Linked Data Business Cube – Modelling Semantic Web business models

Tassilo Pellegrini[1], Christian Dirschl[2], Katja Eck[2]

[1]St. Pölten University of Applied Sciences, Matthias Corvinus Str. 15, A-3100 St. Pölten, Austria, tassilo.pellegrini@fhstp.ac.at

[2]Wolters Kluwer Deutschland GmbH, Freisinger Straße 3, 85716 Unterschleißheim, Germany, {CDirschl}{Keck}@wolterskluwer.de

Abstract: With the increasing availability of semantic data on the World Wide Web and its reutilization for commercial purposes, questions arise about the economic value of interlinked data and business models that can be built on top of it. This paper introduces the Linked Data Business Cube, a heuristic approach to model and analyze business models for Linked Data assets. The Linked Data Business Cube provides an integrated view on Linked Data assets, revenue models and stakeholders, thus allowing investigate the specificities and interdependencies of business models under conditions of networked infrastructures and collaborative value creation.

Keywords: Linked Data, Business Model, Content Value Chain, Media Economics, IPR, Data Licensing

1. Introduction

From the large amount of information produced every day just about 5% is "structured" [23]. But 92% of all analytical activities are exercised on top of structured data [21]. The remaining data is currently hardly utilized or discarded as a whole. Hence new approaches are being developed to improve the machine-processability of available data ideally by not just creating more structured data but also by applying structural principles that support interoperability at an infrastructural level. One of these approaches is called Linked Data [29].

Anecdotal evidence supports the hypothesis that Linked Data is an enabling technology to improve workflow efficiency and trigger business diversification [24]. Accordingly, Linked Data strategies can be very diverse and context-specific, covering the spectrum from improving access to open data provided by governmental bodies, to improvements of workflow efficiency in various industrial sectors like automotive, media & publishing or the life sciences. A look at existing case studies [20][9][5][25][26][28] reveals that Linked Data is implemented along the incremental IT development practices of enterprises and public organizations, but additionally brings along disruptive technological effects that pose significant challenges to and opportunities for business development [1]. This includes foremost an appropriate licensing strategy that takes

adfa, p. 1, 2011.

account of the various asset specificities of Linked Data as intellectual property. Additionally legacies and policies deriving from historically grown information architectures influence the transition from silo-based systems to networked databases and repositories as a precondition for smarter products and services. And third, the Linked Data infrastructure in terms of quality-approved and commercializable datasets, tools and services has not yet reached a critical mass but benchmarks indicate a steadily growing adoption rate [12].

This line of argument motivates to take a closer look at the business model implications of Linked Data and the added value derived from it.

To do so this paper is structured as follows: Chapter 2 briefly discusses the ecosystem created by Linked Data. Various traffic patterns and stakeholders are involved in the corresponding value creation process from raw data to Linked Data. Chapter 3 introduces the Linked Data Business Cube, an OLAP-inspired model to visualize the interdependencies between Linked Data assets, revenue models and stakeholders. Chapter 4 provides a practical case study about the Linked Data business model of the publishing house Wolters Kluwer. Chapter 5 closes this paper with a conclusion and reflexion on the commercial aspects of Linked Data.

2. The Linked Data Ecosystem – Related Work

Linked Data marks a transition from hierarchies to networks as an organisational principle for data and knowledge [11]. Hence, the primary value proposition of Linked Data is rooted in its modularity and connectivity to generate network effects at the data level [2]. By sharing the Resource Description Framework (RDF)[1] as a unified data model, Linked Data provides the infrastructure for publishing and repurposing of data on top of semantic interoperability [24]. In this paper we refer to the concept of business model as an *architecture of revenues* in a multi-stakeholder environment built around Linked Data assets, also commonly referred to as business ecosystem [31].

Taking the network characteristics of Linked Data into account we can identify three scenarios of traffic patterns in the utilization of Linked Data [3]:

Scenario 1: Internal Purposes: Enterprises make use of Linked Data principles to organize information within bounded organizational settings. This is especially relevant for organizations that have to deal with an increasing amount of dispersed databases, federated repositories and the legacy issues deriving from it. Linked Data is can be utilized to consolidate these infrastructures without necessarily disrupting existing systems and workflows.

Scenario 2: Inbound Purposes: Organizations use external data sources for purposes like content pooling or content enrichment. This trend is basically backed by the increasing availability of (open) data i.e. provided by governmental bodies, community projects like Wikipedia, Musicbrainz or Geonames and an increasing amount of commercial data providers like Socrata, Factual or Datamarket. Most services provide access to their data via an application programming interface (API), which can be used

[1] See also http://www.w3.org/RDF/, accessed July 4, 2014

either free of charge (according to the Terms of Trade) or as a paid service according to its service levels.

Scenario 3: Outbound Perspective: Organizations apply Linked Data principles to publish data on the web either as open data or via an API that allows the fine granular retrieval of data according to a user's needs. Linked Data Publishing allows an organization to become part of a Linked Data Cloud [24] and thus participate in the surrounding and maintaining ecosystem. Data publishing strategies often go hand in hand with the diversification of business models and require a good understanding of the licensing issues associated with it [18][19].

Latif et al. [15] propose a model that describes the value creation process underlying the various traffic patterns. The model distinguishes between various stakeholder roles an economic actor can take in the creation of Linked Data assets and various types of data and applications that are being created along the data transformation process. Herein, raw data – which is provided in any kind of Non-RDF format (i.e. XML, CSV, PDF, HTML etc.) – is being transformed into Linked Data. In the next step the Linked Data is being consumed and processed by a Linked Data application. Finally the end user consumes the human readable data via functionally extended applications and services. This pattern can be found in many Open Government Data projects as illustrated by Archer et al. [1].

Kinnari [13] extends this view with an orthogonal layer called "support services and consultation", stressing the fact that apart from the value creation process itself, Linked Data also creates an environment for added value services that transcends the pure transformation and consumption of data. Such services are usually provided by data brokers [7], who collect, clean visualize and resell available data for further processing and consumption.

For the time being it is difficult to estimate the cost-effectiveness of Linked Data but several inquiries indicate that depending on scale and scope of a Linked Data project the saving potential in the management und reutilization of data can be noteworthy [16] [8] [17]. Herein Linked Data is expected to reduce technological redundancies thus lowering maintenance costs, improving information access, reducing search and discovery efforts and provide opportunities for service and business diversification due to the higher granularity and increased connectivity of content and services [17].

3. Business Model Perspective on Linked Data – The Linked Data Business Cube

The Linked Data Business Cube (Fig. 1) provides an integrated view on a Linked Data business model. The X-axis lists the stakeholders involved in a business transaction. The Y-axis lists potential revenue models. And the Z-axis lists various Linked Data assets that occur along the content supply chain. In the following sections we will discuss these dimensions in more detail and illustrate how the Linked Data Business Cube changes its shape when mapping revenue models to stakeholders.

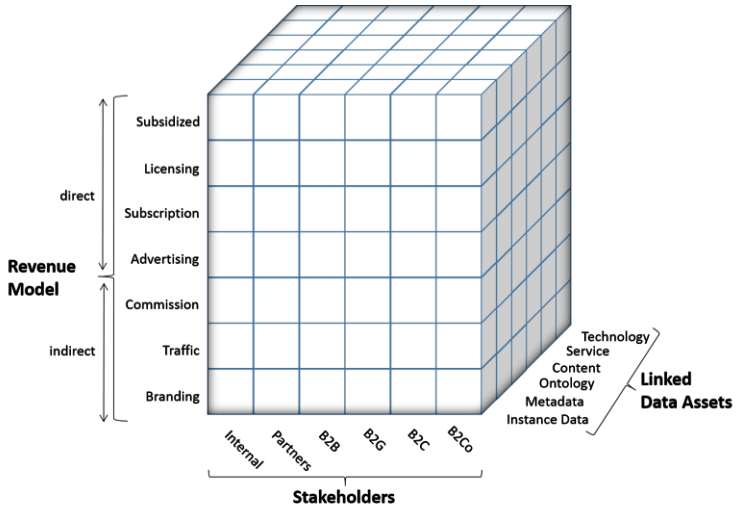

Fig. 1: Linked Data Business Cube

3.1 Linked Data Assets - From Instance Data to Technology

Linked Data is comprised of various asset types that emerge in the process of semantic data processing. Each asset type contributes to the value creation process and thus can be protected by intellectual property rights [18][19]. In the European Union the legal framework of property rights related to Linked Data comprises Copyright[2], Database Right[3], Competition Law[4] and Patent Law[5]. These appropriative legal regimes are being complemented by open access policies and according licensing instruments.[6] Creative Commons[7] allows to define tired licensing policies for the reuse of work protected by copyright. Open Data Commons[8] does the same thing for assets protected by database right. And open source licenses complement the patent regime as an alternative

[2] See also Directive 2001/29/EC. See also http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32001L0029, accessed April 20, 2014

[3] See also Directive 96/9/EC. See also http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31996L0009:EN:HTML, accessed April 20, 2014

[4] See also Consolidated versions of the Treaty on European Union and the Treaty on the Functioning of the European Union - Official Journal C 326, 26/10/2012 P. 0001 – 0390. See also http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:12012E/TXT&from=EN, accessed April 20, 2014

[5] See also http://www.epo.org/law-practice/legal-texts/html/epc/1973/e/ar52.html, accessed April 20, 2014

[6] A detailed discussion of licensing issues related to Linked Open Data is provided by [13].

[7] See also http://creativecommons.org/, accessed May 21, 2014

[8] See also http://opendatacommons.org/, accessed May 21, 2014

form of resource allocation and value generation in the production of software and services [10]. Table 1 illustrates the how Linked Data assets and intellectual property rights relate to each other.

	Copyright	**DB Right**	**Comp. Law**	**Patents**
Instance Data	NO	YES	PARTLY	NO
Metadata	NO	YES	YES	NO
Ontology	YES	YES	YES	NO
Content	YES	NO	YES	NO
Service	YES	NO	YES	PARTLY
Technology	YES	NO	YES	PARTLY

Table 1: Linked Data Assets and related Property Rights

Now let's look at the various Linked Data assets in more detail.

Instance Data: Instance Data are the concrete values that comprise a dataset in various formats to perform symbolic or mathematical operations on top of it. To be protected according to database right it must be encoded and electronically accessible.

Metadata: Metadata assets are basically all kinds of symbolic artefacts that provide information about data, objects and concepts. This information can be descriptive, structural or administrative. Metadata adds value to data sets by providing structure (i.e. schemas) and increasing the expressivity (i.e. controlled vocabularies) of a dataset. In case metadata is represented in digital code and made available via technical means for further processing, it is protected by Database Right and Competition Law.

Ontology: With reference to Gruber [27] "a common ontology defines the vocabulary with which queries and assertions are exchanged among agents" based on "ontological commitments to use the shared vocabulary in a coherent and consistent manner." In this sense ontologies function as integration layer for various datasets (comprised of instance data and metadata) and leverage interoperability from a syntactic to a semantic level for the purpose of knowledge sharing.[9] De-referenceable ontologies leverage network effects for datasets by referring to a common data model (i.e. RDF). In case the creation and maintenance of an ontology requires a significant amount of technical and intellectual effort this asset type protected by Copyright, Database Right and Competition Law.

Content: Content shall be understood as the output of an editorial workflow, in which information is being compiled into a technically consumable format. Hence, content is all kind of bundled information for consumption purposes. It is usually encoded as a document for asynchronous use or provided via a service for immediate use (i.e. as stream or API call). Applications Programming Interfaces (API) gain importance in the re-use of content. According to Knowles [14] "APIs enable the automated re-use of a given resource thus making it easier to interface to the proprietary sources of structured

[9] „Ontologies are being considered valuable to classifying web information in that they aid in enhancing interoperability – bringing together resources from multiple sources." [17, p. 657].

information." Content is generally protected under Copyright for the creative value of a literary work.

Service: A service is a technical means to provide access to a resource. It usually adds value to the resource itself by reducing the transactions costs for the consumption of the resource. Hence a service always carries an intrinsic value independent of the resource it grants access to. A service can address an end user (i.e. via a search engine or a recommendation service) or it can address an intermediate user that uses the service to add value to another resource (which can be a service too) i.e. via an API.

In case a service is represented in digital code and electronically accessible, it is protected by Copyright and Competition Law. A service can also be protected under Patent Law in case it fulfils the specific territorial requirements of the according patenting regime.

Technology: For reasons of completeness technology in terms of executable software code, frameworks or developments kits should be considered as another Linked Data asset. These tools support the creation, manipulation and consumption of Linked Data and are a valuable resource in nurturing a lively Linked Data ecosystem.

Software is usually protected by Copyright and Competition Law. In certain cases it is also protected by Patent Law.

3.2 Revenue Models for Linked Data Assets

In the following section we will look at various revenue models in the capitalisation of Linked Data assets. To do so we will refer to an adopted classification of Brinkner [4], who distinguishes between direct and indirect revenue models.[10] Direct compensation takes place where assets are directly being paid for, wherein indirect compensation takes place, where assets are being used to trigger revenue at a later stage in the consumption process.

Let there be the following assumptions: 1) Assets that are easily substitutable generate little incentives for direct revenues but can be used to trigger indirect revenues. This basically applies to instance data and metadata, so called *low-incentive assets*. 2) On the other side, assets that are difficult to imitate and substitute, i.e. in terms of competence and investments necessary to provide the service, carry the highest potential for direct revenues. This applies to assets like content, service and technology, so called *high-incentive assets*. 3) The higher the value proposition of an asset – in terms of added value – the higher the willingness to pay.

[10] In 2010 Scott Brinkner addressed the issue of Linked Data business models on his private blog chiefmartec.com. In his post he lists a handful of revenue models and discusses their relevance for various stakeholders. Brinkner's view is strongly marketing-oriented laying an emphasis on indirect revenue streams as a result of new marketing practices on top of Linked Data. Brinkner approaches the problem from a purely heuristic perspective. His classification lacks an empirical backing. Nevertheless it has been widely cited, i.e. by [13], [24] or [6] and discussed in the Linked Data community.

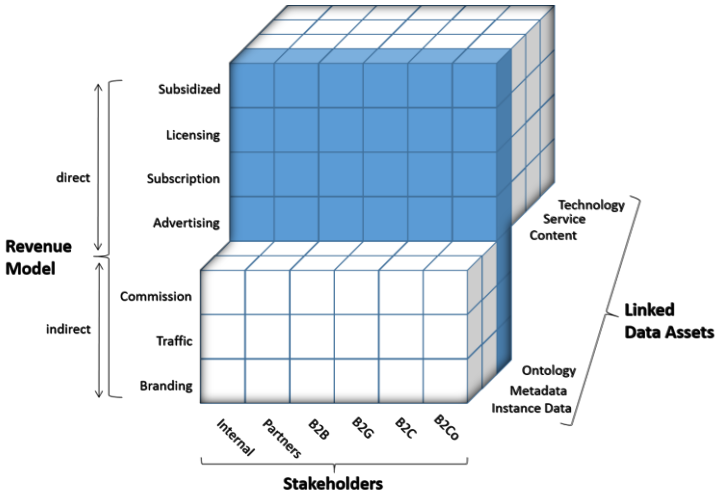

Fig. 2: Revenue Models for Linked Data

By mapping revenue models to Linked Data assets we can modify the Linked Data Business Cube as illustrated in Figure 2. The figure indicates that with increasing business value of an asset the opportunities to derive direct revenues rise.

Ontologies seem to function as a "mediating layer" between "low-incentive assets" and "high-incentive assets". This means that ontologies as a precondition for the provision and utilization of Linked Data can be capitalized in a variety of ways, depending on the business strategy of the Linked Data provider.

In the following we briefly discuss the various revenue models. It is important to note that each revenue model has specific merits and flaws and requires certain preconditions to work properly. Additionally revenue models often occur in combination as they are functionally complementary and can be used to address various stakeholders.

3.2.1 Direct Revenue Models

Subsidy model: An organization might be funded by the government, an NGO or by regulatory mandate to generate and publish Linked Data, e.g. for reuse in the public domain for commercial or non-commercial purposes.

Licensing model: An organization licenses Linked Data assets for commercial or non-commercial purposes, either with a standard license, specific terms of trade or individual agreements. With the increasing importance of community-driven value creation, i.e. as part of a developer program or open innovation policy, dual licensing has become an important strategy in the customer-centric provision of Linked Data assets. Dual licensing provides different user groups with different licenses, depending on the scale, scope and purpose of the assets' utilization. Hence, the reasonable combination of open and closed licenses becomes a crucial competence in the development of Linked Data business models.

Subscription model (incl. micropayments & Freemium): An organization charges for access to a Linked Data asset. This can be done on a subscription basis or pay-per basis. Access can be tired along various service level agreements and corresponding pricing schemes. Herein, the value proposition of the service lies not in the uniqueness of the retrieved data alone, but also in the convenience of accessing and reusing the data via a service. **S**ubscription models often come along in combination with Freemium models (versioned access) and/or micropayment models (transaction-related charges and accounting).

Advertising model (incl. paid inclusion & sponsorship): An organization sells the contact frequency of an information artefact, i.e. a site, a service or a search query a third party. The advertising model is most likely to be relevant for assets like content and services, who are accessible via public interfaces, but it is less likely for ontologies and technology. Nevertheless advertising information can represented in an ontology, hence the paid inclusion of information for advertising purposes might become a revenue stream given the fact that data-driven applications will have plenty of opportunity for contextual ads and sponsorships via data feeds and/or query results. Under circumstances of advertising it is crucial to provide detailed and reliable information on provenance and usage rights of the third party assets involved.

3.2.2 Indirect Revenue Models

Commission model (incl. affiliate model & added value model): An organization charges for an inclusion in its dataset or service o. The data provider then is being compensated in exchange for commissions on related transactions (i.e. views, clicks, sales). In commission models data providers use the service providers brand reputation to signal that their data is trustworthy and/or of high quality, and service providers benefit from the quantity of the data they offer to their customers.

An affiliate program can be seen as a typical expression of a commission model, where affiliate companies combine product links with data to earn commissions on related sales. Affiliates can also use each other's data to provide value added products, i.e. by incorporating free or bonus data as an enhanced feature to win customers for another product or service. In such a case various assets are being bundled to make the overall solution more valuable.

Traffic model: An organization publishes data to earn favourable positions in search engines and other directories to generate traffic, thus boosting visibility and ranking of sites in horizontal and vertical search engines. Additionally this strategy might have a positive impact on advertising revenues due to increased click rates. Brinkner refers to this as "data-enhanced search engine optimization". Initiatives like http://schema.org support such revenue models by providing normalized metadata schemas for the description of datasets and entities.

Branding model: An organization publishes data for branding purposes. This can be either proprietary data that has been opened up via an open license or third party data that is already available for such syndication purposes. According to Brinkner "data branding can use data — and the vocabularies that define and structure data — to position and promote a company's worldview and differentiation strategy."

3.3 Mapping Stakeholders to Revenue Models – A proposal

A Linked Data ecosystem is usually comprised of several stakeholders participating in the utilization of interlinked data. Stakeholders can be distinguished by the context in which Linked Data is being utilized. In the following we identify six prototypical stakeholder settings (views) and propose corresponding revenue models for further discussion. Fig. 3 illustrates possible dependencies between stakeholders and revenue models. The model is purely heuristic and open to discussion.

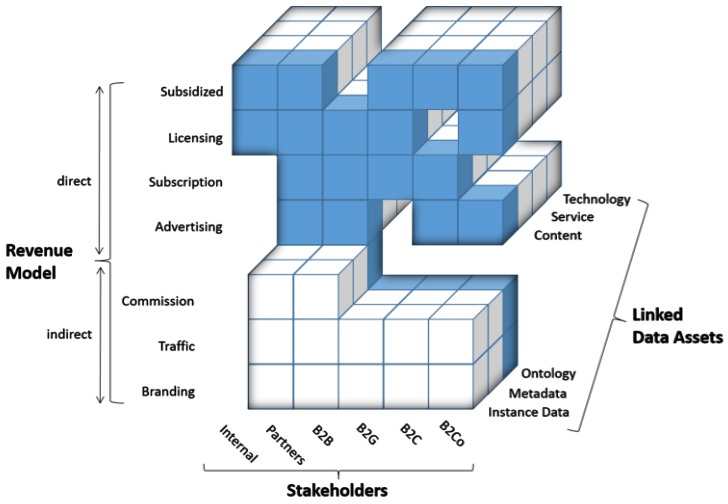

Fig. 3: Mapping Stakeholders to Revenue Models

3.3.1 Internal Use

In most cases the internal use of Linked Data assets is not relevant to generate a business. Nevertheless in certain cases it might be necessary to set up an internal licensing agreement or a subsidization policy if Linked Data assets are being shared between profit centres of the same enterprise or corporation.

3.3.2 Strategic Partnerships

Strategic partnerships are characterized by long-term agreement for sharing of physical and/or intellectual resources to reach a common objective [30]. In a strategic partnership two or more actors engage in a close functional and /or financial relationship that bind the partners structurally together for a certain period of time. In Linked data environments strategic partners can use each other's data to improve branding and traffic along their online channels. In a more advanced manner commissioning can be used for cross-promotion. More sophisticated resources – like ontologies, content, services or technology – can be combined to offer mutual products based on a subscription or licensing model, which is especially relevant for services or technology. And partners can contribute unique products to the public domain by subsidizing them.

3.3.3 The B2B View

Business-to-business (B2B) relationships shall be defined as transactions between firms characterized by (1) relatively large volumes (wholesale), (2) competitive and stable prices, (3) fast delivery times and, often, (4) on deferred payment basis.[11] No deeper functional or structural integration in the value creation process takes place.

Similar to the partnership scenario the B2B scenario allows a capitalization of Linked Data assets along all types of revenue models. But there exists a significant difference between these two stakeholder groups: while in the partnership scenario the focus of cooperation will most likely be on non-monetary forms of cooperation, in a B2B scenario partners will be eager to generate more direct revenues then indirect ones. Hence revenue models like advertising, subscription and licensing will be more relevant under B2B circumstances, whereas subsidies as a financing model occurs to be irrelevant (at least for the time the asset itself can claim exclusivity).

3.3.4 The B2G View

Business-to-government (B2G) relationships shall be defined as supplier-relationships between firms and governmental bodies for purposes like procurement, public corporate governance or intergovernmental communication [1].

The statements made in general to B2B also hold true in B2G environments. But differences exist with respect to authority-based revenue models like commissioning and advertising. These revenue models can cause conflicts with respect to credibility and integrity of their Linked Data assets, in case third party content (like ads or proprietary data) is being added to a government's original dataset – even if marked as such. Beside that all other revenue models can easily be applied to build direct and indirect revenue on top of governmental datasets.

3.3.5 The B2C View

Business-to-consumer (B2C) relationships shall be defined as acts of retail products to end-consumers on a monetary or non-monetary payment basis. Under B2C circumstances a variety of revenue models can be applied to generate value. Customers could use Linked Data assets to improve branding and traffic of their own sites. Herein companies could place advertising information within ontologies, content or services or give these resources away for free by subsidizing certain assets as part of a freemium model. Additionally customers can subscribe to Linked Data assets, which might be a reasonable revenue model for unique assets that are difficult to substitute. Licensing of assets that go beyond obligatory terms of trade and that include direct payments between two parties, is of minor importance in a B2C environment as end customers usually do not consume assets for further commercial purposes.

3.3.6 The B2Co View

The last view addresses the scenario of open innovation and shall be coined business-to-community (B2Co). These stakeholders are quite often ignored when talking about Linked Data in professional environments. Still, open source, crowdsourcing and developer programs are actively being maintained to co-create value within collaborative settings and API ecosystems. Instance data quality and the need for further development

[11] See also: http://www.businessdictionary.com/definition/business-to-business-B2B.html, accessed January 15, 2015

of metadata are prevalent phenomenon when using Linked Data. Data acquisition and cleansing often need human expertise and crowdsourcing initiatives can help to create added value on important data with low data consistency. Additionally communities are often influential forces in establishing de-facto standards by generating a critical mass for certain assets.

Hence providing adequate licensing models is crucial for leveraging a flourishing developer community. By sharing data organizations can improve branding and traffic. They can apply advertising, licensing and subsidization techniques to refinance their community activities. It is less likely to see subscription or commission models for community-derived assets as these revenue models usually coincide with high administrative efforts.

4. The Linked Data Business Case of Wolters Kluwer

The Linked Data Business Cube has been applied to analyse the utilization of Linked Data at the publishing house Wolters Kluwer Germany (WKG). Table 2 gives a snap shot of WKGs Linked Data business model. The model itself is open to evolve.

4.1 Traffic Patterns

Wolters Kluwer utilizes Linked Data along all three information traffic patterns. In the Inhouse-Scenario Linked Data is being used to support the content supply value and associated business processes. In the Inbound-Scenario Wolters Kluwer aggregates low-incentive assets (instance data, metadata and ontologies) from third party sources. In the Outbound-Scenario Wolters Kluwer publishes itself a broad range of low- and high-incentive assets to the public. Ontologies play an important role in the acquisition and publishing of data.

4.2 Stakeholders

To fulfil the purposes described above, WKG's Linked Data ecosystem entails several stakeholders: Generally, Linked Data is being applied to share data and streamline processes within the internal settings of WKG. But in the current situation the most important stakeholders are strategic partners with which Wolters Kluwer is sharing all kinds of Linked Data assets. These partners are basically affiliated companies of Wolters Kluwer within their corporate structure or closely associated content syndication partners. At the current stage of development B2B and B2C stakeholders play a minor role with respect to low-incentive assets like instance data, metadata and ontologies. These stakeholders are basically served with conventional assets like content and services.

With the introduction of Linked Data principles B2G and B2Co have gained significant importance for WKG's data publishing strategy. In the area of B2G WKG harvests

large amount of governmental data to fuel editorial processes. And by actively serving specific communities with low-incentive Linked Data assets WKG is nurturing a collaborative production environment for open innovation purposes. This also includes the advancement of Linked Data technologies and standards.

4.3 Revenue Models

WKG's main economic asset is content. Content assets serve multiple purposes in the value creation process and thus is being capitalized via various direct and indirect revenue models. In certain cases instance data is being licensed to interested parties. Metadata is being used for branding and advertising purposes, especially in the improvement of search engine marketing. Ontologies also play an important role for branding purposes (i.e. developer programs), but ontologies are also being sold via subscription. The latter aspect also holds true for services that are fuelled by Linked Data.

		Instance Data	Metadata	Ontology	Content	Service	Technology
Traffic Patterns							
	Inhouse	+	+	+	+	+	
	Inbound	+	+	++			
	Outbound	+	+	++	+	+	
Stakeholders							
	Internal	+	+	+	+	+	+
	Strat. Partners	++	++	++	++	++	++
	B2B				+	+	
	B2G	+	+	+	+	+	
	B2C				++	++	
	B2Co	+	+	+			+
Revenue Model							
	Branding		+	+	++		
	Traffic				++		
	Commission				++		
	Advertising		++		++		
	Subscription			++	++	++	
	Licensing	+			+		
	Subsidy			++			

Legend: ++ = very important; + = important

Table 2: The Linked Data Business Model of Wolters Kluwer Germany

5. Conclusion and Critique

This paper introduced an analytic framework to investigate the asset creation and commercialization of interlinked data. Summing up, Linked Data business models are diverse and context specific. Revenue models change in accordance to the various assets involved and the stakeholders who make use of them. Knowing these circumstances is crucial in establishing successful business models, but to do so it requires a holistic and interconnected understanding of the value creation process and the specific benefits and limitations Linked Data generates at each step of the value chain. This is sometimes

difficult to achieve. Technological legacies, division of labour among specialized business units, institutional arrangements that foster intra-company competition and last but not least the reactance to change in established organisational structures sometimes hinder the adoption of new business models.

Hence, the economic viability of Linked Data is equally a matter of technological feasibility, as a matter of organisational adaptability to new forms of data management. This becomes especially obvious when licensing issues are concerned not just in the reutilization of third-party datasets but also in the provision of Linked Data to the public or a specific business community be it for profit or for altruistic reasons. Developing licensing policies for Linked Data and reusing Linked Data according to the various involved licensing policies, poses a major challenge in the commercialization of Linked Data assets. Automatic clearing of machine-readable terms and conditions (with respect to permissions, prohibitions and obligations) in the compilation and marketing of new, federated datasets will be a necessary precondition, if the network dynamics of Linked Data shall unfold.

Heuristic models as the one described in this paper have their weaknesses. They leave plenty of room for improvement and empirical falsification and therefore should be seen as a proposal for further discussion and elaboration of the economic specificities that occur when data becomes a network good. If and how Linked Data will gain economic importance as a driver of service diversification and new business opportunities is open to observation and future research.

6. Acknowledgements

This work was partly funded by the project "LOD2 – Creating Knowledge out of Interlinked Data". Call: FP7-ICT-2009-5; Grant Agreement No: 257943 and the project "NoLDE – Network of Linked Data Excellence". Austrian Research Promotion Agency (FFG); Grant Agreement No: 3592880

7. References

[1] Archer, Phil; Dekkers, Max; Goedertier, Stijn; Loutas, Nikolaos (2013). Study on business models for Linked Open Government Data (BM4LOGD - SC6DI06692). Prepared for the ISA programme by PwC EU Services See also: http://ec.europa.eu/isa/documents/study-on-business-models-open-government_en.pdf, accessed May 10, 2014

[2] Berners-Lee, Tim (1998). Web Architecture from 50,000 feet. In: http://www.w3.org/DesignIssues/Architecture.html, accessed May 20. 2014

[3] Blumauer, Andreas (2014). Linked Data in Unternehmen. Methodische Grundlagen und Einsatzszenarien. In: Pellegrini, Tassilo; Sack, Harald; Auer, Sören (Eds.). Linked Enterprise Data. Berlin: Springer (to appear)

[4] Brinkner, Scott (2010). Business models for linked data and web 3.0. In: chiefmartec.com, 10.03.2010. See also: http://chiefmartec.com/2010/03/business-models-for-linked-data-and-web-30/, accessed April 10, 2014

[5] Cardoso, Jorge; Hepp, Martin; Lytras, Miltiadis (2008). The Semantic Web: Real-World Applications from Industry (Semantic Web and Beyond). New York: Springer

[6] Cobden, Marcus; Black, Jennifer; Gibbins, Nicholas; Carr, Les; Shadbolt, Nigel (2011). A Research Agenda for Linked Closed Data. In: Second International Workshop on Consuming Linked Data, Bonn: Ceur Workshop Proceedings. See also: http://ceur-ws.org/Vol-782/CobdenEtAl_COLD2011.pdf, accessed April 10, 2014

[7] Committee on Commerce, Science and Transportation (2013). A Review of the Data Broker Industry: Collection, Use and Sale of Consumer Data for Marketing Purposes. Staff Report for Chairman Rockefeller, December 18, 2013. See also: https://law.ku.edu/sites/law.ku.edu/files/docs/media_law/2014/Data_Broker_Industry_Senate_Report.PDF, accessed May 5, 2014

[8] Cranford, Steve (2009). Spinning a Data Web. In: Price Waterhouse Coopers (Ed.). Technology Forecast, Spring 2009. See also: http://www.pwc.com/us/en/technology-forecast/spring2009/index.jhtml, accessed September 20, 2012

[9] Dodds, Leigh; Davis, Ian (2009). MP Data SPARQL Editor. In: http://www.guardian.co.uk/open-platform/apps-mp-data-sparql-editor, accessed April 20, 2013

[10] Ghosh, Rishab Ayer et al. (2006). Economic impact of open source software on innovation and the competitiveness of the Information and Communication Technologies (ICT) sector in the EU. Final report prepared on November 20, 2006 (Contract ENTR/04/112.). See also http://ec.europa.eu/enterprise/sectors/ict/files/2006-11-20-flossimpact_en.pdf, accessed May 10, 2014

[11] Halford, Susan; Pope, Catherine; Weal, Mark (2012). Digital Futures? Sociological Challenges and Opportunities in the Emergent Semantic web. In: Sociology, 47/1

[12] Jentzsch, Anja (2014). Linked Open Data Cloud. In: Pellegrini, Tassilo; Sack, Harald; Auer, Sören (2014). Linked Enterprise Data. Berlin: Springer Verlag (to appear)

[13] Kinnari, Tomi (2013). Open data business models for media industry. A Finnish case study. Master Thesis, Department of Information and Service Economy, Aalto University. See also: http://epub.lib.aalto.fi/fi/ethesis/pdf/13166/hse_ethesis_13166.pdf, accessed March 29, 2014

[14] Knowles, Chris (2002). Intelligent agents without the hype: why they work best with structured content. In: Business Information Review, 19/4, p. 22-28

[15] Latif, Anif; Us Saeed, Anwar; Höfler, Patrick; Stocker, Alexander; Wagner, Claudia (2009). The Linked Data Value Chain: A Lightweight Model for Business Engineers. In: Proceedings of I-Semantics 2009, the 5th International Conference on Semantic Systems. Graz: Journal of Universal Computer Science, p. 568–577

[16] McHugh, Linda (2009). Measuring the Value of Metadata. White Paper: Baseline Consulting

[17] Mitchell, Ian; Wilson, Mark (2012). Linked Data. Connecting and exploiting big data. Fujitsu White Paper, March 2012. http://www.fujitsu.com/uk/Images/Linked-data-connecting-and-exploiting-big-data-%28v1.0%29.pdf, accessed September 12, 2012

[18] Pellegrini, Tassilo; Ermilov, Ivan (2013). Guide and Best Practices to Licensing Interlinked Data. Public Deliverable 7.4. EU-Project LOD 2. Grant Agreement No: 257943. See also http://svn.aksw.org/lod2/WP7/D7.4/public.pdf, accessed 03.01.2014

[19] Pellegrini, Tassilo (2014). Linked Data Licensing – Datenlizenzierung unter netzökonomischen Bedingungen. In: Schweighofer, Erich; Kummer, Franz; Hötzendorfer, Walter (Hg.). Transparenz. Tagungsband des 17. Internationalen Rechtsinformatik Symposium IRIS 2014. Wien: Verlag der Österreichischen Computergesellschaft, S. 159-168

[20] Rayfield, Jem (2012). Sports Refresh: Dynamic Semantic Publishing. In: http://www.bbc.co.uk/blogs/bbcinternet/2012/04/sports_dynamic_semantic.html, visited April 20, 2012

[21] Russom, Philip (2011). Big Data Analytics. TDWI Research Report, 4. Quarter 2011. http://tdwi.org/bdr-rpt.aspx (accessed September 14, 2012)

[22] Saumure, Kristie; Shiri, Ali (2008). Knowledge organization trends in library and information studies: a preliminary comparison of pre- and post-web eras. In: Journal of Information Science, 34/5, 2008, p. 651–666

[23] The Economist (2010). Data, data everywhere. A special report on managing information. http://www.emc.com/collateral/analyst-reports/ar-the-economist-data-data-everywhere.pdf (accessed March 30, 2013)

[24] Vafopoulos, Michalis N. (2011). A Framework for Linked Data Business Models. 15th Panhellenic Conference on Informatics (PCI) 2011. Available at SSRN: http://ssrn.com/abstract=1850365, accessed April 10, 2014

[25] Wood, David (Ed.) (2010). Linking Enterprise Data. New York: Springer

[26] Wood, David (Ed.) (2011). Linking Government Data. New York: Springer

[27] Gruber, Thomas R. (1993). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. In International Journal Human-Computer Studies 43, p.907-928.

[28] W3C (2012). Semantic Web Case Studies and Use Cases. In: http://www.w3.org/2001/sw/sweo/public/UseCases/, accessed January 10, 2015

[29] W3C (2014). Linked Data. In: http://www.w3.org/standards/semanticweb/data, accessed July 5, 2014

[30] Roberts, John (2004). The Modern Firm. Oxford: Oxford University Press

[31] Zott, Christoph; Amit, Raphael; Massa, Lorenzo (2011). The Business Model: Recent Developments and Future Research. In: Journal of Management, 37/4, pp. 1019-1042