is the largest legal source index in the world, and its advertising-funded business model further distinguishes itself from subscription databases.[54]

Legislation

Legislation is in general not subject to copyright and can be freely reused – though there are exceptions such as the United Kingdom. Australia reformed its copyright for legislation very recently, permitting reuse and the creation of the AustLII database.[55] A relatively comprehensive European source is Eur-LEX, detailed in the EU Cases study.[56] Comparative studies of European legislation show widely divergent practices in publication[57], as do studies of common-law (Anglo-American) legal systems.[58]

In addition to national and European legislation, there is a growing body of subsidiary legislation, from that devolved to nations and regions (for instance autonomous regions such as Comunidad Valenciana or the nation of Scotland), as well as a body of secondary or enabling legislation/regulation that is in many countries much larger than primary (i.e. fully deliberated) legislation. In the UK, there are

53 *Glassmeyer, Sarah and Smith, Peter* (2014) Open law: technology in service of the rule of law, Legal Information Management, 14(3) at p6 in draft: http://shura.shu.ac.uk/view/types/article.html#group_G

54 *Hunter, I.* (2013) "Free legal and official information on the web: is it time to stop Google bashing?" Legal Information Management. 13(4), 256259.

55 *Rubacki Michael* (2013) Free access online legislation in a federation: Achievements of Australian Governments and issues remaining, Presented at the AustLII Research Seminar, 7 May 2013, Australasian Legal Information Institute (AustLII), Sydney, reproduced in 2 Journal of Open Access to Law 1 (2014) at http://ojs.law.cornell.edu/index.php/joal/article/view/9

56 See also the visual representation of the 'Cellar": http://www.w3.org/International/multilingualweb/rome/slides/18-schmitz.pdf

57 See *Donelan, E.* (2009) "European Approaches to Improve Access to and Managing the Stock of Legislation" Statute Law Review. 30(3), 147

58 *Cox, N.* (2006) "Copyright in Statutes, Regulations, and Judicial Decisions in Common Law Jurisdictions: Public Ownership or Commercial Enterprise?" Statute Law Review. 27(3), 185

3300 Statutory Instruments per year, a substantial rise since the 1980s. All instruments since 1987 are available in a much-used public database, legislation.gov.uk.[59]

Database rights and updating procedures vary widely across member states, and the statutory databases may not be up to date at any one time given the restrained resources available to national records offices. As a result commercial offerings may be more reliable indicators of current legislation as updated, such as Halsbury's Laws of England, published since 1929 and now owned by Lexis-Nexis.[60]

The format and clarity of legislation is an important element in its presentation to the legal community and public. If legislation is presented in proprietary formats or in a manner which prevents effective linking (for instance with insufficient XML mark-up to individual sections), that can significantly compromise its usability. The use of common identifiers is essential to allowing re-use of legislative material, and may other good practices have been identified by Poulin.[61] As he explains, better access to legislation is of most use to that section of the public most engaged with legislative interpretation: professional lawyers.

RIS & RIS:App

The Austrian RIS (Rechtsinformationssystem des Bundes) is the official legal information system of Austria, containing legislation and case law. The award-winning system has opened its interfaces via ao-pen data connector, allowing for independent third party application to use the data. The governments provideds the raw material so that new innovative solutions can grow.

These open interfaces led to the implementation of the RIS:App, an app for smartphones and tablet devices, so that legislation is better accessible on mobile devices. Even though the project was set up as a research project and for testing acceptance by users, the RIS:App was already downloaded over 50,000 times.

https://www.ris.bka.gv.at

Legislation is not merely consumed by citizens. Experiments are taking place in 'crowd-sourcing' legislative proposals via the Internet, but citizen-inspired petitions for law-making date to the earliest civilizations. Experiments with using Internet-based discussion to initiate law include the European Citizens' Initiative introduced under the Lisbon Treaty 2009.[62] The most well-known recent examples are

59 See http://www.legislation.gov.uk/uksi

60 See *Hetherington, Simon* (2007) Halsbury's Laws of England: centenary essays, LexisNexis Butterworths.

61 *Poulin, Daniel* (2014), Meaningful public access to legislation, presentationto Nudging Regulations conference, Canadian Institute for the Administration of Justice in Ottawa, September 9th, at https://lexum.com/en/publications/meaningfulpublicaccesslegislation

62 See Regulation (EU) No. 211/2011 of 16 February 2011 on the citizens' initiative, and http://ec.europa.eu/citizensinitiative/public/initiatives/finalised/answered?lg=en

the proposed though non-enacted Icelandic constitution of 2011-12,[63] the many 'federal popular initiatives' in Switzerland where direct democracy has an extensive unbroken history and which actually mandate legislators to enact legislation approved by referendum within a year,[64] and in California and other members of the United States, where the Fifth Amendment to the Constitution permits constitutional amendments subject to ratification by three-quarters of states or of State Ratifying Conventions.

Preparatory works

Alongside primary legal materials, there has been huge growth in secondary material which explains and aids interpretation of that primary material. This may include digitised parliamentary records, which now includes minutes of evidence to parliamentary committees and drafts of reports made available via parliamentary websites. There is also a large amount of material made available for parliamentarians to aid their debates which is then made public via the libraries of parliament[65], for instance detailing Impact Assessment methods used and the intention and debate which led to legislative initiatives.[66] This secondary evidence is vital to the courts and lawyers assessing test cases under new legislation.

In many jurisdictions, the decision to open up parliamentary records/documents in legislative processes was informed by potential of digital technology and shift to digital within administrations/legislatures. Freedom of information (FoI) law has had a transformative effect on the amount of such material placed in the public domain, as has digitisation. Because of trias politica, FOI traditionally do not also cover documents from parliaments, courts (see e.g. Netherlands). The amount of material released through executive agencies at arms' length from ministries has also grown enormously, with for instance the websites of the communications regulators containing a huge amount of regulatory orders (e.g. Body of European Regulators of Electronic Communications). The appeals processes and tribunal data from these agencies has also substantially increased as the regulatory agencies' work has evolved, such that competition lawyers are often more concerned with decisions of the national competition agencies than the courts.

63 *Wasserman, Todd* (2011) Iceland Unveils Crowdsourced Constitution, Mashable, 29 July at http://mashable.com/2011/07/29/iceland-crowdsourced-constitution/

64 For instance, the Executive Pay Law of 2013: http://en.wikipedia.org/wiki/Swiss_executive_pay_referendum,_2013

65 See e.g. European Parliament Research Service (2014, 25 March) Net neutrality in Europe, 140773REV2.

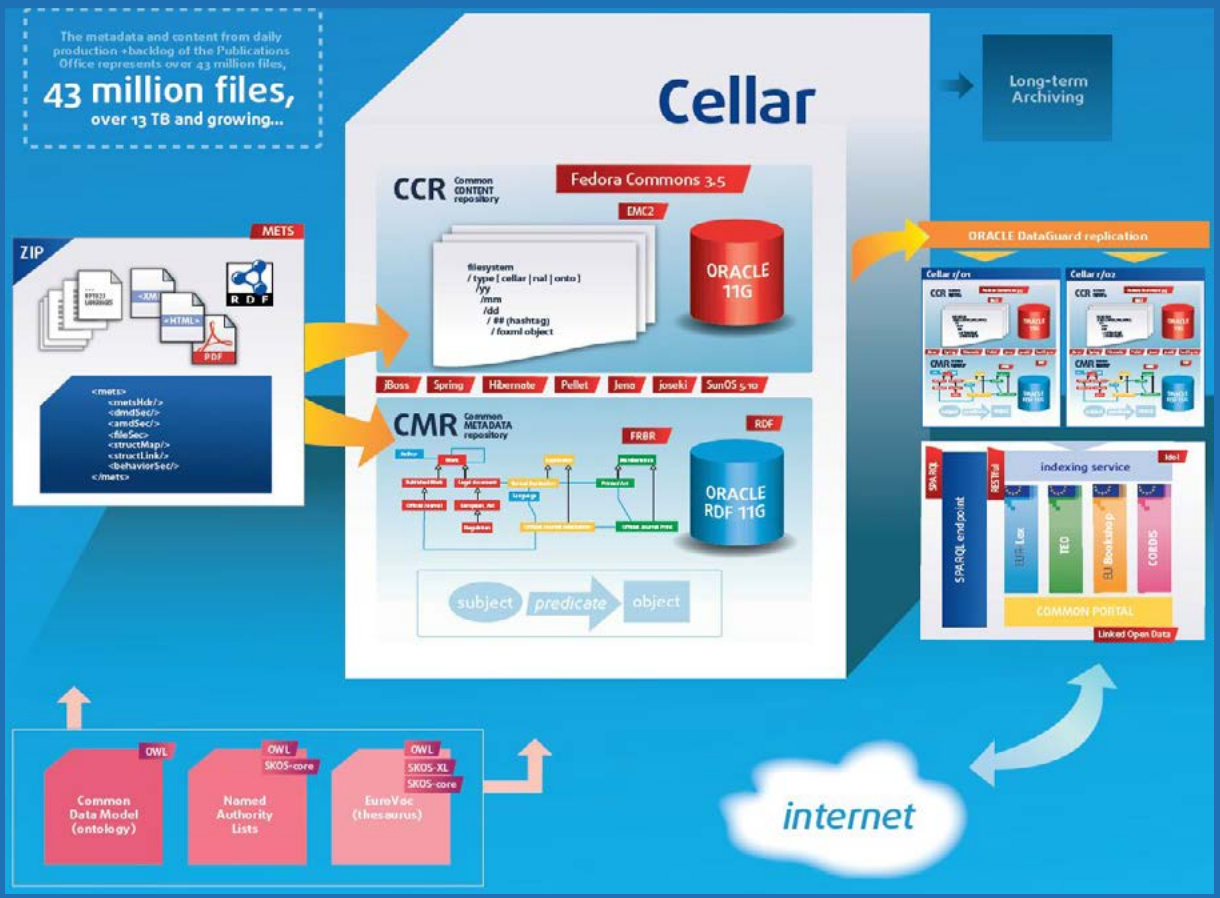
66 For instance, the recent detailed evidence in PE514.071 (2013, October) "Initial appraisal of a European Commission Impact Assessment (SWD (2013) 331, SWD (2013) 332(summary) of a Commission proposal for a Regulation of the European Parliament and of the Council laying down measures concerning the European single market for electronic communications and to achieve a Connected Continent, and amending Directives 2002/20/EC, 2002/21/EC and 2002/22/EC and Regulations (EC)

The Publications Office of the European Union

The Publications Office of the European Union (Publications Office) is an interinstitutional office whose task is to publish the publications of the institutions of the European Union (Decision 2009/496/EC, Euratom). Its core activities include production and dissemination of legal and general publications in a variety of paper and electronic formats, managing a range of websites providing EU citizens, governments and businesses with digital access to official information and data from the EU, including the EU Open Data Portal and EUR-Lex, and ensuring long-term preservation of digital content produced by EU institutions and bodies.

The EU law and publications website offers easy access to EU law. In the future, it will expand its services and give the users a single access point to all the publications, EU law and data managed by the Publications Office. The 'Cellar' is the central document storage for the Publication Office and a valuable source for openlaws.eu.

http://publications.europa.eu/



Furthermore, legislatures increasingly pay close attention to, or delegate detailed legislative research work to, statutory Law Commissions, which investigate areas of the law in need of reform. Added to such legal reform work is that of Parliamentary Commissions of Inquiry, often judge-led into particular incidents. Such inquiries can be enormously valuable but also enormously comprehensive, and were extremely difficult to research in the pre-digital era. However, recent advances in digital evidence submission have resulted in easy searching of for instance the Leveson Inquiry Part I, whose report alone amounted to 2500 pages.[67] Often these inquiries and commissions have their own website to establish their independence, which leads to problems in archiving the material available.

Case law

Case law is an area of legal information publishing which has very significant challenges in different jurisdictions. Recently the 7th Framework Programme EU Cases project has completed a state-of-the-art report in this area.[68] In some countries, case law of at least the constitutional or Supreme Court is published by government, with private publication of High Court or superior court proceedings, and no reporting except in unusual circumstances of lower court proceedings. Other countries have a far more comprehensive approach, with publication of judgements in even employment, immigration and other tribunal cases, and family court disputes. The latter categories involve sensitive personal information which on balance judges and court officials may not wish to see published, though interpretation of the transparency of justice requirement balanced against the protection of sensitive personal data varies substantially, driven in part by different traditions of transparency in publication.

We can therefore identify officially published, unofficially commercially published, and unreported cases. These vary so much by territory that they are dealt with in depth in country case studies for England and Wales, Netherlands and Austria. Historically, such materials were made available privately, with legal historians documenting both trial reports and customary law declarations by judges and monarchs dating to the early mediaeval period, in for instance Flanders, Germany and England.[69] Much legal information also survives from the Roman period, which proved vital to formative pre-modern European conceptions of justice. Or-

67 *Leveson LJ* (2012) An Inquiry Into The Culture, Practices And Ethics Of The Press. The full report is only 2500 pages, the evidence into the report is closer to 100,000 pages or 5,000,000 words.

68 EUCases (2014) Deliverable 1.1: Report on state-of-the-art and user needs at: http://eucases.eu/d1_1/ summarised at http://eucases.eu/fileadmin/EUCases/documents/Presentations/Torino2014/APIS_LOD_State-of-the-art.pdf

69 See generally *Musson, Anthony and Chantal Stebbings Eds.* (2012) Making Legal History: Approaches and Methodologies, Cambridge University Press. Also see *D. Heirbaut* (2001) The Belgian legal tradition: does it exist?, in: *H. Bocken and W. De Bondt eds.* Introduction to Belgian law. *Mechelen* pp 1-22. *Brand, P.* (1996) The Earliest English Law Reports, Voumes I and II, London: Selden Society.

Global Open Data Index - Legislation

The Open Data Legislation Index is published by Open Knowledge International (http://index.okfn.org/dataset/legislation/). UK is ranked #1, together with six other countries. Austria and the Netherlands (participating in openlaws.eu) are ranked on #7, together with seven other countries (because data is not available as bulk download, but "only" via interfaces).

Rank	Place	Breakdown	Location (URL)	Format	Info	Prev. (2014)	Score
1	United Kingdom		http://www.legislation.gov.uk/...	XML, ...		#1 100%	100%
1	Brazil		http://dadosabertos.senado.gov...	metada ...		#33 45%	100%
1	Romania		http://legislatie.just.ro/Publ...	TXT		#33 45%	100%
1	Taiwan		http://law.moj.gov.tw/PublicDa...	XML		#12 75%	100%
1	Finland		http://www.ldf.fi/semantic-fin...	RDF		#19 70%	100%
1	United States		http://www.gpo.gov/fdsys/bulkd...	XML		#1 100%	100%
7	Kenya		http://www.kenyalaw.org:8181/e...	XML, ...		#6 90%	90%
7	Chile		http://datos.bcn.cl/	HTML, ...		#6 90%	90%
7	Qatar		http://www.almeezan.qa/	HTML		n/a	90%
7	Netherlands		https://data.overheid.nl/data/...	Webserv ...		#6 90%	90%
7	Korea, Republic of		http://www.law.go.kr/	XML, ...		#25 55%	90%
7	Spain		https://www.boe.es/legislacion...	XML		#12 75%	90%
7	Italy		http://www.normattiva.it/	XML		#6 90%	90%
7	Austria		https://www.ris.bka.gv.at/defa...	HTML, ...		#6 90%	90%
7	Germany		http://www.gesetze-im-internet...	HTML, ...		#1 100%	90%
16	Japan		http://law.e-gov.go.jp/cgi-bin...	n/a		#12 75%	75%

Table 3: Global Open Data Index: Legislation

ganised records of case law remained fragmentary until the eighteenth century in England, and in nations with less need of precedent-setting court reporting, this was even more the case. Therefore, the modern era of court reporting typically began with commercial publication on behalf of professional lawyers in the capital city, reporting on the higher courts. Customers were both lawyers and judges, but particularly the libraries of legal societies such as the Inner Temple in London. Edmund Coke's "Institutes of the Laws of England" dates to 1628-44 and forms an early example of such works.

An indication of the enormous volume of material which can be made available as judges adopt digital authorship of their judgements is available from Canada. Crown Court relationships were established for the province of Quebec in 1982/83 to make law available publicly.[70] There are now 80,000 Quebec decisions published each year, not counting anonymised family law cases (A v. B etc). As a result,

70 The system is known as Soquij – see http://www.caij.qc.ca/en/library/libraries

BAILII

The British and Irish Legal Information Institute (BAILII) provides access to the most comprehensive set of British and Irish primary legal materials that are available for free and in one place on the internet. In August 2012, BAILII included 90 databases covering seven jurisdictions. The system contains around 36 gigabytes of legal materials and around 297,513 searchable documents. BAILII is legally constituted in the UK as a company limited by guarantee (No 4131252) and as a charitable trust (registered charity no 1084803).

http://www.bailii.org/bailii/

55% of CANLII decisions are from Quebec, which shows the potential for overwhelming numbers of published decisions were all Canadian provinces – or European Union member states – to do so. Quebec has only 15% of the Canadian population, which could mean over 500,000 cases reported each year were all provinces as diligent as Quebec.

Privacy issues arise in particular connection with family law cases, particularly for individuals whose only significant Internet indexing is to a family law dispute which they would prefer to be forgotten. A 2003 decision was made by CANLII to prevent Google scraping the index of such cases, in order that financing of the system could continue. Whether to publish lower court family decisions can be very difficult to decide. The CANLII policy is to publish on a Notice and Take Down basis: if the judge decides in each case that confidentiality is maintained, the case is taken down. Checks are made with clerks of courts, a time consuming but responsible process. Registrars of courts in some provinces can make the decision on privacy grounds to publish only in subscription-based commercial services such as Lexis, rather than with open data repositories such as CANLII.[71] A Romanian citizen/'entrepreneur' has scraped CANLII database, displaying divorce and custody cases amongst others, with the result that family law disputants have paid him via Paypal to remove their Google search result. His website, Globe24h.com, has very strong Google metrics.[72]

Lower US court decisions are now also very widely available, financed by credit rating agencies which have strong interests in ensuring transparent access to data on personal insolvency and debt judgments.[73] The US system is fragmented, partly due to the lack of a common citation system until the late nineteenth century when adopted by the founder of what became Westlaw. Similarly, English court

71 Information courtesy of Openlaws interview with *Colin Lachance* of CANLII, 18 August 2014 Ottawa.
72 *Lachance, Colin* (2014) 26 May: Google, Gonzalez and Globe24h http://www.slaw.ca/2014/05/26/google-gonzalez-and-globe24h/
73 See https://www.pacer.gov/psco/cgi-bin/links.pl

cases relied on competition between private reporters to provide case summaries and judgements until the twentieth century. Even in the twenty-first century, there is fragmentation below the appeal court level.[74]

Note the development of online case law by Legal Information Institutes as part of the Free Access to Law movement (FALM)[75], such as BAILII in the UK. BAILII is particularly well documented over its history[76], including in its foundation[77], establishment, reform and funding difficulties.[78]

74 To take an example, Bailey Solutions Ltd provided the technical solution for Legal Online Research Databases (LORD) – a BIALL project designed to remove fragmentation of English law reports, by pooling metadata for titles. The problem remains unresolved due to lack of resources. Source: Openlaws interview with *Penny Bailey*, 28 May 2014.

75 See *Greenleaf, Graham*, *Mowbray, Andrew and Chung, Philip*, The Meaning of 'Free Access to Legal Information': A Twenty Year Evolution (2012). Law via Internet Conference, 2012. Available at SSRN: http://ssrn.com/abstract=2158868 or http://dx.doi.org/10.2139/ssrn.2158868 See further *Greenleaf, Graham*, Free Access to Legal Information, LIIs, and the Free Access to Law Movement (2011). Iall International Handbook Of Legal Information Management, *R. Danner and J. Winterton, eds*., *Ashgate*, 2011; UNSW Law Research Paper No. 2011-40. Available at SSRN: http://ssrn.com/abstract=1960867

76 *Ury, J.* (2011) "Ten years of BAILII." Internet Newsletter for Lawyers. 2011 May/Jun, 11-12. *Ury, J.* (2004) "British and Irish Legal Information Institute (BAILII) - where we are now." Internet Newsletter for Lawyers. 2004 Sep/Oct, 3-4

77 *Eastham, L.* (1999/00) "Free the law - inspiration and motivation." Computers & Law. 10(5), 3-5 *Leith, P* (2000) "Owning legal information." European Intellectual Property Review. 22(8), 359-365 *Holmes, N.* (2000) "A page on the web." Solicitors Journal. 144(11), 227 *West-Knights, L. J.* (2000) "The law online: we are getting there" Judicial Studies Board Journal. 10, 17-18

78 *Butcher, D.* (2002) "Electronic sources of UK legislation: BIALL & SCOOP Joint Seminar Report." Legal Information Management. 2(3), 48-50 *Furlong, J.* (2003) "Free and easy? The development of Internet access to Irish legal materials." Legal Information Management. 3(2), 95-98 *Brooke, H.* (2004) "BAILII looks for help..." Journal of the Law Society of Scotland. 49(9), 35. *Leith, P.* (2007) "BAILII - towards a national law library?" Legal Information Institute. 7(1), 42-45 *Freedman, C. and Ury, J.* (2008) "BAILII's continuing expansion." Computers & Law. 19(3), 5-6 *Brooke, H.* (2008) "Judgments for all." European Lawyer. 78, 3. *Holmes, N.* (2010) "Free case law - an overview." Internet Newsletter for Lawyers. 2010 Jul/Aug, 1-3 *Holmes, N.* (2011) "Is Free Law Good Enough?" Internet Newsletter for Lawyers. Nov/Dec 2011, 1-2 Miller, J (2011) "Funding BAILII" New Law Journal. 161, 824(2) *Phillips, J.* (2011) "But who will pay?" Journal of Intellectual Property Law & Practice. 6(9), 589. *Harris, J* (2011) "Bailii - minimum cost for maximum benefit." Lawyer. 25(25), 6. *Fellows, C., Leith, P. and Ury, J.* (2012) "Assessing BAILII 2012." Legal Information Management. 12(3), 148-164 *Leith, P. and Fellows, C*. (2013) "BAILII, legal education and open access to law." European Journal of Law and Technology. 4(1)

European Union Open Data Portal (ODP)

The European Union Open Data Portal is the single point of access to a growing range of data from the institutions and other bodies of the European Union (EU). Data are free to use and reuse for commercial or non-commercial purposes.

By providing easy and free access to data, the portal aims to promote their innovative use and unleash their economic potential. It also aims to help foster the transparency and the accountability of the institutions and other bodies of the EU.

The EU Open Data Portal is managed by the Publications Office of the European Union. Implementation of the EU's open data policy is the responsibility of the Directorate-General for Communications Networks, Content and Technology of the European Commission.

http://data.europa.eu/

European Data Portal (EDP)

The difference between the ODP and EDP is that the ODP contains datasets that are collected and published by the European Institutions. EDP is a European portal that harvests metadata from public sector portals throughout Europe. EDP therefore focuses on data made available by European countries. In addition, EDP also harvests metadata from ODP. The European Commission is currently exploring how to bring those two portals closer together.

http://www.europeandataportal.eu

From Big Open Legal Data to Legal Intelligence

The process has already started: Legal data is more and more becoming available as open data, and even though the amounts of such information can hardly be considered 'big data' in terms of volume, velocity and variety, the legal data flood is still impressive. However, data alone is not good enough. Data needs to be converted into information, which then will lead to knowledge and ultimately intelligent systems can evolve. This model is commonly known as the 'DKIW (data - information - knowledge - wisdom) Pyramid' and has been adapted for several purposes and occasions.[79] The term 'wisdom' is often replaced with other names, depending on the context. In the context of this BOLD Vision 2020, the ultimate goal would be intelligent legal information systems.

Legal Data

The data layer includes legal data as described above. Legislation and case law are good examples for data sets which are supposed to be open. For the openlaws.eu project a strong focus was put on these publically available data sets. However, the spectrum of data ranges from open data via shared data to closed data. For the future we see data integrations and data mashups that combine and link private data with open data. For example, references in private documents, such as offical notices by authorities or contracts, could be linked to open data sources. Again, such links would make access to justice easier.

Legal Information

Data is machine-readable, but as a human user we require some kind of representation of this data to consume it as legal information. Data has to be made 'useful'. As of today, legal information is shown almost exclusively as text - and we are used to it. A famous exception are traffic signs, which were a legal innovation as such when they were introduced end of the 17th century in Portugal.[80]

However, textual representations may be supplemented by additional information, such as metadata shown as tags or even visualizations. Additional possibilities to create useful information out of legal data are described further down below.

79 https://en.wikipedia.org/wiki/DIKW_Pyramid
80 https://en.wikipedia.org/wiki/Traffic_sign

Legal Knowledge

The knowledge component of DIKW is generally agreed to be an elusive concept which is difficult to define. Knowledge is typically defined with reference to information.[81] Knowledge requires some for form of structure and organization. In this respect reference can be made to the concept of linked open data. One single legal text may be important piece of information, but usually one can only understand the meaning and 'know' the law in its full compexity, when this text is linked with related information.

Acquiring legal knowledge is a complicated thing. Legal experts study and practice for years in order to be able to give legal advice. The experience capured in their brains can hardly be transferred to a computer system. As of today, ICT systems can only support legal experts, they cannot replace them. The 'end of the lawyers' is not here yet, and they will continue tohelp their clients long beyond 2020. What legal ICT tools can do though is to strengthen the legal knowledge of experts. A good lawyer may know many cases and may be well-informed about the latest developements in his or her field of expertise, but with the increased complexity and volume of legal information, no lawyer can know it all. Often it is said that a good lawyer does not have to know information by heart, but that it is more important to know where to look it up. Again, this shows the importance of the network, the context and the organisation. In other words: While laypersons are able tosearch and find legal information in today's databases, legal experts are able to convert this information into knowledge.

It can be expected that legal ICT systems will be able to help experts to enhance and deepen their knowledge within the next five years. Creating links and saving comments and summaries - in either a public or a private setting - will help experts to create their personal knowledge-base. On the ICT side, it can be expected that legal knowledge graphs and thesauri will be further developed in the upcoming years. However, the creation and modeling of such knowledge graphs still requires the input of legal experts.

Legal Intelligence

Problem solving requires intelligence. While complex legal problems require a legal expert, others may be solved by laypersons and simple problems may already be solved by the machine. For example, IBM Watson is a technology platform that uses natural language processing and machine learning to reveal insights from large amounts of unstructured data.[82] The results in the legal domain may be still

81 *Rowley, Jennifer; Richard Hartley* (2006). Organizing Knowledge: An Introduction to Managing Access to Information. Ashgate Publishing, Ltd. pp. 5–6. ISBN 978-0-7546-4431-6.

82 http://www.ibm.com/smarterplanet/us/en/ibmwatson/what-is-watson.html

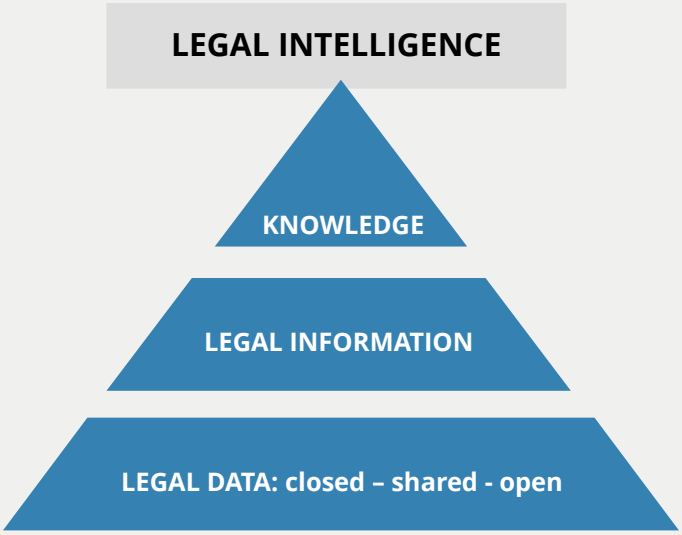


Figure 3: Building Legal Intelligence

far from satisfying, but it shows where ICT research is heading. It is hard to imagine that the machine will be able to apply a given set of legal norms to the facts of a case and to come up with a solution. However, machines are useful when it comes to processing of big amounts of data. A very simple task is to monitor the legal and regulatory framework for changes and amendments. This is rather straight forward and does not require a high degree of intelligence or underlying knowledge. On the other hand, such a task would require many hours when done manually, which would lead to unnecessary high costs.

Over time the level of knowledge and intelligence of the machine will increase. It can be expected that this will already be the case during the upcoming five years, given the efforts of research institutions, commercial publishers and legal tech startups. Again, this does not mean the 'end of the lawyer' but possibly a shift of the activities of the lawyers. Simple tasks can be done by the machine, complex -

Read more about openlaws.eu

- Analysis of legal networks (Deliverable 2.1.d1)
- Requirements for enrichments tools (Deliverable 2.2.d1)
- Final specification and vision of enrichment tools (Deliverable 2.2.d3)

Deliverable are available online via http://ww.openlaws.eu.

I want us to ask ourselves every day, how are we using technology to make a real difference in people's lives.

Barack Obama

Status Quo & Actors

Who is interested in legal information anyway? In the first phase of openlaws.eu, a lot of effort has been put into stakeholder analysis. Several end-user groups and their needs were identified :

- Citizens need easy access to legal information as a first entry point and to find an expert in case further advice is needed.
- Businesses need legal information on a regular basis in order to be able to fulfil their obligations and to exercise their rights. The larger the business, the more jurisdictions and the more legal professionals need to be managed.
- Legal professionals need expert access to premium content, to increase their productivity and to demonstrate their skills to potential clients. In order to stay competitive, lawyers organize themselves more and more in formal or informal legal networks.
- Researchers need access to legislation, case law and premium legal content in order to generate new knowledge. They need to publish and disseminate their findings. Open access publication is required in many publicly funded research projects.
- Public bodies and civil servants have to inform citizens and have to use legislation, case law and commentary in their own daily work.

Depending on the level of legal expertise, users have different needs when it comes to ICT-enabled legal information systems. While legal professionals seek legal expert systems that provide in-depth knowledge, citizens and businesses require easy-to-use systems that provide a quick overview. However, other markets show that the lines between pure expert systems and information tools for the broad public are blurring.

Citizens

There are over 500 million people living in the European Union. They are all affected by thousands of national and European regulations. The vision of openlaws is to provide citizens with better access to justice.

Legal information is complicated and hard to understand. Even with the latest technology, cases cannot be solved automatically and there will be plenty of work for legal experts in the future. The vision of openlaws.eu is though that a few simple tasks can - and in fact should - be automated. Monitoring changes in the regulatory framework are a good example. No citizen would hire a lawyer to inform him/her about the latest changes in legislation that might affect such person. In the next five years, new legal tech services could inform EU citizens about such changes automatically - if they want to be informed. This means platforms like openlaws will not replace a lawyer in case of a legal dispute, but they can provide helpful initial information.

Citizens are a very under-resourced group in accessing law. Where there is great demand for access to social welfare law for non-specialists, Member State governments and the EU provide basic information in a "digested" manner. Consumer protection law may be one example. Austria has a comprehensive online help portal for citizens, containing a lot of legal information.[1] The United Kingdom for example provides basic access to landlord-tenant law. CANS service in the UK has provided wide access to social legislation summaries since 1939.[2] This is intended to help members of the general public to understand the law in areas such as welfare, property and so on.

Citizens are of course the ultimate arbiters of legislation as the electorate. The potential for citizen inspired laws is of great interest. In Hamburg, Germany citizens have created a draft Transparency Act, which was later enacted in a formal legislative process. The users wrote a new law in a wiki-like environment, at first even without the support of a legal expert. In a second stage, a former judge reviewed the draft before it was finally presented to the governmental authorities. The group could engage 'the crowd' so that a sufficient number of supporting signatures could be collected. Finally, the Transparency Act was accepted and enacted by the city of Hamburg. This is one of the first 'co-created' laws, a first signal for what collaboration and participation could do to inform legislative processes.[3]

1 https://www.help.gv.at
2 CANS: 1939 launched alongside Citizens Advice Bureaus. See e.g. bomb shelter advice: http://www.cans.org.uk/librariespublic/archive
3 ttp://de.hamburgertransparenzgesetz.wikia.com/wiki/Transparenzgesetz_selber_machen

Legal expertise (lower to higher)			
	judges lawyers notaries general counsels legal scholars	judiciary legislative authorities large law firms	**BOLD Vision:** **Collaborative Network**
	small enterprises law student semi-professionals	medium enterprises gov. administrations legal publishers	large (international) enterprises
	citizen (individual)	citizens (groups/associations/ etc.)	society (as a whole) Member States
	Individual/institutional size (smaller to larger)		

Table 4: Stakeholder matrix

Businesses

Businesses are 'end-users' of legal information like citizens, but their behaviour and their needs are quite different. These different legal roles are explicitly recognized by the EU and its member states, by differentiating between them in many legal acts and even by enacting dedicated laws (e.g. consumer protection laws or trade acts). Generally speaking, a company is less protected than a citizen, because it is assumed that a company is better educated than a citizen. If a citizen starts a micro business, such person has to comply with legal obligations just as medium or large enterprises.[4]

4 With the exception of certain regulations that only apply once a company has reached a certain size.

Businesses and Lawyers in Europe

Employees	Total	0 - 9	10 - 20	20 - 49	50 - 249	> 250	Lawyers
EU 28	22.098.422	92.7%	4.0%	2.1%	1.0%	0.2%	
United Kingdom	1.703.562	88.9%	6.0%	3.1%	1.5%	0.4%	191,498
Austria	308.411	87.1%	7.1%	3.8%	1.7%	0.3%	5,805
Netherlands	862.697	93.6%	3.0%	1.8%	1.0%	0.2%	17,486
France	2.882.419	94.7%	2.6%	1.7%	0.7%	0.1%	58,224
Italy	3.825.458	94.8%	3.3%	1.2%	0.5%	0.1%	233,852
Spain	2.385.077	94.4%	3.2%	1.6%	0.6%	0.1%	144,159
Germany	2.189.737	82.3%	9.9%	4.8%	2.5%	0.5%	163,690

Table 5: Businesses and lawyers in Europe

Eurostat: http://ec.europa.eu/eurostat/

CCBE: http://www.ccbe.eu/)

Council of Bars and Law Societies of Europe (CCBE)

The Council of Bars and Law Societies of Europe (CCBE) represents the bars and law societies of 32 member countries and 13 further associate and observer countries, and through them more than 1 million European lawyers.

The CCBE was founded in 1960, as the ramifications of the European Economic Community on the legal profession started to be seriously considered. During the decades which followed and through to the present day, the CCBE has been in the forefront of advancing the views of European lawyers and defending the legal principles upon which democracy and the rule of law are based. The CCBE is an international non-profit-making association incorporated in Belgium.

The CCBE is recognized as the voice of the European legal profession by the national bars and law societies on the one hand, and by the EU institutions on the other. It acts as the liaison between the EU and Europe's national bars and law societies. The CCBE has regular institutional contacts with those European Commission officials, and members and staff of the European Parliament, who deal with issues affecting the legal profession.

The CCBE was also developing the Find-A-Lawyer database, which is available on the European e-Justice portal (https://e-justice.europa.eu/)

http://www.ccbe.eu/

Micro & Small Businesses

Outsourcing and self-employment have created an extremely large overlap of micro-businesses (with less than 10 staff, typically only a single individual) and citizens. According to Eurostat, small enterprises and micro enterprises vastly outnumber medium and large companies in the EU. The most usual encounter for micro and small businesses with the law is in the filing of annual tax returns and sales tax returns, and incorporation of the business. Thus, their default professional advisor is often an accountant, though the amount of self-filing of income tax also suggests the online advice given by the taxation authorities is a more common encounter with legal information. They are as likely to contract with a lawyer or notary when letting, buying and selling property, a making a will, getting divorced, as in the course of their professional self-employment.

Dealing with legal questions and risks is also important for micro and small enterprises even if it is cumbersome. A lawsuit or violation of public law (e.g. environmental or safety regulation) can threaten the whole business.

The easiest and fastest way to access legal information for them is via the Internet (as also shown by the openlaws.eu survey). Depending on the EU member state, there are free governmental platforms. In addition businesses will find information in law blogs, in wikis and in different forums. However, they will face three issues with such a search.

- First, is the information complete and up-to-date?
- Second, is it quality controlled?
- Third: Does the information fit the specific situation the company encounters?

If the business person decides that the legal question cannot be answered sufficiently internally, they have two options: they can consult a legal professional or not. However, acting with due care may require that the director contact an external expert, if he/she does not want to risk personal liability. What is actually done (explicitly or in a more tacit manner) is a kind of legal risk assessment. In this particular case, is it worth hiring an expert (with an extremely high hourly rate) or do we simply take the risk? If the risk is considered higher than the legal expenditures that will occur, the expert/intermediary/gatekeeper will be consulted. The company will pay for the advice and receive legal information that is 'catered' to the specific issue.

Outlook for Businesses

With more automated and personalized legal tech services, business will be able to handle the legal and regulatory framework much easier and at affordable prices.

However, such services will not replace the professional advice of internal or external legal professionals, since the machine is still not capable to solve complex legal problems.

Medium and Large Corporations, Corporate Counsels

With the growth of a business, the complexity of legal information increases. First, medium and large corporations have to deal with large amounts of legislation and case law because their activities often become broader. On the other hand, multiple jurisdictions become an issue, because of the geographic expansion of the business. This is in particular a problem in the European Union. Also, 24 official languages in the EU make it harder to deal with the information flood. Usually a company will consult multiple legal experts in each jurisdiction where the company is active. Innovative law firms have recognized this problem and build networks themselves, so that they can offer legal services across Europe as a one-stop-shop solution.

Dealing with legal information in medium and large corporations is not limited to the legal department though. Directors and officers have to ensure that legal information is spread throughout the company, on a "need-to-know" basis. Otherwise they enter unnecessary financial risk for the company and might even risk personal liability if they do not fulfil their organizational duties. For example, the human resources department has to know everything about labour law, the IT development department has to know about privacy, sales and marketing has to know about competition law and so forth.

Legal compliance and risk management are therefore essential to corporate counsels, which makes sharing of information on diverse subjects important. These include labour law, corporate governance, but also competition law, tax law and many other aspects of regulation. This means corporate counsel need broader, and arguably shallower, knowledge than experts in law firms, though with significantly advanced knowledge within specific sectors of industry or subject domains.

Experts

With their high level of legal expertise, legal professionals are gatekeepers to the law. This expert knowledge is gained through a special legal education in combination with practical experience. As described in the first part of this BOLD Vision, machines will not be able to solve complex cases in the future. Still, the quality

Gartner Legal IT Scenario 2020

Gartner has published a legal IT scenario for the year 2020 in 2014 (http://blogs.gartner.com/french_caldwell/2014/02/28/gartner-legal-it-scenario-2020-smart-machines-and-lpo-radically-disrupt-legal-profession/). According to the report, many lega tech disruptions are already on the way, such as the increasing demand for legal process outsourcing (LPO) and the use of advanced analytics. Gartner's predictions:

- By 2020, 75% of U.S. and U.K. corporations will use LPO.
- By 2019, 75% of corporate legal and IT departments will have shared staff.
- By 2018, legal IT courses will be required for the graduates of at least 20 U.S. Tier 1 and Tier 2 law schools.

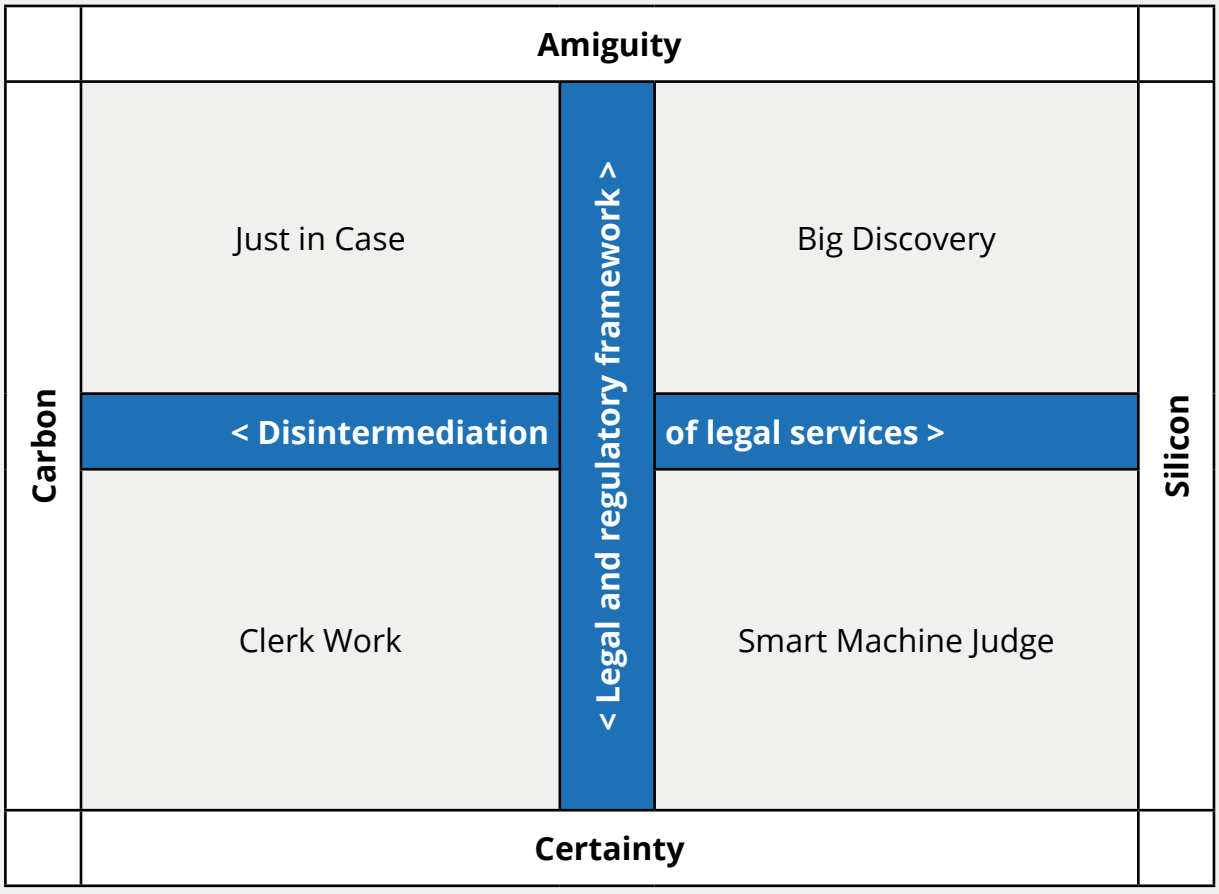


Figure 4: Legal IT scenario 2020 (Gartner)

of legal databases and information systems will increase over time, which gives the expert the opportunity to focus on more complex problem solving. Gartner has published a Legal IT Scenario 2020, where legal process outsourcing (LPO) is described, from traditional "Just-in-Case" work to IT supported "Smart Machine Judges" and "Big Discovery".

Lawyers: Small Law Firms and Sole Practitioners

According to Eurostat there are approximately 500 million legal enterprises within the European Union.[5] Most of these are solo practitioners or small law firms with only a few employees. They are either working as generalists who have a broad overview on various legal areas or they are very specialized in a particular field. Over the past years and decades they have are built and maintained – usually very good – relations to their clients, i.e. citizens and/or businesses. Hourly rates of legal experts are high, the workload is high as well, leading to constant revenue streams. In our field research, we have not found one single lawyer, who complained that there was not enough work. Usually the problem is that they do not have enough time and that their clients want "more for less", leading to declining margins, a trend that can be observed in other industries as well. Still, many of these small law firms and sole practitioners believe that their consulting business will continue in the same way as it has always been.

Critics argue that small law firms and sole practitioners do not have a bright future, in particular in liberalized regimes that do not protect the legal industry anymore.[6] Small legal enterprises cannot use economies of scale like larger legal enterprises. For example, purchasing access to commercial legal databases is relatively more expensive for a small legal enterprise than for a large one. Other examples are internal IT systems, advanced CRM systems, document management systems, collaboration and communication systems, all of which will be more and more expected by the 'end user'. Parallels may be drawn to the accountancy business, where there is a concentration of the 'Big Four' players.

However, there are many IT solutions available, that may be used by small legal players often at very affordable prices. Cloud computing theoretically enables sole practitioners to use economies of scale, namely the benefit from the community sharing one central IT system online. The providers of these solutions do not necessarily emerge from legal IT providers. Legal enterprises may use these solutions like any other small business. Furthermore, with the commoditization of online information services, access to legal information may also become easier. Reference is made to the providers of free legal information, like

5 Eurostat, Annual detailed enterprise statistics for services (NACE Rev. 2 H-N and S95) [sbs_na_1a_se_r2], M691, http://epp.eurostat.ec.europa.eu

6 See variously *Susskind, R.* (OUP, 1987) Expert Systems in Law, *Susskind, R.* (OUP, 1996) The Future of Law; Susskind, R. (OUP, 2000) Transforming the Law, Susskind, R. (OUP, 2008) The End of Lawyers? Rethinking the Nature of Legal Services, *Susskind, R.* (2013) supra n.37, all Oxford University Press.

Legal Information Institutes (LIIs), the Free Access to Law Movement, the more and more powerful governmental legal databases, and free resources from large stakeholders like Google (see Google Scholar for Case Law).

A few of these small law firms team-up with other small law firms in other countries of the EU in order to form networks or alliances. So maybe in the future we will see more networks in terms of collaboration, rather than only a concentration of market player in terms of traditional mergers and acquisitions. Legal information providers will certainly adapt their solutions to the needs of this user group. Note the declining margins and commoditization of small law firm work, and that some repetitive work that is commoditized needs little research – for instance conveyancing, wills and probate.

Lawyers: Large Law Firms

Large law firms enjoy economies of scale and therefore they may dominate the future of commercial legal information services. Big law firms often cover all legal fields, so the 'end user' will always find an answer there. The people working in a large law firm are a closed community and they may not even know each other colleague in the law firm anymore. Certain large law firms employ more experts than a whole country. In particular openlaws field research revealed that some US based law firms have more experts than there are lawyers in Austria (i.e. approx. 5,800). Companies in these dimensions can afford an IT department and many productivity tools. Furthermore, it is also easier and more attractive for commercial legal information providers to address one large law firm at once, rather than several hundred sole practitioners individually. Accordingly, such providers may adapt their IT solutions towards the needs of these big players.

The sheer size of large firms creates a greater need for information sharing to prevent isolation of individuals and teams within the firm. Directories are the solution to at least find a colleague. The knowledge within

Outlook for Lawyers

There is a place for sole practitioners and small law firms in the future of society of networks. Smaller may be more beautiful, more flexible, more agile, more on demand. However, larger firms remain likely to be the core market for commercial legal information. Juniors/associates spend a large percentage of time researching for senior colleagues. Large firms are also the largest producer of material for analysis – journal/blog authors. As margins are being squeezed, specialised offerings and tailor made, value-added services may continue to expand as a result. Internationalisation is spreading rapidly and contributing to this differentiation trend.

Green and Gold Open Access

Open access can be provided in different ways, with the two most common methods usually categorized as either gold or green open access.

Gold open access refers to publication in an open access journal. The journal may or may not charge a fee to the author.

Green open access refers to self-archiving in open repositories such as for example OpenAIRE, which is large-scale EU initiative that aims to promote open scholarship.

https://www.openaire.eu

such a huge organization is extremely high on an aggregated level, but the question is how this knowledge can and should be managed, shared and maintained.

Large law firms often have offices across the EU, sometimes even worldwide. Here again, the difference to small law firms collaborating in networks may not be too big (apart from the finance and accounting aspects in the background; administrative issues that do not increase legal competence). Marketing of large law firms can be quite different compared to small law firms. While the big ones may have a well-known name and even registered trademarks and marketing budgets, smaller law firms will rather prosper on the personal reputation of one single expert and his/her achievements in the past. In his respect, large law firms are more anonymous. This aspect is important, because in times of social networks and transparency, the competence of an individual is becoming again increasingly important (e.g. "endorse functionality" in LinkedIn or other networks). So already today in our network society, the advantages of a large law firm may not be so enormous as it might have looked in the past.

Legal Scholars

Academics need access to law through a combination of free, open and commercial databases. Academic research interests will vary from the highly resource-intensive (original research) to the more mundane (textbooks, legislation and case reports) to the relatively trivial (single queries from compendia).

One major change in the upcoming years may be open access publication. This approach is still very uncommon in the legal field, where publications in traditional and well-established journals add to the reputation more than pub-

lications in open access journals. This is despite predictions by Hibbitts almost two decades ago that the traditional law journal may be dying.[7] Note that 37 US law reviews have signed up to open access principles, but only a few European journals.[8]

European academic commentary is opening access in very much an incremental fashion. By contrast, US law journal subscription is falling rapidly (7580% in the forty years from 1972) due to free online readership, but consumption is growing extremely rapidly – a victory for an open access to law model in the cash-rich US law school review publishing system.[9]

Public Sector

There are many legal experts employed in the public sector. Combined, they have an enormous legal knowledge. Unfortunately, this knowledge cannot be shared adequately today. Judges in general have access to law libraries and free and commercial databases but their sharing of legal information relies on proprietary solutions in many jurisdictions, to increase perceived security and confidentiality of the legal decision-making process.

Other categories of public servants who require access to law, largely without localised access to law libraries, are civil servants, policemen, prison and probation officials, immigration and enforcement officials, local council officers, etc. While all can access information on legal changes via the national and specialist press,

Outlook for Legal Scholars

More legal commentary will be published under an Open Access regime by 2020. However, open access will remain a niche in the upcoming five years, compared to the established publication channels, i.e. commercial publishers.

Collaboration and networking will become easier though, just as more and better free legal information systems (based on open data) will make work faster and more productive.

"Legal data scientists" may arise and produce new insights based on aggregated Big Open Legal Data, which can lead to a new type of legal papers and publications.

7 See *Hibbitts, Bernard J.* (1996) 'Last Writes? Reassessing the Law Review in the Age of Cyberspace' 71 NYU L Rev 615. *Rier, David A.* (1996) 'The Future of Legal Scholarship and Scholarly Communication: Publication in the Age of Cyberspace' 30 Akron L Rev. 183, 188210; *Bruce, Thomas R.* (1996) 'Swift, Modest Proposals, Babies, and Bathwater: Are Hibbitts's Writes Right?' 30 Akron Law Rev. 243, 243. *Pearson, Shawn G* (1997) 'Comment, Hype or Hypertext? A Plan for the Law Review to Move into the TwentyFirst Century' 1997 Utah L Rev 765, 798.

8 See https://wiki.creativecommons.org/Open_Access_Law_Adopting_Journals

9 See *Davies, Ross E.* (August 24, 2013) The Increasingly Lengthy Long Run of the Law Reviews: Law Review-Business 2012 – Circulation and Production Journal of Law, Vol. 3, No. 2 (Journal of Legal Metrics, Vol. 2), pp. 245271, at p258; *George Mason* Law & Economics Research Paper No. 1351. Available at SSRN: http://ssrn.com/abstract=2315382

Outlook for Public Bodies

Public bodies have the power to make large amounts of data accessible to the public for re-use. The pressure to increase economic growth and to reduce public spendings will lead to more innovative approaches, which include open innovation and open data.

trades associations, briefings from government and using local libraries, their knowledge of law must to some extent depend on use of generic legal information online.

Member States are also in charge of the implementation of the PSI Directive, as described further above. It can be expected that more data sets will be published in the next five years. Such data may not only include legislation and case law, but also other data that can be used for combination with legal data (for example geographic data, data from public registers, financial data for governments, etc.).

Open data is part of open government. Increasing information and knowledge exchange, enhanced connectivity, openness and transparency provide new opportunities for public administrations to become more efficient and effective, provide user-friendly services, while reducing costs and administrative burden. According to the European Commission open data and open government have an enormous potential to save costs while providing better services:[10]

> *The availability of open data can facilitate the creation of new services, stimulate new markets, businesses and jobs, by adding value to the original data provided by government. The full use of big data in Europe's 23 largest governments can reduce administrative costs by 15% to 20%. Open and modular public services can be re-used by different administrations, but also by businesses and citizens, in order to create and deliver personalised, user-friendly and innovative services.*

With respect to law, the Commission points at the example of Iceland. The government used social networking sites to crowd-source provisions to their new constitution.

10 Digital Single Market, Open government, https://ec.europa.eu/digital-single-market/en/open-government; DG Connect, A Vision for Public Services, https://ec.europa.eu/digital-single-market/en/news/vision-public-services

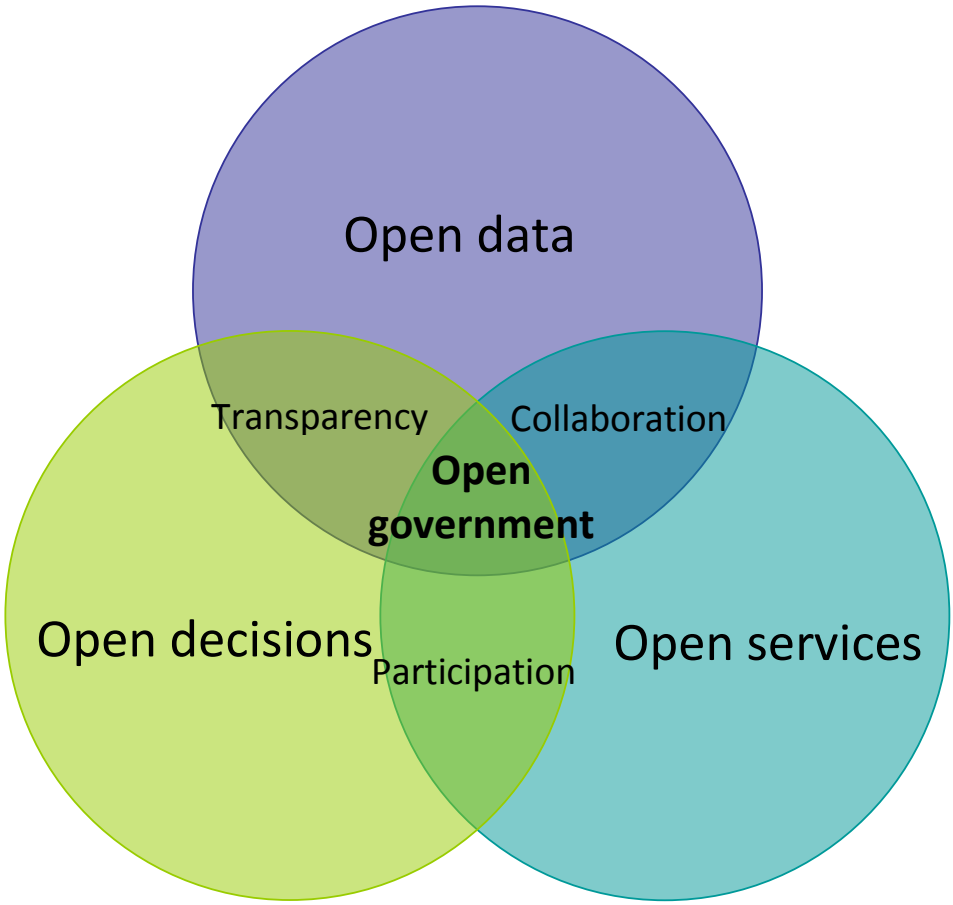


Figure 5: European Commission: A Vision For Public Services

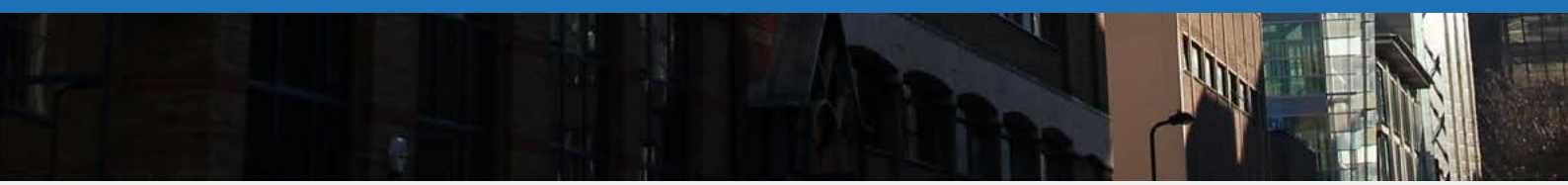
Read more about openlaws.eu

- State-of-the-art report for legal, social and business aspects of re-use of legal information (Deliverable 1.1.d1)
- Case study 1: European institutions (Deliverable 1.2.d2)
- Case study 2: UK (Deliverable 1.2.d3)
- Case study 3: Austria (Deliverable 1.2.d4)
- Case study 4: Netherlands (Deliverable 1.2.d5)
- Synthesis report (Deliverable 1.3.d1)
- Comparative country report: White paper on the OpenLaws.eu open innovation community (Deliverable 1.3.d2)

Deliverable are available online via http://ww.openlaws.eu.

Prediction is very difficult, especially if it's about the future.

Niels Bohr



General Trends

General trends help us to predict the future. Demographic changes for example are relatively easy to measure and monitor and therefore lead to good and stable predictions. Trends in innovation and ICT development are much harder to predict. The legal domain is adopting new solutions rather late, and there are good reasons for that.[1] But because of these late adoptions, predictions become easier to a certain extent. Innovations that have proven to be successful and useful in other domains, are likely to be adopted in the legal domain as well.

Adopting new technology earlier or later is not a question of right or wrong. There are sound reasons why the legal profession is very careful before jumping on the bandwagon. Legal information is important and mistakes could lead to severe consequences. Take for example privacy. Clients expect the highest trust level when consulting a legal expert.

With more and more digital natives coming out of universities, ICT-enabled services are likely to become more common in the legal area. This can already be observed in the USA where legal tech companies are highly active. Looking at Silicon Valley, a full range of legal tech start-ups can be observed.[2] Venture capitalists have already successfully financed a few of these legal tech companies.

In comparison to traditional law firms, businesses are more open to innovative solutions, especially if they help them run their core business better or if they reduce the risks or the costs for the company. With an abundant number of new Internet services, citizens are also used to explore innovative solutions, especially if they are provided free of charge and if they help them make their lives easier.

A few general trends can be observed in the area of ICT. All of these are visible in legal information technologies. Bearing the progress of such trends in other markets in mind, it becomes clearer where the road of legal tech will lead us, even in a short period like the next five years.

1 *Susskind, Richard*, The End of Lawyers? Rethinking the Nature of Legal Services, Oxford University Press (2010). This is also in line with the OpenLaws.eu field research, interviews and focus groups. Even lawyers consider themselves often as laggards with respect to new technology

2 The openlaws.eu team attended the Stanford CodeX Future Law conference in April 2015 and conducted inter-views with several start-ups in the San Francisco Bay Area.

Digitalization and Mobility

As everywhere, digitalization is a so-called megatrend. Legal databases replace traditional libraries, e-mails have replaced most letters (with the exception of "official" letters from/to authorities). Social networks are everywhere, and even if there are only few dedicated social networks in the legal market, people are used to creating connections with each other (e.g. via LinkedIn or Facebook). Smartphones and tablet PCs make it easier for everybody to access ICT services anytime and anywhere.

Mass Customization and Personalization

Personalization is a trend that is used almost in every online-based service in order to provide services in accordance with the personal profile and the personal needs of users (e.g. a personal dashboard for the user after log-in). Customized services help users to save time and to increase their productivity. Mass customization means that a service is designed in way so that it can be easily adjusted to the individual requirements of a person. Even if the underlying platform is thesame, the service will be slightly different for each user, depending on the user's adjustments. A lawyer working in the area of privacy has certainly different interests compared to a civil servant working on environmental law.

Combining big data and open data with personal data, enables us to come up with new legal services. The large amounts of legislation and case law can suddenly be organized and structured in line with a user's individual preferences.

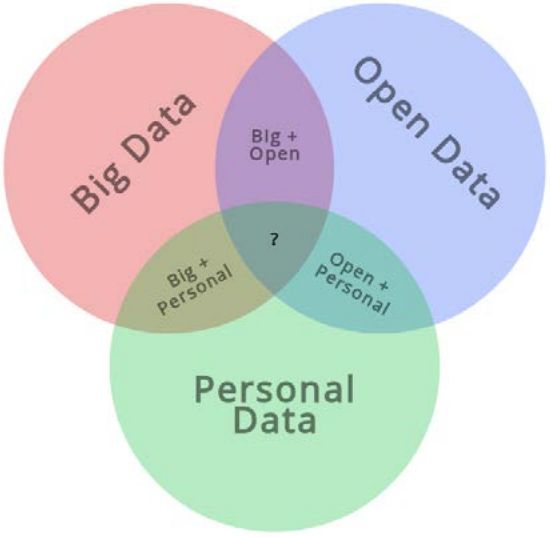


Figure 6: Big - open - personal (ODI)

Legal Tech Fields of Activity

Legislation databases	Databases to access legislation, typically operated by governments
Case law databases	Databases to access case law, typically operated by governments and/or publishers
Literature databases	Databases to access literature, typically operated by publishers
Registers	Registers to find information on companies, patents, land, etc.
Case management	Services to manage different cases
Contract databases	Services to manage contracts
Electronic filing tools	Services for electronic filing of documents, e.g. in at court or at public administrations
Communication tools	Services for communication between different parties, in particular legal professionals
Dispute resolution tools	Services that can be used to solve conflicts between two parties without having to go to court
Document assembly tools	Services to create new documents, e.g. in legislative processes
eDiscovery tools	Services that help to collect relevant information, in particular in cases of litigation or governmental investigation

Table 6: Legal tech fields of activity

Big Data Analyses

Big Data has been described already in detail above. Other domains are already relying on big data to gain new insights. The European Commission considers both open data and big data as important trends. Accordingly they are part of the 'Digital Agenda for Europe'.[3] Expectations of the European Commission are high. They see the potential that big data and improved analytics will make it possible to:

3 See for example: Making Big Data work for Europe, https://ec.europa.eu/digital-single-market/en/making-big-data-work-europe

- transform Europe's service industries by generating a wide range of innovative information products and services;
- increase the productivity of all sectors of the economy through improved business intelligence;
- more adequately address many of the challenges that face our societies;
- improve research and speed up innovation;
- achieve cost reductions through more personalised services
- increase efficiency in the public sector.

More for Less

Another trend is to receive 'more-for-less'.[4] Citizens and businesses alike are used to receive more performance for a much smaller price. This trend is of course a problem for industries with only limited economies of scale or economies of scope. The more-for-less challenge is also an issue in the legal market, when clients expect more legal advice for less money.

Automated services and the support of ICT finally help lawyers to provide more insight for less money. A few years ago it would not have been possible to analyze the data of a whole organisation. Today, e-discovery is part of many company acquisitions during a legal due diligence. If lawyers or paralegal had to read each document, they would have to invest a lot of time and effort.

It can be assumed that all stakeholders will receive "more for less" in the upcoming years. Legal experts will receive legal ICT tools with higher performance. Parts of this new value will be passed on the experts' clients, i.e. citizens and businesses. The entry barrier to the law for citizens and businesses will also be lowered as a result of better search mechanisms and more information online. Legal experts may argue with good reason, that the knowledge collected via the Internet may even be dangerous and misleading. Again, there are other domains, where these developments can already be observed. One example is the medical field, where many patients challenge their doctor's opinion with information found online. The legal domain may learn from these experiences.

Open Innovation

Open innovation is a paradigm that assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as the firms look to advance their technology.[5] On the other hand, knowledge that is generated within the company should be licensed out in order to enable third parties to build additional services, which might address the customer needs even

4 See for example *Matthias Horx*, http://www.zukunftsinstitut.de (26.6.2015).
5 *Chesbrough, Henry William* (1 March 2003). Open Innovation: The new imperative for creating and profiting from technology. Boston: Harvard Business School Press.

better or in another segment. The concept has become more and more popular in recent years in order to create solutions that are close to the users. At a European level, this trend is for example being analysed by the Open Innovation Strategy Policy Group (OISPG) of the European Commission.[6]

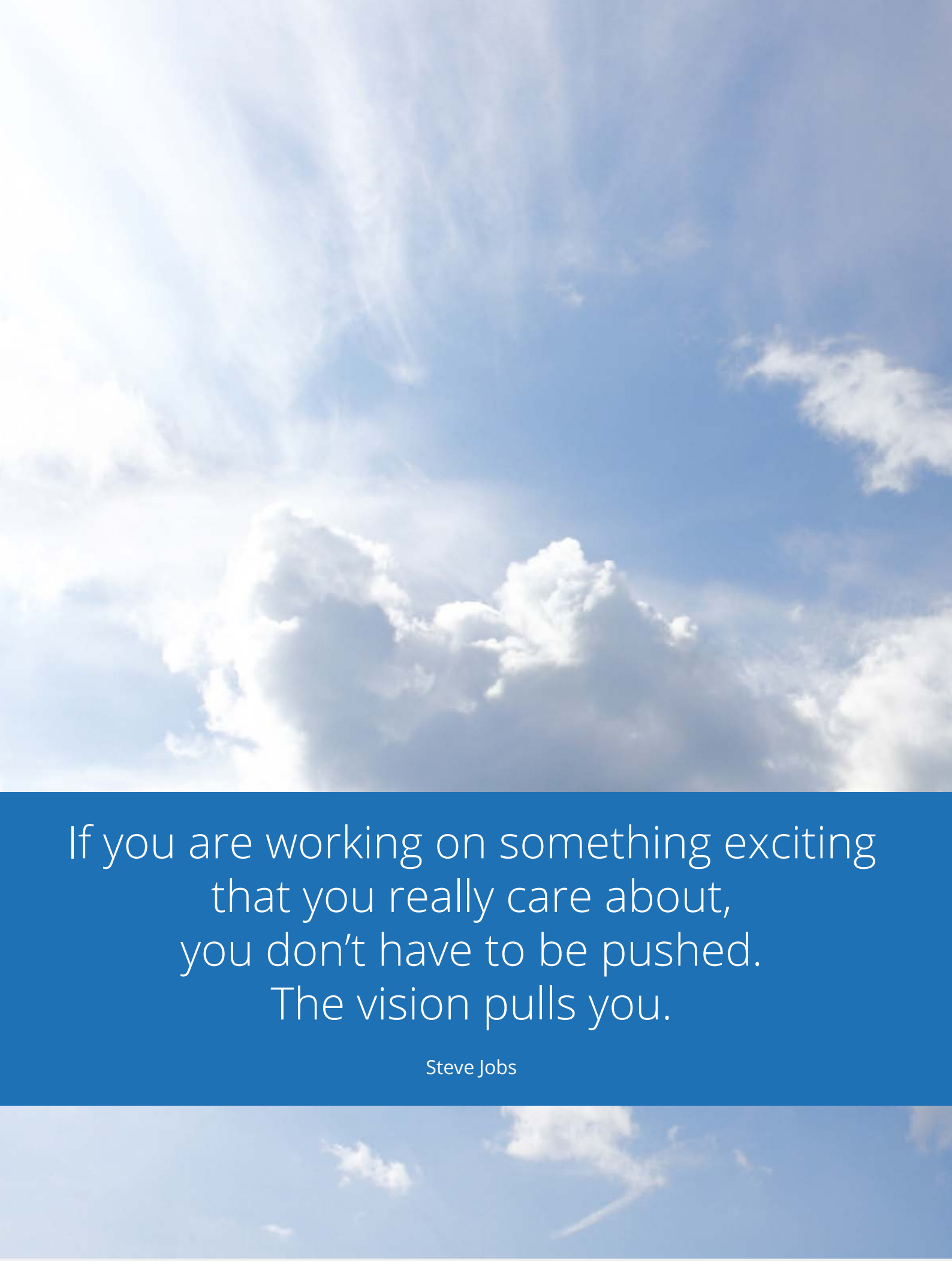
Open innovation has the potential to change the future of the law dramatically. Open APIs and community input can lead to aggregated legal databases. Different datasets can then be re-combined, personalized, enhanced with personal data, etc. Take this raw material to implement new ideas, and a broad variety of new legal services will start to flourish.

Read more about openlaws.eu

- Analysis of legal networks (Deliverable 2.1.d1)
- Requirements for enrichments tools (Deliverable 2.2.d1)
- Final specification and vision of enrichment tools (Deliverable 2.2.d3)
- Final BOLD business models for stakeholders (Deliverable 2.4.d2)

Deliverable are available online via http://ww.openlaws.eu.

6 OISPG, http://ec.europa.eu/digital-agenda/en/open-innovation-strategy-and-policy-group/ (26.6.2015).



If you are working on something exciting
that you really care about,
you don’t have to be pushed.
The vision pulls you.

Steve Jobs

Opportunities

The vision of openlaws is to provide access to legal information and legal experts in the most user-friendly way possible, both on a national and on a European level.

Legal information means primary sources of the law, i.e. legislation and case law, as well as secondary sources, i.e. literature and commentary. These two categories of sources will be complemented with a social layer, i.e. the legal community in its broadest sense.

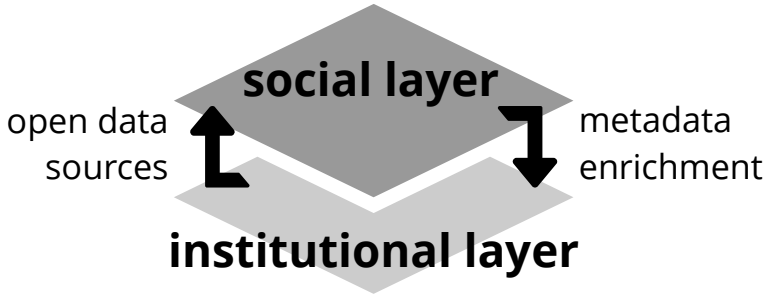


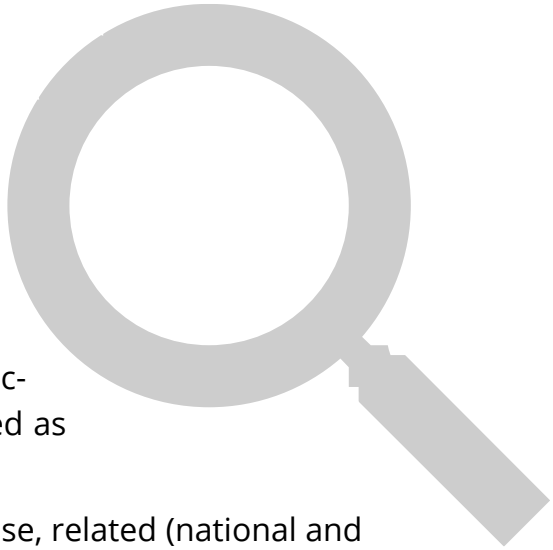
Figure 7: Social and institutional layer

Citizens will learn more about their rights. For companies it will be easier to act in accordance with the law. Legal experts can demonstrate their knowledge and intensify the customer-relationship. Publishers can integrate their premium legal content via open interfaces. Governments can support openlaws and the creation of a community-driven legal knowledge base by making even more data sources available as open data.

For this purpose, openlaws is building a cloud-based online service, where users can search for legal information, organize it the way they need it (e.g. by creating personalized folders, by adding tags, highlights and comments, etc.), and share it with others, either entirely openly or in private groups.

The following pages describe a few legal tech features that will make the lives easier for legal experts, businesses and even citizens. Some features have already been successfully implemented as a prototype during the openlaws.eu project. Some have already been published under the working www.openlaws.com platform. Some may be implemented by the openlaws team in the future, some by innovative legal tech start-ups. Some may be standalone. Some may be built on top of the aggregated legal information of openlaws in accordance with open innovation principles (“do not re-invent the wheel”).

Search



Including a search functionality is essential in every information system. The main difference compared to other legal databases is the meta-search approach of openlaws. Legislation, case law and literature databases can be searched in one simple step. In the future, directories of legal experts and other datasets can be included as well.

This means a search for a certain case will provide the case, related (national and potentially EU) legislation, related literature where the case has been discussed and also legal professionals who are specialists in the subject matter.

openlaws is open. It will be possible for governments, publishers or open data software developers to connect additional sources to openlaws, making the search results even more comprehensive. This means openlaws will not be restricted by the contents of only one data provider, but combine the results and redirect to the original source.

While legal professionals are typically using (expert) search functionalities to find what they need ('search-field approach'), citizens and businesses will need a more structured approach to be able to browse through legal topics ('directory approach'). In addition, recommendation systems will ensure that the users find related legal information.

The ambitious but realisic target of openlaws is, that each year two to three additional EU Member States and their legislation and case law databases can be integrated. By 2020 15 EU Member States could be fully integrated into openlaws. This means that the contents can not only be searched, but that users - and Member States who decide to publish legislation and case law in line with commonly used standards - will benefit from the features as listed below.

Personalisation

openlaws is centred around personal legal assistance services. Registration is not required, the basic openlaws features like search and browsing can be used free of charge and without submitting personal information. However, customization is only possible if users create an account. Users can adjust basic settings like language or preferred jurisdictions, databases and filters. This means every time the users comes back and log-in, the system will be adjusted.

Personal user accounts will also make it possible to claim legal expert status. Legal experts will benefit from a few special options catered for their needs. For instance, experts will be enabled to publish legal commentary. The qualification as a legal expert is necessary to keep the level of published content high. With this approach will be a counter-balance to the common criticism on Wikipedia-like platforms and other forums.

Bookmarks

Users do not want to search for the same content over and over again. They want to organize information according to their personal needs. Bookmarks are nothing new and they are an early feature of personalization. They exist since the early days of the Internet and every browser supports them. So how can bookmarks be special? In combination with structured and linked data as described before, bookmarks can become active. This means that the openlaws platform will be able to inform the registered user automatically whenever there are changes to the bookmarked object. This could be an amendment of legislation or even a new case, which refers to bookmarked data.

Bookmarks can not only be put on existing legislation or case law, but can also be combined with search features. So every time a new case appears that features a bookmarked keyword (e.g. privacy), the user can be notified. Such intelligent search bookmarks exist of course also since quite a while. Good examples are Google alters or large databases from commercial legal publishers.

Portfolios

Closely related to bookmarks are personal folders or entire portfolios. The bookmarks are simply stored in folder structures. Again, this approach is not ground-breaking per se, but in combination with linked data and data analytics this feature becomes interesting and has a few indirect benefits. First of all, the organisation in a folder leads to a certain personal cluster of legislation and case law. If clusters of the same legislation and case law appear several times, it is likely that those users did not accidentially combine such legal information, but that there is probably some kind of relation between those objects. This can be used for recommendations and collaborative filtering (see below).

For example, users can collect everything about copyright law in one folder. Or they can create a folder that includes all the cases for the next exam. Or create a list with all legislation that is relevant in a specifc domain.

Highlingting

Legal experts love to highlight text. This is an easy and simple way to structure information and to spot central text elements swiftly. As a result many legal text books are rather colourful. The visual markings help users to jump to the important positions. Legal publishers offer this feature in apps and also a few US legal tech start-ups have introduce highlighting in their solutions.

Again, highlighting is a useful productivity tool and a personalization feature. Given the popularity of this feature in the "real" world, we believe that this feature can and should be more broadly used in legal ICT systems. Member states can for example swiftly introduce highlighting by connecting their systems to openlaws. Or at least by making legislation and case law available as open data, which means that somebody else could establish the connection.

Tags

Tags are commonly used to describe content and to find information more easily. Legal information is usually represented in text, which means it is searchable. However, in order to be able to find a relevant document, the user would have to look for exactly the right expression. One solution to overcome this problem are thesauri, knowledge graphs, synonyms and machine-generated tags. The other more traditional approach is to provide keywords. This is already happening today and public bodies are adding keywords to legislation and case law. So do commercial publishers. The problem with this approach is, that resources of those central database operators are limited - even if it is a large publisher with a large editorial team. It is unlikely that he internal staff of those organisations will think about all the applicable keywords.

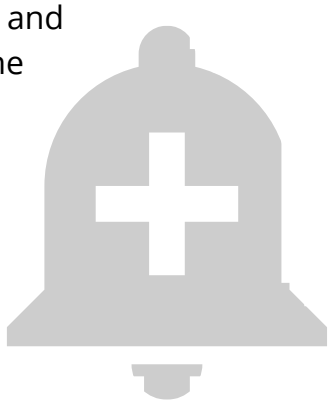
One possible solution to this is open innovation. Users can be offered the option to add tags themselves. Just like the features before, this is already happening in other areas. A most prominent service where users add tags is Twitter. "Hashtags" describe the information and help others to find the content easily. A problem for all community-generated information is quality and the same applies to tags. As a counterbalance, other users can be enabled to rate or vote the relevancy of tags. Ideally, the more relevant tags would appear on top, and tags, which don't apply, would disappear. LinkedIn is following this approach for its endorsement functionality, where users can say whether a person has knowledge in a specific area.

Tags which are created by the community could also be re-used and integrated by public bodies or commercial publishers - if necessary after a quality check. Still, the effort of accepting or rejecting proposed keywords is lower than adding them without any guidance or suggestion. Like all open innovation mechanisms, the community based creation of tags is related to Linus' Law: Given enough eyes, all bugs are shallow.

Alerts

Being actively informed about legal and regulatory changes can save a lot of time and effort for legal experts, businesses and public administrations. As of today, many experts read news, journals, e-mail newsletters, the official gazettes etc. in order to keep track of new legislation and new landmark cases. This process is also described in the ISO 19600 guideline about compliance management systems.

Receiving customized e-mail alerts which inform users about such changes would be a time-saving benefit for many people. Based on personal settings and bookmarks as described before, users would stay well-informed and the risk of missing relevant updates is reduced. Most modern ICT systems offer some kind of alerting mechanism, why shouldn't this be broadly used in the area of law? Commercial publishers already offer alerts, but there is still a high potential for moving from passive to active information.



The essential basis is of course structured data. Only when there is a new amendment of an act, users are interested. Whenever there is a new case in the field of interest of an expert, the system should provide a call-to-action. This requires aggregated (ideally open) data, structured and linked data, as well as some basic personalization for each user.

Comments

Legal information is often not self-explanatory. Notes and comments can help to understand legal texts. This is true for laypersons, but also for legal experts. User-generated content becomes more and more valuable and essential in online platforms. Again, public commenting is a feature which is offered by innovative legal tech companies in the USA already.

Quality of public information is an important aspects for comments. In every community-driven collaboration system there have to be certain mechanisms to avoid misuse or misconduct. Wikipedia for example has a systems of checks and balances in order to approve or reject edits. Still, the quality of Wikipedia is often discussed, especially when it comes to complex areas with relatively few contributing users - such as the legal domain. The approach of openlaws is therefore that only

legal experts can make available information for the general public. Standard users can still make comments, but only for themselves. In addition, legal experts can only publish under their real name. This has the effect, that the motivation to provide high quality information increases. One the other hand, experts can use comments to increase their visibility and reputation. This system requires a certain upfront effort to verify expert users manually (unless they can be matched with public registers and some form of identification).

Privacy is the another concern. Obviously, a lawyer would not want - and is not permitted by law - to publish insights about an actual case. As described earlier, data sits on a scale. Even though openlaws is advocating for openness and transparency, there is a clear distinction between open, shared and closed data. This means comments can be kept entirely closed, they can be shared in groups, or they can be broadly published.

With the general trend toward open access and open access publication - also promoted by the European Commission in the Horizon 2020 funding scheme - it is realistic that more legal content will be available for free. It would be useful for the general public, if public administration would share their comments and findings on case law and legislation in comments that are available for everybody.

Red Flagging

What if users could actively inform public bodies about problematic legislation? For software it is pretty common that users can provide a bug report. For (legal) code there is no such mechanism on a large scale. Of course, legal scholars, lawyers and other experts discuss legal problems in journals, books or reports. But identification “on-site” does not happen. It would be relatively easy to introduce feedback forms or simply buttons, where (expert) users could “red-flag” provisions.

For example, experts could inform public bodies about implementations, which are not fully in line with the underlying European directive. Or about legislation, which is not in line with the respective constitution of the Member State.

Legal ICT information can empower users easily - may it be the governmental databases or systems like openlaws. The more interesting question then is of course, what to do with such information?

Collaboration

openlaws is designed around collaboration. In the „social layer" of openlaws users can collaborate in open or closed groups, always centred around legal information. Users can share links to legislation and case law, they can share their legal portfolios and their comments. Users can follow certain public portfolios, so that they do not have to re-invent the wheel.

For example, a law student might want to work together with others for the next exam. A legal scholar might want to share legal research with others. A lawyer might want to collaborate with colleagues in a law firm to manage their combined knowledge. Or the lawyer might want to publish a few case discussions in order to increase visibility and credibility. A company might want to inform employees about the relevant labour law provision.

Find a Lawyer

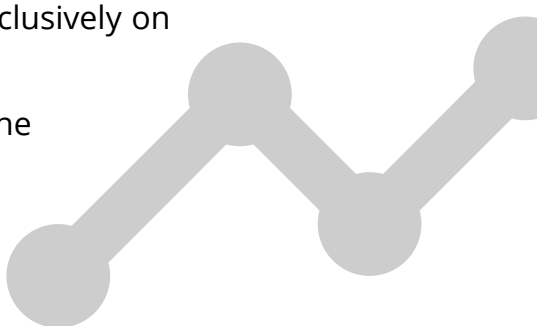
Services to find a lawyer or a notary exist on a national and on a European level, sometimes even with a user-friendly visualisation on a map. However, the combination of big open legal data with information about legal experts could be of interest. Lawyers who are appearing often in case law that deals with competition law, are likely to have a high expertise in competition law. Winning or losing a case is probably not a strong indicator, simply appearing in such cases shows that experts are active in this field. The names of experts can be extracted from the data and linked to the expert's profile as a strong competency. This in return can generate new clients.

Visualization

Visualization of legislation and case law can be a useful tool for navigation through large amounts of legal information. Linked case law with all its references and citations is a promising candidate for graphical representations, other than in traditional text form. The openlaws graph database already contains linked case law, which is a pre-condition for such applications.

As of today, visualization are rather seen as nice-to-have features and there are few universities, start-ups or publishers who focus exclusively on visualization.

However, given the latest developments in the USA and the increase in data analytics, visualisation has a high potential for the future. By 2020 there may be more user-friendly



visual applications, that help to navigate through large corpora of legal documents or that structure the contents of databases in a different way.

Translations

Language is a classic barrier in the legal domain, especially in the European Union, with its 24 official languages. Even though less challenging compared to answering legal questions, this task is still extremely complex. Legal language is special. The style of formulations and expressions and the special terms used are different than the average language. In addition, high quality translations are needed. It is just not enough to get the meaning of an article, the legal expert needs to understand the smallest differences in a legal text. A machine-translation may just not be enough to reach this high level. Still, automated machine translation could be a first step in the right direction. While the machine makes the first attempt, humans could do the fine-tuning in a second step. The aggregated openlaws database could be a good starting point for the developments and explorations in the upcoming five years.

Text and Data Mining (DTM)

Text and Data Mining, which is roughly equivalent to text analytics, refers to processing of text for deriving and retrieving high-quality information. Given the large amounts of legislation and case law, systematic analyses can be beneficial, not only from a scientific but also from a practical perspective. A large database with a lot of content is predestined for data and big data analysis. Even though legal databases cannot be considered big data under today's definitions (volume, velocity, variety; see above), but the amounts of text are still "too big" for being simply read, monitored and analysed by humans. New services could analyse the data automatically (usually it is text) and transform raw data into digested information. The idea is to understand text information by analysing such text with automated algorithmic tools. Such analyses are mathematically complex and rely on well chosen data structures to achieve acceptable response times and results. Analyses might include counting, basic average calculation or complex statistical analysis like a complexity index for legal documents.

A look to the USA may provide insights where the road is taking the legal domain. Nonetheless, services are already on the market. Lex Machina, a LexisNexis company, is offering analytics on intellectual property (IP) law. The service will tell the user the chances for approval or rejection of a patent application, based on the information collected from the existing databases.

Content Recommendation

Completeness of information is important in many professions. The same is true for law. Experts do not want to miss relevant information. A digital legal assistance system, that comes up with recommended content is therefore very useful. Content recommendation is based on so called recommender systems. They can be seen as some kind of information filtering method, aiming to predict certain aspects of user behaviour, preferences or profile data to promote interesting or interesting information which someone might not even have thought of.

Content recommendation is very popular in electronic market places, such as for example Amazon. The system will tell the user which other products may be interesting for her or him. Building a bridge to legal information, a system could suggest related legislation, case law, literature or even legal experts.

Artificial Intelligence

Wouldn't it be nice if you could ask the machine a legal question and it could come up with the answer right away? This is of course science fiction, but large players are already working on training machines. Data sources are an important basis.

For example, ROSS is a legal service built on top of IBM Watson. The value proposition is that users can ask questions in plain English and ROSS then reads through the entire body of law and returns a cited answer and topical readings from legislation, case law and secondary sources. This task is inherently difficult because various jurisdictions have grown over maybe hundreds of years and are - although maybe quite similar in their implications - very differently formulated and also written in different languages. Consequently an algorithm needs to recognize and extract the real meaning of a legal document - a problem where even legal experts struggle. This is also the good news for legal experts: The machine will not have replace lawyers by 2020 - but it will assist them.

Read more about openlaws.eu

- Requirements for enrichments tools (Deliverable 2.2.d1)
- Final specification and vision of enrichment tools (Deliverable 2.2.d3)
- Final BOLD business models for stakeholders (Deliverable 2.4.d2)

Deliverable are available online via http://ww.openlaws.eu.

We must all obey the
great law of change.
It is the most powerful law of nature.

Edmund Burke

Policy Recommendations

PSI - Open Data

The primary EU regulatory instrument in place that promotes the availability of free to use national legal data is the EU Directive on the Re-use of Public Sector Information of 2003/2013 (or PSI Directive).[1] The PSI Directive stimulates public sector bodies to allow the commercial exploitation and non-commercial use of their data resources. Broadly speaking, it requires public authorities to be transparent about what information it has available, and to limit conditions and charges wherever feasible. The instrument itself does not mandate public access to legislation, case-law, the legislative record or commentary. It only applies to information (technically: documents) that is public under national law and in which third parties have no intellectual property rights. For these reasons, the PSI Directive has less actual 'punch' than might seem at first. Arguably, as official legal texts are of fundamental importance to democratic states and the rule of law, this is a category for which the PSI Directive could be stricter, by imposing that official texts be made available for re-use free from any constraints (e.g. creative commons zero), or under a liberal open license. At the very least,[2] the European institutions could jointly push the voluntary application to legal information of open licenses as per the EC Guidelines on Recommended Standard Licences, Datasets and Charging for the Re-Use of Documents.[3]

Copyright and Database Rights

The public sector is the core source of legal information on which linked open data systems rest. For pan-European services to develop, a measure of legal certainty is required that is currently lacking. There is no uniform notion of the kind of legal works that are subject to copyright. Every Member State has its own rules. There are some similarities among certain Member States, e.g. that legislative acts are not subject to copyright. Overall however, the status of much legal information like court records, parliamentary records, technical standards incorporated into

1 Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the Re-use of public sector information, OJ 2003, L 345/90, as revised by Directive 2013/37/EU, OJ 2013, L175/1.

2 To what extent the EU can legislate how and when official texts of the Member States must be made public is a contentious issue. Arguably the legal basis for the PSI Directive (art. 114 TFEU, i.e. approximation of laws for the improvement of the conditions for the functioning of the internal market) is insufficient in this respect.

3 These guidelines accompany the revised PSI Directive: 'Commission Notice: Guidelines on Recommended Standard Licences, Datasets and Charging for the Re-Use of Documents', OJ 2014, C 240/1–10.

health & safety or building laws differs per jurisdiction. In addition, in many jurisdictions the precise scope and operation of copyright provisions concerned with legal information is unclear. One salient question is what the legal status is of official texts that emanate from supranational or foreign authorities. Such uncertainty creates transaction costs. In light of the advanced state of harmonization of EU copyright law, it makes sense for the EU legislator to develop uniform standards. Since legal information can be subject to both copyright ('works') and the so-called sui generis database right, the latter must be included for any reform to be effective. The ECJ has ruled that the fact that the contents of a database (legal texts) are not subject to copyright does not rule out that the database as such qualifies for sui generis protection (or protection as a collective work under copyright). But the precise applicability and necessity of the sui generis database right for official legal data is controversial.

Unfortunately international copyright treaties are of little help in the drafting; the WIPO Copyright Treaty and TRIPS do not address official texts directly. The Berne Convention leaves it to States to determine whether official documents are copyrighted (Article 2(4) BC) and to what extent speeches delivered in political or legal proceedings are public domain (Article 2bis BC).[4] On a positive note, because the Berne Convention leaves much to the discretion of parties, the EU can easily formulate its own policy in compliance with international law.

The so-called Wittem Copyright Code provides an excellent starting point. It proposes that excluded from copyright are a) Official texts of a legislative, administrative and judicial nature, including international treaties, as well as official translations of such texts; and b) Official documents published by the public authorities.[5] Inspiration for sui generis database reform can be taken from national examples. The Dutch implementing act of the Database Directive for example stipulates that databases of official texts that are exempt from copyright protection (laws, decrees or ordinances issued by public authorities, judicial and administrative decisions) are not protected sui generis when produced by government.[6]

Privacy

Of the types of legal public sector data discussed here (legislation, case law, preparatory works), the release of case law as open data poses perhaps the biggest challenge from a data protection and privacy perspective. Legislation is unlikely to contain personal data or have privacy impacts. Preparatory works, e.g. parliamentary records, will typically contain personal data such as names of politicians/

4 The (foreign) author of such speeches however must be granted the exclusive right of making a collection of such works (art. 2bis(3) BC).
5 http://www.copyrightcode.eu
6 Art. 8(1) Databankenwet (Database Act).

members of parliament, public officials, and sometimes also of citizens or persons acting for private sector organizations. As a rule, how and at what point in time information enters the official public legislative record is subject to safeguards (e.g. rules of procedure of the senate or house) that balance privacy interests against the general interest in transparency of proceedings and outcomes. Public scrutiny and debate is of course of utter importance in lawmaking processes. Transparency is universally regarded a precondition for the effective exercise of political rights and freedoms, and for ensuring accountable public authorities. In the case of a politician objecting to the disclosure (mandatory under national law) of expenses claims, the European Court of Human Rights assessed the freedom of expression interest and the privacy interest. It stressed the importance of internet access to the claims:

> *"[t]he general public has a legitimate interest in ascertaining that local politics are transparent and Internet access to the declarations makes access to such information effective and easy. Without such access, the obligation would have no practical importance or genuine incidence on the degree to which the public is informed about the political process."*[7]

This reasoning is even more convincing where the information concerns law making processes.

The situation for case law is arguably somewhat different. To be sure, there is a strong general interest in transparency of judicial proceedings; this also shows in a variety of constitutional norms, statutory provisions and court practices. But the information contained in court decisions (let alone the wider category of court records) is often sensitive as it pertains to the private lives of parties or affects reputation. Judicial authorities deal with the tension between transparency and private interests differently. In light of the far-reaching effects of publishing case-law as open data, online and for everyone to use without time restriction, the more likely route to a linked open data environment includes the use of standardized, predictable and robust anonymization guidelines. A point of attention is the independence of the judiciary; in some Member States this also means that courts develop their own publication policies (including anonymization). This can affect what the most promising forum is for the development of common standards.

7 Wypych v. Poland, ECtHR 25 October 2005 (Admissibility decision (dec.), no. 2428/05)

The possibilities to make available case-law as open data without anonymization seem limited under the coming General Data Protection Regulation. One central problem is that making personal data available for uses other than the public task it was collected for, easily breaches the purpose specification principle, which is central to current and coming EU data protection law. Also, releasing case law to the public for alternative uses is an act of processing personal data. This in itself requires either that the data subject gives consent (which cannot easily be really freely given by e.g. litigating parties, victims or persons prosecuted) or that the disclosure is lawful on one or more other grounds for processing that data protection law recognizes.

Of note, even with anonymization processes in place, it is the linked aspect of open data that can cause problems in the end. The more data sets governments disclose, the richer the possibilities for re-identification. Irreversible anonymization of judgments might be difficult to achieve.

Standardization

Standardization should be continued and strengthened. There are several standards for legal information that can be applied. The European Case Law Identifier (ECLI)[8] is promising and several initiatives are already working with it (for example the "Building On ECLI" EU project).[9] ECLI needs to be promoted not only among European and Member State institutions, but also among legal experts, such as lawyers, judges, etc., so that ECLI is actually used in texts and metadata. Only then a broad network of references can be established. This includes that ECLI is also used correctly. For example, skipping the "ECLI:" element in the reference, destroys the potential of ECLI. The reference cannot be detected automatically anymore. Judges from the European Court of Justice as well as judges from the Member States should be trained.

Another problem for the general uptake of ECLI is the individual ordinal number that is decided upon by each Member State without any guidance. An essential requirement would be that the ECLI can be created by users simply by knowing the national court case number. Subsequently, ECLIs could be created for legacy cases easily. Assigning an ECLI top-down from a central position is not recommended.

8 https://e-justice.europa.eu/content_european_case_law_identifier_ecli-175-en.do
9 http://bo-ecli.eu

Akoma Ntoso defines a set of technology-neutral electronic representations in XML format of parliamentary, legislative and judiciary documents.[10] Akoma Ntoso was designed to be compliant with CEN Metalex. Akoma Ntoso is used in the UK Statute Law Database, the United States Library of Congress and the European Parliament.[11] More structure in parliamentary, legislative and judiciary documents is strongly recommended, as this creates more and better opportunities for re-use.

Just like for ECLI, the standardization for the ELI (European Legislation Identifier) should be continued.[12] As of today, the openlaws project could no rely on this identifier. Since the implementation of ELI is optional, Member States should not only be informed about what they can do to implement it, but they need to better understand why they should use it at all.[13] The openlaws project will continue to use its best efforts to demonstrate to Member States the benefits of a linked legal infrastructure.

Open Innovation & Open Government

Open innovation and entrepreneurship are a key dimension of the Digital Agenda for Europe which is the first of seven flagships initiatives under Europe 2020, the EU's strategy to deliver smart sustainable and inclusive growth.[14] Based on openlaws.eu research results, we see a big potential in the combination of the Digital Agenda and big open legal data. While ICT technology and innovation capacity may lack behind in Europe compared to the USA (the most famous ICT services are based in the USA), the EU is clearly ahead in certain areas of open data. The legal domain is one of them. As described in this report, there are already concepts and best practices for open legal data in Europe. Legal tech is clearly gaining momentum and the EU and its Member States should use open innovation principles to stay ahead. While open innovation is still relatively unknown in the legal domain, legal experts can benefit from this approach. During the openlaws code camps and hackathons, we have seen the power of bringing together innovative ICT experts and legal experts. During these events legal experts have presented several problems, and ICT experts have tried to come up with possible ICT-centered solutions. Of course, the machine cannot solve everything, but identifying areas where digital legal assistance can be applied, can serve time and

10 http://www.akomantoso.org
11 https://blog.law.cornell.edu/voxpop/tag/akoma-ntoso/
12 http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52012XG1026(01), http://www.eli.fr/en/
13 See http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV%3Ajl0068
14 http://www.oi-net.eu/m-oinet-network/m-news/m-news2/153-a-eu-digital-agenda-open-innovation

money. Accordingly, we suggest to include more interdisciplinary open innovation events and workshops in the training sessions of DG Justice as well as collaboration between DG Justice and DG Connect. This suggestion is also in line with the Commission's "Vision for Public Services":[15]

> *The current social, technological and economic changes create challenges and new expectations for public services. Given that these challenges are largely intertwined, any vision for the future of public services needs to have a multi-disciplinary approach. A solution may be embracing open government, based on the principles of collaboration, transparency and participation within an appropriate governance framework. Such an open government model builds on open data, open services and open decisions. The provision of public services results in the creation of public value. Empowering individually and collectively all actors that play a role in the constitution of society and sharing resources between all stakeholders will contribute to the creation of public value.*

Read more about openlaws.eu

- State-of-the-art report for legal, social and business aspects of re-use of legal information (Deliverable 1.1.d1)
- Synthesis report (Deliverable 1.3.d1)
- Comparative country report: White paper on the OpenLaws.eu open innovation community (Deliverable 1.3.d2)

Deliverable are available online via http://ww.openlaws.eu.

15 https://ec.europa.eu/digital-single-market/en/news/vision-public-services, https://ec.europa.eu/programmes/horizon2020/en/news/ict-enabled-public-sector-innovation-h2020-video

Related EU Projects

BO ECLI

Building on ECLI (BO-ECLI) is a project involving sixteen partners from ten Member States (Italy, Greece, Croatia, Estonia, Belgium, the Netherlands, Germany, the Czech Republic, Spain, Romania) that aims to broaden the use of ECLI and to further improve the accessibility of case law. The first objective of the project is to (further) introduce ECLI into the case law repositories of Belgium, Italy, Greece, the Netherlands, the Czech Republic, Germany, Estonia and Croatia and to connect them to the ECLI Search Engine of the European e-Justice Portal ('ESE-EEJP').

BO-ECLI is co-funded by the Justice Programme of the European Union. Having started at 1 October 2015 the project is planned to be finished within 18 months with a maximum duration of 24 months.

http://bo-ecli.eu

Big Data Europe

Big Data Europe will undertake the foundational work for enabling European companies to build innovative multilingual products and services based on semantically interoperable, large-scale, multi-lingual data assets and knowledge, available under a variety of licenses and business models.

https://www.big-data-europe.eu/

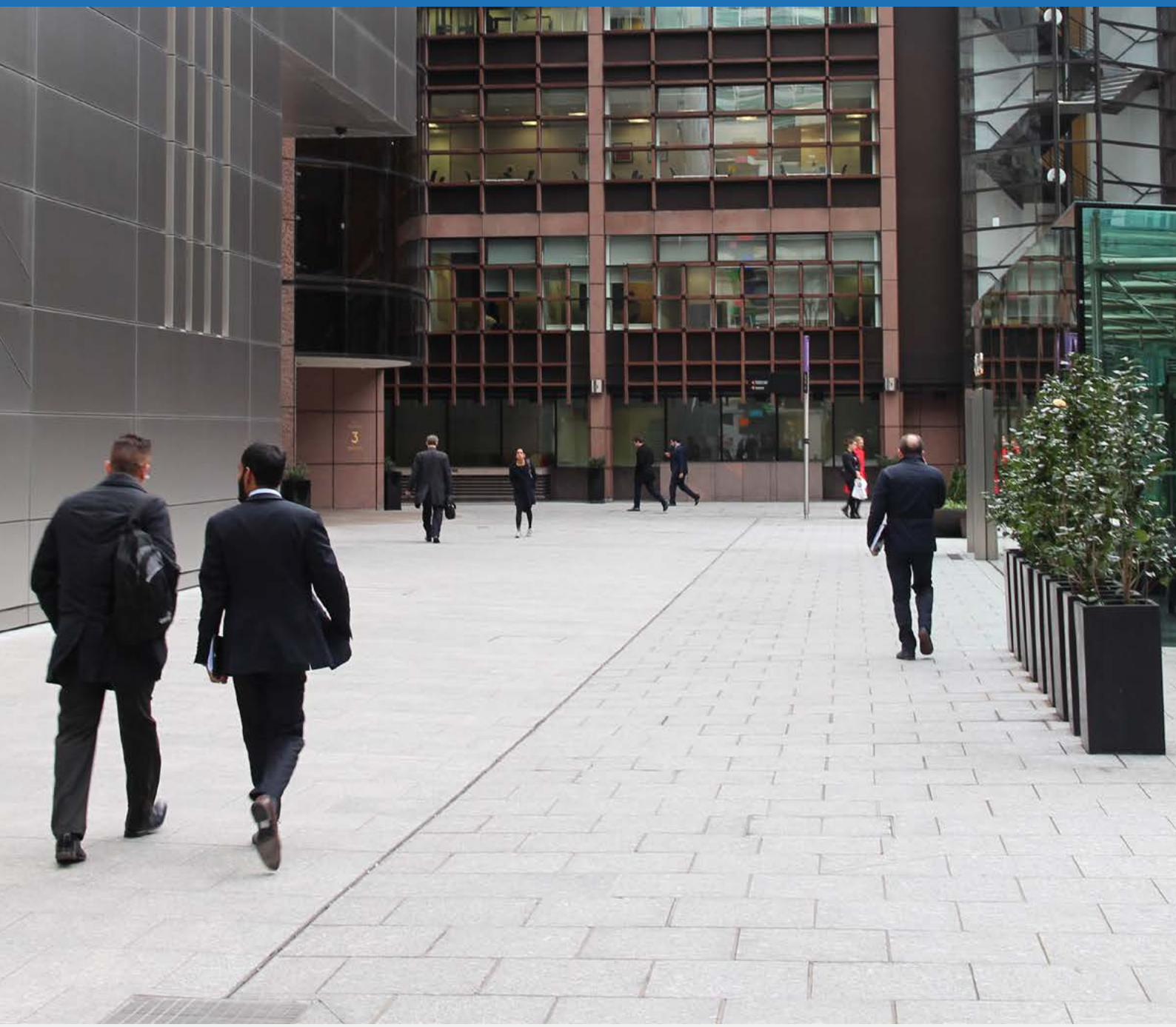
EU CASES

EUCases was a collaborative Research Project supported by Seventh Framework Programme (FP7) funding. The project developed a unique pan-European law and case law Linking Platform transforming multilingual legal open data into linked open data after semantic and structural analysis. EUCases-Legal Linked Open Dataset is represented in RDF graphs and uses the SPARQL query language. The input documents have been encoded in the legal XML scheme Akoma Ntoso. Thus, the schema has been converted into an appropriate RDF presentation for the purposes of linking.

http://www.eucases.eu/

Finally, strategy must have continuity. It can't be constantly reinvented.

Michael Porter



openlaws Strategy

In recent years a new category of start-ups has emerged, the so-called social entrepreneurs. There is no generally-accepted definition of social entrepreneurship, but the general idea behind the concept is that “doing good” and “doing well” can and should be combined. Business mechanisms are introduced in order to solve social problems.

For the research project “openlaws.eu” this means that a “social entrepreneurship” business model can be applied to the project spin-off “openlaws.com”. This spin-off can be built on a social entrepreneurship mind-set and on an online solution that offers free legal information for citizens and businesses that is relevant to their everyday lives.

However, the fact that legal information is supposed to be free leads to considerable restrictions. Still, in the age of the Internet many services and business model innovations have been implemented. In his book “Free”, Chris Anderson describes different ways how “free” can work.

Developing more sophisticated features for experts, researchers, administrations and other power-users will create added value and generate revenues for the spin-off. Additional features will not necessarily all be developed by openlaws.eu or by the project spin-off directly, but may be implemented by third parties, who may know the needs of the users even better than the research openlaws.eu core team or the spin-off team (especially in specific expert domains and in specific legal settings of different jurisdictions). The basis for such third-party applications will be the aggregated database (the “BOLDbase”).

The openlaws.eu solution that is deployed to users under the EU project funding has reached the level of a working prototype (Technology Readiness Level 6, or “TRL 6”). The strategic business plan foresees additional development activities in a second step, in order to take the research prototype to a fully operational level (TRL 9) and on-going exploitation, is done by the project spin-off “openlaws.com”.

Business Model

The end of the openlaws.eu project, supported by DG Justice, is just the beginning of the implementation of the business model for openlaws. During the project the team has spent time and effort on investigation about how a venture with a social impact – namely better access to justice – can be made sustainable. Legal information is a precious "public good" in every democracy and citizens and businesses should know their rights and obligations. This is also key in European Union, where the law should not be used to build walls, but rather to form a common single market. However, if providing access to law is a such an important public task, then shouldn't it be up to governments and European institutions? And don't they do it already to a sufficient extent? While one can always argue about the right quality, the simple answer to the question is: Yes, they do. Even better, as a result of the PSI directive, the EU and its member states provide the information as open data, allowing others to build new and innovative solutions on top of primary legislation and case law.

Whenever value is created, a smaller part of such value can be captured for the purpose of ensuring the continuation of this value-generation process. This is what business models do. They describe the way how value is built and the mechanisms to maintain the venture.

In today's Internet economy, a predominant business model for cloud services has emerged. So called "freemium" services offer free services to the general public, while premium services are provided to paying customers. Criticism on free services exists of course, arguing that whenever something is free, then "the user is the product" that is being sold elsewhere. While this – unfortunately – may be true in many cases, it is not a given fact that it has to be this way. In the area of open source software or in the area of Creative Commons works, people are voluntarily making software code or other works (photographs, publications) available for free.

For a business, the situation is different of course. In order to run a venture, bills have to be paid. Cost drivers for IT companies are salaries and IT infrastructure costs (even though becoming cheaper, the costs can be still very high, depending on the traffic and storage). In addition, tax has to be paid, even if it the company is a venture with a social purpose. Tax money is of course the income of the EU and its member states, which makes it possible to develop public legal information systems and open data interfaces. A social venture has no tax income, only tax payments. It has to rely on other sources of income to make the venture sustainable.

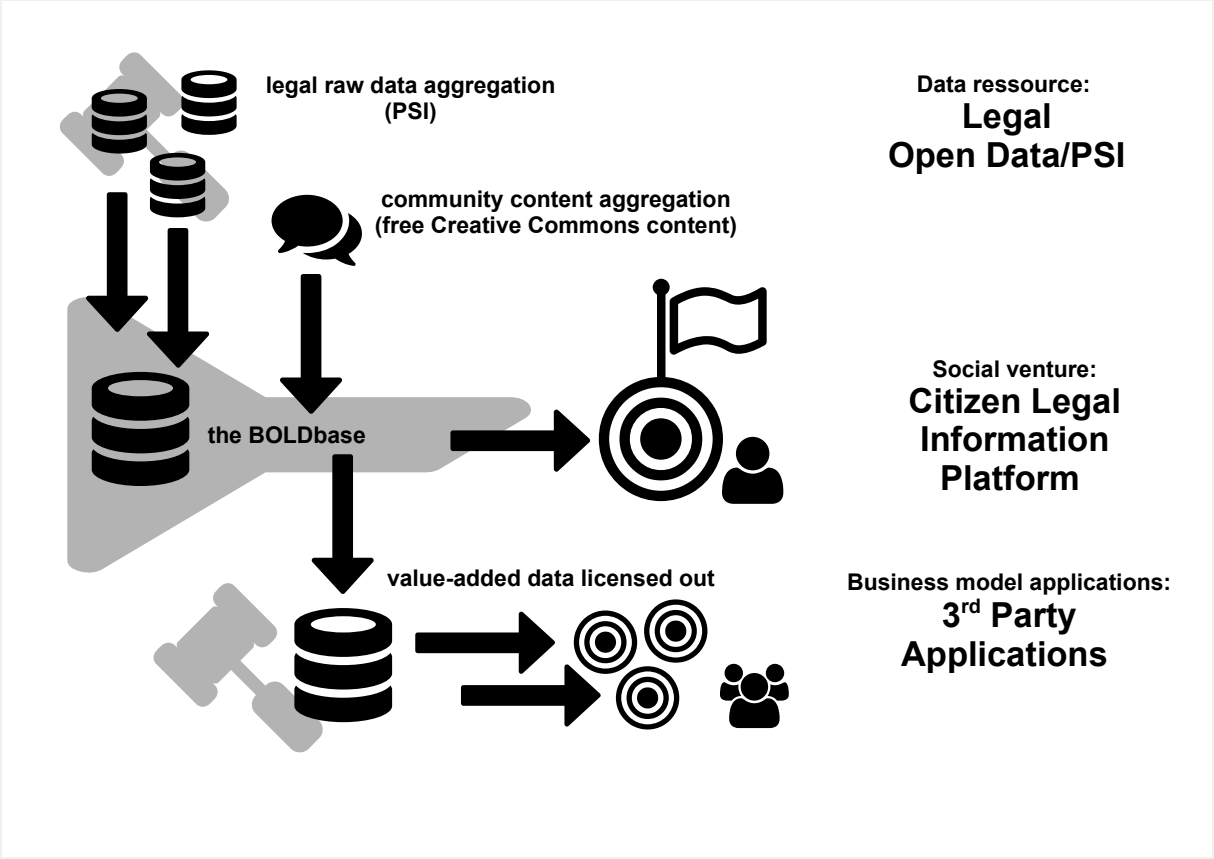


Figure 8: Open innovation scenario

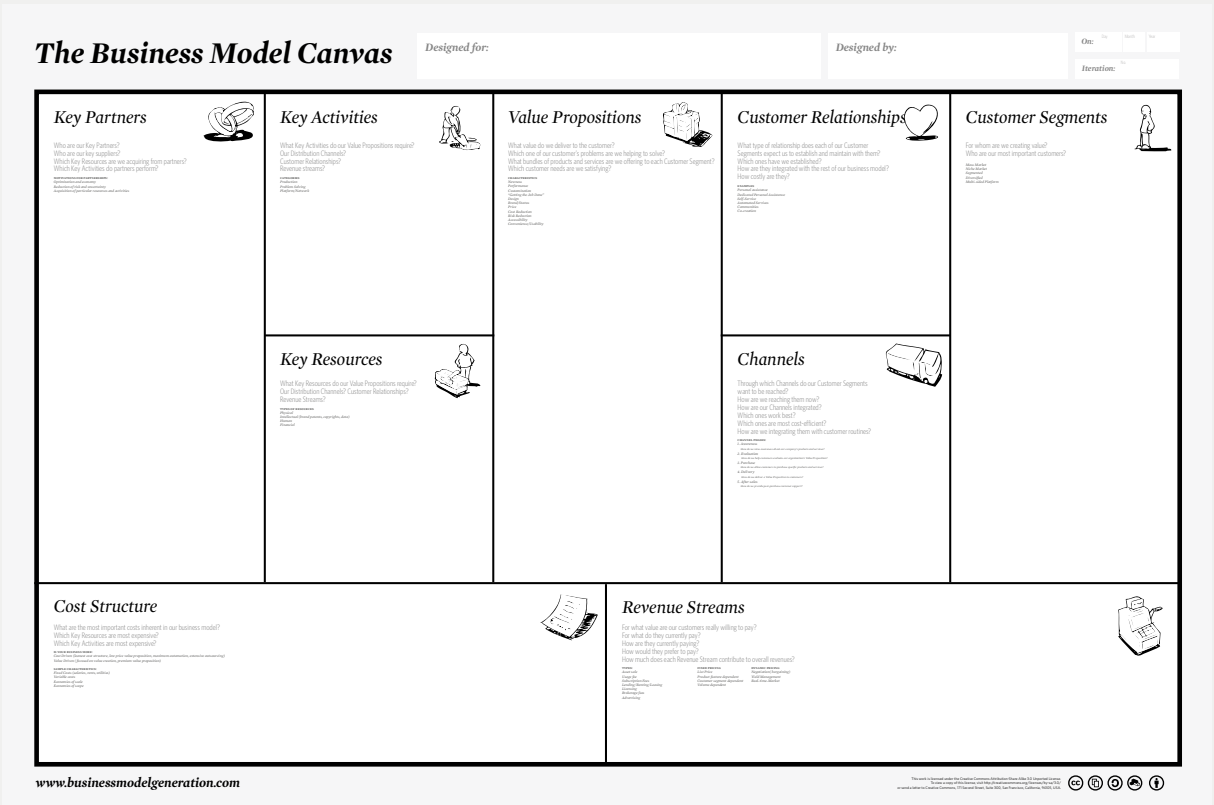


Figure 9: Business model canvas

Freemium is way to generate income. Such model has to be transparent enough so that users do not become the product and do not even get the impression of becoming the product. Such trust can only be built over time and with a lot of transparency and communication. Instead, it should be a common joint effort, where users see and understand that money is needed to operate the service.

One example is Wikipedia. People reveal information for free, know that their knowledge is "the product" and understand that donations are necessary to pay for the infrastructure of Wikipedia. openlaws is no Wikipedia of course. Replicating a model like Wikipedia and building large networks is one of the most challenging activities. Also, the legal community is much smaller compared to the general public. For example, there are 8 million citizens in Austria, but only 6,000 lawyers. When it comes to legal informatics, the situation becomes even more tense. The number of people who can code AND have the sufficient knowledge about the law is very limited. Such people usually work at legal information departments at universities or for legal publishers.

For these reasons, a slightly different model has been chosen for the openlaws spin-off company, which was incorporated in 2015 in Austria. While an ongoing stream of donations or – even better – a payment by each and every lawyer (see CanLII case study) would be preferred, it is no feasible option for the time being. Therefore, the solution is to offer free services for the general public, while the more advanced features are premium features under a subscription payment model. This ensures that the general public has better access to law, while the experts and large business users with legal departments will pay for the service (e.g. for closed group functionality).

In November 2015 the openlaws spin-off received ODI start-up status. The ODI (Open Data Institute) was founded by Sir Tim Berners-Lee and Sir Nigel Shabolt. The institute promotes the usage and creation of open data. With the ODI as a strategic partner, openlaws is confident that a sustainable business can be created, which also has a positive social impact in Europe - even beyond the EU project.

Read more about openlaws.eu

- Handbook for stakeholders (Deliverable 4.1.d3)
- Socio-economic framework for BOLD stakeholders (Deliverable 2.3.d1)
- BOLD socio-economic and governance framework (Deliverable 2.3.d2)
- Final BOLD business models for stakeholders (Deliverable 2.4.d2)

Deliverable are available online via http://ww.openlaws.eu.

openlaws Technology

Responsive Frontend: Angular JS

openlaws can be used on desktops as well as on mobile devices. The frontend adjusts to the device automatically (responsive design) thanks to Angular JS. AngularJS is a complete JavaScript-based open-source client-side web application framework mainly maintained by Google and by a community of individuals and corporations.

https://www.angularjs.org

Graph Database: Neo4J

Neo4j is a graph database management system developed by Neo Technology, Inc. Neo4j is the most popular graph database. Neo4j is implemented in Java and accessible from software written in other languages using the Cypher Query Language through a transactional HTTP endpoint.

http://neo4j.com

Search: Elastic Search

Elasticsearch is a search server based on Lucene. It provides a distributed, multitenant-capable full-text search engine with an HTTP web interface and schema-free JSON documents. Elasticsearch is developed in Java and is released as open source under the terms of the Apache License. Elasticsearch is the most popular enterprise search engine followed by Apache Solr, also based on Lucene

https://www.elastic.co/products/elasticsearch

Programming Language: Java

Java is a general-purpose computer programming language that is concurrent, class-based, object-oriented, and specifically designed to have as few implementation dependencies as possible. As of 2016, Java is one of the most popular programming languages in use, particularly for client-server web applications, with a reported 9 million developers.

https://en.wikipedia.org/wiki/Java_(programming_language)

FAQ

Why do we need openlaws?

Everybody has a legal problems from time to time. More and more legal questions have not only a national but also a European dimension. Since law is the basis of our democracy, legal professionals, citizens and businesses should have adequate access to legal information.

Isn't there anything similar already?

There are free governmental databases with basic functionalities and subscription professional databases, but no community driven open platform that includes legal experts.

What's different compared to Internet search engines?

openlaws will provide very specific functionalities for law. While it can be difficult in a normal search engine to narrow the results down to legally relevant documents, openlaws will only provide legal results.

How does it differ from research networks in other sciences?

While research networks typically combine experts and the community, there is no legislation or case law included. These element are of utmost importance when you have a legal question.

Will I find the full text of legal papers and books?

A typical answer from a lawyer: It depends. If an author chooses to publish under an open access license then yes. However, there will be premium content from commercial publishers who will charge for the article or e-book. We aim for a comprehensive overview so that you do not miss anything.

What does "map the law" mean?

One basic principle of openlaws is open innovation. We believe that we can create a great legal information system together where we collect (or "map") the available legal content jointly.

Which software and tools are you using?

openlaws will use a lot of open source software. We will on an open source content management system and state-of-the-art graph databases. openlaws will use proprietary solutions only in cases where it is not feasible to use open source solutions.

Where is openlaws located?

The platform itself will be operated only on servers that are located within the European Union. The core team members are based in the Netherlands, the UK, Austria and Italy. A project spin-off was established in Austria.

Isn't legal information too sensitive to work with in an open environment?

openlaws does not share confidential information to you or your clients. It helps you search, organise and share legislation, case law and legal literature that is publicly available.

Does open also mean free?

Access to case law and legislation will be free. Legal literature may be copyright protected and therefore will not always be free of charge. The platform will also host premium features for legal professionals who wish to access those publications.

Read more about openlaws.eu

- Leaflet (Deliverable 4.1.d1)
- Information brochure (Deliverable 4.1.d2)
- Handbook for stakeholders (Deliverable 4.1.d3)
- Infographic (Deliverable 4.1.d4)
- Promotional video (Deliverable 4.1.d6)

Deliverable are available online via http://ww.openlaws.eu.

Coming together is a beginning;
keeping together is progress;
working together is success.

Henry Ford

EU Project Participants

University of Amsterdam

The University of Amsterdam is the lead partner of the openlaws.eu project. The Leibniz Center for Law of the University of Amsterdam started in 1988 and currently employs about 10 people. It is part of the Faculty of Law and maintains strong connections with the Science Faculty. The Leibniz Center develops intelligent technology to support legal practice both in the private and in the public sector. It applies Artificial Intelligence (AI) techniques to problems in legal theory, legal knowledge management and the field of law in general. In this capacity, it participates in many (inter)national research initiatives and maintains strong ties to the international research community and government agencies. The Leibniz Center has participated in and coordinated numerous national and European (applied) research projects and has coordinated various European Commission sponsored projects including the 6th framework project ESTRELLA, the Leonardo project TRIAS and the eParticipation project SEAL.

http://www.leibnizcenter.org

Salzburg University of Applied Sciences

The school of Information Technologies and Systems Management (ITS) of Salzburg University of Applied Sciences (SUAS) offers two BA and three MA programs to a total of 350 students and is focused on software engineering, systems engineering, network technologies, and the cross-cutting application field of eHealth. SUAS's staff comprises faculty with long track records in ICT research, mathematical and biomedical modelling. SUAS has participated in several EU projects and is eager to extend its competences in the fields of self-adaptive software systems and symbiotic system architectures.

http://www.fh-salzburg.ac.at

University of Sussex

Law at Sussex was rated 16th in the UK in the 2008 Research Assessment Exercise. Research within the Sussex Law School is organised around thematic research groups, supporting research by hosting conferences and workshops, engaging in interdisciplinary research and collaborative projects. Law's expansion has enabled

the development of new research areas, notably information law, pursuing an ambitious research agenda. The Information Law cluster runs two major EC projects, Internet Science and openlaws, led by Professor Marsden.

http://www.sussex.ac.uk

London School of Economics

The LSE is regarded as one of the world's leading academic institutions and remains a specialist single-faculty constituent college of the University of London, the only such institution in Britain. The aim of LSE's Media and Communications Department is to keep pace with rapid change in media, technology and society demands through dynamic and imaginative re- search. Based on the results from the 2008 Research Assessment Exercise, on grade point average the department is rated third-best in the UK.

http://www.lse.ac.uk

Alpenite srl

Alpenite is an IT software consulting and system integration company with headquarters in Venice, Italy. The major areas of expertise of Alpenite are in portals, mobile applications, e-commerce websites and business intelligence, using open source technologies and integrating with enterprise components where needed. Alpenite has developed strategic relationships with open source vendors and has competencies in the main open source products available on the market.

http://www.alpenite.com

BY WASS GmbH

BY WASS has initiated the openlaws project. The focus of the company is to innovate in the legal sector. BY WASS has developed the RIS:App, the offcial mobile interface for the award-winning Austrian legal information system. The top-ranking app is based on Austrian Open Government Data and was downloaded over 50,000 times. Considering that there are only 5,000 lawyers in Austria, BY WASS has successfully proven that a user-friendly legal service with added value can attract many users – including average citizens – in a short period of time. BY WASS is shareholder in the openlaws.eu project spin-off company openlaws gmbh.

http://www.bywass.com

Project Spin-Off: openlaws gmbh

openlaws gmbh is an Austrian limited liability company. The start-up was incorporated in 2015 as a direct result of the openlaws.eu project with the ambition to take the research results of openlaws.eu to market-readiness and to operate the venture beyond the project duration.

The company is lead by Clemens Wass (law) and Christian Sageder (technology). They are supported by a team of software developers and legal experts. The early-stage investor AC & Friends is providing additional information and background, which is necessary to make the venture sustainable.

The offices are is based in Salzburg/Austria. Salzburg has a long tradition with respect to legal informatics. The International Legal Informatics Symposium (IRIS) is one of the largest legal tech conferences in Europe and held in Salzburg every year (https://www.univie.ac.at/RI/IRIS16/information-in-english/). The IRIS will celebrate its 20th anniversary in 2017. In addition, Austria is one of the leading e-government countries in the world. This makes Salzburg an excellent strategic location for openlaws.

https://www.openlaws.com

openlaws

openlaws will help you find legal information more easily, organize it the way you want and share it with others. The Internet platform is adding a “social layer” to the existing “institutional layer” of legal information systems.

This “BOLD Vision 2020” provides a picture of the near future. The document brings together some of the key conceptual insights behind open data, open innovation, big data and legal technology.

BIG OPEN LEGAL DATA