Open and Linked Data in Business: Disruptive Marketing Mix 3.0
by

Milena Krumova
Technical University, Sofia

Lyudmila Paunova
Technical University, Sofia

Stefka Yotova
Technical University, Sofia

Abstract

This research aims to explain how the use of Open and Linked Data impacts the business practices. Building a business based on Open Data may seem counterintuitive, but new models such as Marketing 2.0/3.0 are emerging with greater frequency and demonstrate integration of Open Data into business operations in a useful, profitable manner.

The paper starts with state of the art about using Open Data in business. The next part explains Linked Data and then presents case studies about some frequently used Open Data in business practices. The paper discusses the opportunities for integration of Open and Linked Data towards sharing marketing information, and proposes a conceptual Disrupting Marketing Mix 3.0 Model, adopting the Business 3.0 approach.

Highlighting the increasing role of Linked Open Data and pointing out potential benefits and risks, the paper analyses smart innovative practices and is a source of innovation itself.

Keywords: Open Data, Linked Data, Business, Model, Marketing
Introduction
Open Data can be small, medium or big data - the latter is usually defined as very large datasets that require massive computing power to be processed. Open Data can come from variety of sources, often from government data and private firms (Bonina, 2013). On one hand, Open Data can become an instrument for breaking down information gaps across industries, allowing companies to share benchmarks and spread best practices that raise productivity. Blended with proprietary data sets, it can propel innovation and help organizations replace traditional and intuitive decision-making approaches with data-driven ones. On the other hand, Open-data analytics can also help uncover consumer preferences, allowing companies to improve new products and to uncover anomalies and needless variations. That can lead to more reliable processes, If actionable insights have been derived from the data. In recent years, the Web has evolved from a global information space of linked documents to one where both documents and data are linked. Underpinning this evolution is a set of best practices for publishing and connecting structured data on the Web known as Linked Data. Reasons for opening up to businesses is to stimulate economic growth by encouraging businesses to exploit the Open Data, to add value to business data and improve business decisions. Reasons for businesses to need different kinds of data are: increase fidelity and utility of own data, offer new data-enabled services to customers and satisfy journalistic interest (Deloitte, 2012).
Open Data in business

Open Data is accessible public data that people, companies, and organizations can use to launch new ventures, analyze patterns and trends, make data-driven decisions, and solve complex problems (Gurin, 2013). Similarly, Open Data Institute states that Open Data is data that anyone can access, use and share. Without a license, the Open Data can’t be reused. The license outlines that people who use the data must credit whoever is publishing it (this is called attribution) and people who mix the data with other data have to also release the results as Open Data (this is called share-alike). Open Data can be linked to, so that it can be easily shared and talked about, it is available in a standard, structured format, so that it can be easily processed. Open Data has guaranteed availability and consistency over time, so that others can rely on it and is traceable, through any processing, right back to where it originates, so others can work out whether to trust it (Open Data Institute, n.d.c).

Data is open if it satisfies both to be technically open: i) available in a machine-readable standard format, which means it can be retrieved and meaningfully processed by a computer application, and ii) legally open: explicitly licensed in a way that permits commercial and non-commercial use and re-use without restrictions. N.Shadbolt points to attention that Open Data offers the prospect of instant connectivity between partners, as in open supply chains, where businesses source form places they might never have considered or even suspected could be a source.
Open Data use
Open Data can reduce integration costs, improve transparency and harness the innovation of others” (Deloitte, 2012). C.M. Bonina defines Open Data as data released in open access format, generally coming from the government but also other public and non-public bodies. Open Data can be used by businesses, nonprofits, governments and citizens to launch new initiatives and ventures, analyze trends, make data-driven decisions, and solve complex problems in all sectors of the economy (Open Data Enterprise). Open Data is publicly available data – often released by governments, scientists, and occasionally private companies – that is made available for anyone to use, in a machine-readable format, free of charge. Considerable attention has been devoted to the economic potential of Open Data for businesses and other organizations, and it is now widely accepted that Open Data plays an important role in spurring innovation, growth, and job creation. From new business models to innovation in local governance, Open Data is being quickly adopted as a valuable resource at many levels (Govlab, 2014). The more data is opened, the more it can be used, reused, repurposed and built on - in combination with other data - for everyone’s benefit. As the economy and society become more knowledge-based, data are core assets, creating value in their own right and driving social and economic innovation, growth and development (Gruen, Houghton, & Tooth, 2014). Several benefits of the use of Open Data are identified, and consist of direct and indirect benefits. Direct benefits are monetised benefits that are realised in market transactions in the form of revenues and Gross Value Added, the number of jobs involved in producing a service or product, and cost savings. Indirect economic benefits are i.e. new goods and services, time savings for users of applications using Open Data, knowledge economy growth, increased efficiency in public services and growth of related markets (Carrara & San Chan, 2015). Businesses in all industries find relevant Open Data and use it to improve their products and services, such as leading organizations Google and Asos. Although businesses need to overcome the inevitable tension between the value they traditionally ascribe to proprietary data and value of opening up, making their data more accessible and combining it with other sources will dramatically increase its value to the economy (Deloitte, 2012). There are many uses of Open Data:

• Open Data possess a large potential economic value from its benefits, including increased efficiency, new products and services, and a consumer surplus;
• enhances big data’s impact by creating transparency, exposing variability, and enabling experimentation; helping companies to segment populations and thus to customize actions directed at them; replacing or supporting human decision making; and spurring innovative business models, products, and services;
• creates multiple business opportunities, such as the potential to raise productivity, to improve new products and services, and to enable entirely novel lines of business for both established companies and entrants;
• benefits consumers even more than businesses, by creating price and product transparency as well as new channels to provide feedback that improves the quality of goods and services.
Apart from the value and benefits, the open nature of the data is the reason for different challenges and requires caution, because using Open Data:
• entails business risks, including reputational issues related to the potential release of negative information; the potential consumer backlash from aggressive open-data use; and the inadvertent release of confidential information, such as benchmarking data
• requires governments to play a central role by developing and implementing policies to mitigate consumer and business concerns about the misuse of Open Data and to help set standards that will allow the potential economic and social benefits to materialize
• faces barriers, including privacy concerns and the need for legal and regulatory frameworks (Manyika et al., 2013).

Open Data and Master Data Management in business
The benefits of Open Data are diverse and range from improved efficiency of public administrations, economic growth in the private sector to social development. The kind of data might be concerning users, customers, products, suppliers, departments, geographies, sites, cost centers, and business units. In large organizations, this data is often held in many different places, with lots of overlap and redundancy, in many different formats, and with varying degrees of quality and means of access. Master Data Management (MDM) is the practice of identifying, cleaning, storing and most importantly, governing data. Its key concerns include managing change over time as organizational structures change, businesses merge, and business rules change; incorporating new sources of data; supplementing existing data with externally sourced data; addressing the needs of reporting, compliance, and business intelligence consumers; and versioning data as its values and schemas change. Graph databases don’t provide a full MDM solution; they are, however, ideally applied to the modeling, storing, and querying of hierarchies, master data metadata, and master data models. Such models include type definitions, constraints, relationships between entities, and the mappings between the model and the underlying source systems. A graph database’s structured yet schema-free data model provides for ad hoc, variable, and exceptional structures - schema anomalies that commonly arise when there are multiple redundant data sources - while at the same time allowing for the rapid evolution of the master data model in line with changing business needs (Robinson, Webber, & Eifrem, 2015).

Big Data in business
Big data is mainly characterized by the three V’s: Volume, Velocity and Variety (Kalcheva, 2015). It is part of the future web infrastructure. Where massive amounts of data are not only available, but also connected and identifiable via Uniform Resource Identifiers, big data is related to Linked Data, and they both become an integral part for achieving enterprise purposes. When creating a semantic layer over the big data initiative, it is of great importance to include the following elements:
• Flexible, universal data model based on industry standards: Using standard industry models with a semantic platform, allows for big data solution developers to quickly create industry or company specific solutions that can be used with big data stores and where the solutions can evolve as data needs evolve.
• Use of semantic RDF standards to make the data “self-describing”: By using semantic RDF standards, instance data and meaning (meta-data) travel together so that both humans and machines can understand and use the data.
• Graph representation and management of data: By using a graph representation, big data gets contextualized with entity and relationships that can be used for search and analysis.
• Service-Oriented Architecture (SOA) infrastructure: A SOA infrastructure over big data and existing data stores allows in run time to bring in data into the big data store as necessary.
• Post-ingestion data characterization: Big data is about collecting data without worrying about schemas and data descriptions but the problem is that usually the data never gets any sort of description so it stays of limited utility (Prasad, 2014).

Barriers for the use of Open Data
Publishing Linked Data on the public Web, which categorizes it as Linked Open Data, can potentially provide numerous benefits. However its adoption also entails a number of barriers which can be related to institutional, task complexity, use and participation, legislation, information quality and technical (Janssen, Charalabidis, & Zuiderwijk, 2012). According to the European Data Portal barriers can be identified in several areas, including political, legal, technical and financial issues (Nieuwenhuis, Steenbergen, & Carrara, 2016).

A lack of Open Data license
Datasets that serve as Open Data are required to protect the ability of the users to access, share and use the content. This purpose focuses on the need of a proper license and three legal requirements that involve protection of personal data and sensitive information, preservation of data owner’s rights and promoting the correct use of content (Open Data Institute, n.d.b).
A large content of datasets could not be classified as licensed due to the lack of clarity and understanding on the difference between copyright and licenses. Licenses are granted by an authority to allow a usage. Owning the copyright gives the exclusive right to own the created work, to give permission of use and distribution of resources by the copyright owner. The things that are copyrighted are sometimes referred to as “intellectual property”. Created content is meant to be distributed with a license that sets the guidelines for use - number or uses, bounds of use, length of time until the license expires. In case of “work for hire”, the employer holds the copyright, instead of the author or creator. The copyright owner happens to be a company such as a creative agency or its
client by contractual agreement. The creator retains “moral rights” to their work, including the right of attribution (Bushell, 2011).

**Data cleansing and fixing errors**
Another barrier regarding the use of Open Data is related to the incomplete, improperly formatted, duplicated or simply incorrect data. Common errors that appear in Open Data include: wrong date formats; multiple representations; duplicate records; redundant data; mixed numerical scales; mixed ranges and spelling errors (Open Data Institute, n.d.a).

The solution of these issues is found in data cleansing tools and approaches. Altering data in order to be accurate and correct could be achieved by various software and data storage architectures. They focus mainly on the review of data sets and protocols associated with any particular data storage technology (Techopedia, n.d.). In general, data cleaning can involves several phases: data analysis; definition of transformation workflow and mapping rules; verification; transformation; and backflow of cleaned data (Rahm & Do, 2000).

**Potential threat to privacy**
The potential threat to privacy is one of the risks towards opening up data sets. One of the main concerns is that if arguments for Open Data are applied as a ‘general rule’ without sensitivity to the kinds of data in question, there are some risks that privacy rights might be violated. At the same time reasons behind ‘data protection’, or ‘protecting privacy’ might be actually excuses not to release data, or to only release data in general forms that don’t permit detailed analysis (Hand, 2012).

**Linked Data in business**
The idea of Linked Data is a descendant of the semantic web, whereas the idea of Open Data is built on the concept of a social web (Bauer & Kaltenböck, 2012). The terms Open Data and Linked Data are often used interchangeably and/or combined as in the term Linked Open Data. Open Data is data that has been explicitly published under an open licence. Not all Linked Data is Open Data either because the technical linkage is taking place behind an organisation firewall or because the data being used was not explicitly intended for open use (Jisc, 2013). Linked Data is intended for access by both humans and machines. Linked Data uses the RDF family of standards for data interchange (e.g., RDF/XML, RDFa, Turtle) and query (SPARQL). If Linked Data is published on the public Web, it is generally called Linked Open Data (Hyland, Atemezing, Pendleton, & Srivastava, 2013).

The Linked Data principles were introduced by Tim Berners Lee, he outlined four rules for making human or machine-readable links for the exploration of web of data. The first rule refers to the usage of Uniform Resource Identifier URI for identification of items. The second given rule specifies that only HTTP URIs are meant, so that people can look at them and these can be found by the standard established Domain Name Space (DNS).
The third rule was formulated for the purpose of providing additional useful information for the items defined by URI. The information should be denoted in a standard format, such as Resource Description Framework RDF, in form of the RDF/XML or an alternative serialization (N3, Turtle). The fourth rule concerns providing the linkage of such described items with other related items (data), so that the related information on the Web can be discovered more easily (Robak, Franczyk, & Robak, 2012). Tim Berners-Lee, the inventor of the Web and initiator of the Linked Data project, suggested a 5 star deployment scheme for Linked Open Data. The 5 Star Linked Data system is cumulative. Each additional star presumes the data meets the criteria of the previous step(s). 5 Star Linked Open Data includes an Open License (allows commercial reuse) and assumes publications on the public Web (Hyland et al., 2013).

- Publish data on the Web in any format accompanied by an explicit Open License.
- Publish structured data on the Web in a machine-readable format.
- Publish structured data on the Web in a documented, non-proprietary data format.
- Publish structured data on the Web as RDF.
- In your RDF, have the identifiers be links to useful data sources.

There are generic Linked Data browsers which allow users to start browsing in one data source and then navigate along links into related data sources. There are Linked Data search engines that crawl the Web of Data by following links between data sources and provide expressive query capabilities over aggregated data, similar to how a local database is queried today. The Web of Data also opens up new possibilities for domain-specific applications. Unlike Web 2.0 mashups which work against a fixed set of data sources, Linked Data applications operate on top of an unbound, global data space. This enables them to deliver more complete answers as new data sources appear on the Web (Bizer, Heath, & Berners-Lee, 2009). Linked Open Data (LOD) facilitates innovation and knowledge creation from interLinked Data, it is an important mechanism for information management and integration. There are two equally important viewpoints to LOD: publishing and consuming. Publishing Linked Open Data LOD provides a powerful mechanism for sharing data and information along with metadata and the respective data models for efficient re-use. Consuming LOD enables to integrate and provide high quality information and data collections to mix data and third party information. These enriched data collections can act as single points of access for a specific domain in the form of a LOD portal and as an internal or Open Data warehouse system that enables better decision making, disaster management, knowledge management and/or market intelligence solutions (Bauer & Kaltenböck, 2012).

The huge success and widespread adoption of the Linked Data approach has led to the availability of vast amounts of public data such as DBpedia, WordNet RDF or the data.gov.uk initiative, ranging from domain-specific expert vocabularies to, for instance, data about cultural heritage. More recently, these approaches started to get adopted by education institutions. This has led to the creation of an embryonic “Web of Educational Data” including institutions such as the Open University (UK) or the National Research
Council (Italy), as well as Linked Data about publicly available educational resources, such as themEducator – Linked Educational Resources. Initiatives such as LinkedEducation.org, LinkedUniversities.org and LinkedUp have provided first efforts to bring together people and works in this area (Linked Education, 2014).

The use of Open Data in business and marketing: Case studies
Existing models prove that Open Data begins to result in social improvement when being managed skillfully and purposefully. Innovative entrepreneurs can turn Open Data into new applications. A good example in the Netherlands is OfficeRank.nl, a website that provides information on the quality and cost of real estate. The site has been compiled using data from a variety of sources, including the land registry (Algemene Rekenkamer, 2015). To exploit the value of Open Data, to maximize the benefits, and to enable creation of innovative Open Data products and services, Open Data-driven Organizations should develop and implement Open Data Business Model before starting their business. This is required to ensure that the produced Open Data products and services generate necessary value proposition and meet the needs and demand of the customers/users and eventually generate substantial revenue (Ahmadi-Zeleti, 2016).

Business models using Open Data
Jeni Tennison categorizes two types of business models. Businesses that publish (but do not sell) Open Data and Businesses built on top of using Open Data.

A. Businesses that publish (but do not sell) Open Data.
- Freemium: In this model, the bulk data of open addresses could be made available freely.
- Cross-subsidy: In this approach, the data would be available, and the opportunities to generate income would come from providing extra services.
- Network: In this business model, value is created by generating a network effect around the core business interest, which may not be the Open Data itself.

B. Businesses built on top of using Open Data.
Development Seed Builds solutions for development, public health and citizen democracy challenges by creating open source tools and utilizing Open Data.
- Enigma.io Open Data platform with advanced discovery and search functions.
- Mapbox Enables users to design and publish maps based on crowdsourced OpenStreetMap data.
- Open Bank Project Creates an open source API for use by banks.
- OpenDataSoft Provides an Open Data publishing platform so that cities, governments, utilities and companies can publish their own data portal for internal and public use.
- Snips Predictive modeling for smart cities (Verhulst, 2014).
The innovative startups demonstrate that Open Data implementation could be applied to different fields with a variety of purposes. What they have in common is the contribution to the development of an advanced society. The goal of adding value with improved
content is directly related to the essence of the Marketing Mix 3.0 and the idea behind recent innovations in this field.

**Open Data and web: Marketing Mix 3.0**

Marketing innovations and the implementation of a new marketing method involving significant changes in product design or packaging, product placement, product promotion or pricing (UNESCO Institute for Statistic, 2009). The plan to incorporate the advances in marketing science, technology or engineering to increase the effectiveness and efficiency of marketing, to gain competitive advantage and increase shareholder value (Gaskin, 2013).

The changes in marketing are based on the influence of web technologies, which evolve rapidly for the last decades, from Web 1.0 to Web 3.0, and currently research show that show that Web 6.0 is emerging. Concepts such as Web 4.0 and Web 5.0 are still in development, but Web 6.0 is already in early stages. Web 4.0 is a concept based on interactions between humans and machines in symbiosis (Choudhury, 2014) and Web 5.0 is referred as web of Decentralized smart communicator (Patel, 2013). The next step is Web 6.0 (Khanzode & Sarode, 2016). Web 1.0 is defined as web of information connections, world wide web, anonymous users download content. Web 2.0 is web of people connections, social web, registered users download and upload content. Web 3.0 is the web of knowledge connections, semantic web, authorized users assess and improve content – a defined profile in social networks (Polańska, 2014). This huge amount of social data attracts researchers who want to use it to learn more about user preferences and interests, and enhance recommendation and personalization systems (Plumbaum, De Luca, & Albayrak, 2011). Application of Semantic Web technologies to the Social Web is leading towards the "Social Semantic Web" (sometimes also called "Web 3.0"), forming a network of interlinked and semantically-rich content and knowledge (Breslin, Passant, & Decker, 2009). The impact of web on every business functions including marketing, which purposes are mainly to (Costantino, 2011):

- Build awareness, credibility and trust with your preferred prospect or customer;
- Facilitate the decision making process of your preferred prospect or customer;
- Lower the risk for your preferred prospect or customer to take the next step in the buying process.

Since traditional marketing techniques are no longer able to completely grasp modern markets, the implementation of Marketing 3.0 is expected to result in an improved marketing strategy, which not only enables long-term financial performance, but also environmental improvements (Warrink, 2015).
<table>
<thead>
<tr>
<th>Marketing Mix 1.0</th>
<th>Marketing Mix 2.0</th>
<th>Marketing Mix 3.0 (a dozen M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Product - Products meant to be sold to mass buyers</td>
<td>- Product - Products that fulfil individual needs</td>
<td>- Product - Products that fulfil individual and societal desires; Co-creation</td>
</tr>
<tr>
<td>- Price - A price that complements the other three Mix elements</td>
<td>- Price - High price competition</td>
<td>- Price - Initial premium prices - eventually compensated by cost savings from waste reduction and government subsidies, Price advantage for reputable brands</td>
</tr>
<tr>
<td>- Place - Direct mail, TV ads</td>
<td>- Place - Online sales</td>
<td>- Place - Further shift towards online sales</td>
</tr>
<tr>
<td>- Promotion - One-to-many transactions</td>
<td>- Promotion - Individual-value, Mass communication</td>
<td>- Promotion - Vertical communication, Educate about sustainable behavior</td>
</tr>
<tr>
<td>- Physical Evidence - Build a Brand</td>
<td>- Process - Process Efficiency</td>
<td>- Process - Integrate sustainability throughout the supply chain, Integrate sustainability internally, Implement EMS</td>
</tr>
<tr>
<td>- „Customer 2.0“ has specific expectations and attitude. Online customers are divided in four levels of activity: adopter, collaborative, content creator, social</td>
<td>- People - Fair treatment of stakeholders</td>
<td>- People - Share sustainable corporate vision among employees, High ethical norms</td>
</tr>
<tr>
<td></td>
<td>- Physical Evidence - Build a Brand</td>
<td>- Physical Evidence - Build a character for the brand, Accept Corporate Social Responsibility</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Presence in Social Media - Use Facebook, Twitter, Instagram, Pinterest, Google+ to gain recognition</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Pull marketing content – Create valuable content to attract clearly-defined audiences</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Per click campaign - Quality Score, Text Ad Optimization, Click-Through Rate</td>
</tr>
</tbody>
</table>

Table 2. Marketing Mix 1.0, Marketing Mix 2.0 and Marketing Mix 3.0 (Warrink, 2015)
| Prioritize the Open Datasets – Create new opportunities for entrepreneurs, Help established businesses operate strategically, Give investors valuable information | Push the use of Open Linked Data - Incorporate up-to-date, relevant information to push notifications at the time of send |

**Marketing and market intelligence**

The development of Marketing Mix 3.0 affects all aspects in the marketing field, such as marketing strategy, marketing planning, promotions, public relations, and sales, etc. Traditionally, the gathered and analyzed information as a result of keeping track of the competition has been termed "market intelligence." As societal and environmental needs are forming the core of the 7P formula, marketing intelligence expands its targets, focusing no longer only on the competition. Market intelligence describes the set of activities that provide a company with a view of a market using existing sources of information to understand what is happening in a market place, what the issues are, what competitors are doing, what customers or consumers are doing (e.g. social media) and what the likely market potential is for new products or services based on previous activities and responses. Market intelligence areas depend on the source of the data:

- Based on external data;
- Social media monitoring;
- Based on internal data. (Dobney, n.d.).

**Social media marketing**

Social media marketing refers to the process of gaining traffic or attention through social media sites (Marketing Land. n.d.): Facebook; Instagram; Twitter; Pinteres; Google+; Linkedin; YouTube; etc. Social media intelligence can help firms track brand health and market structure and can even provide a leading indicator of shifts in consumer sentiment (Marketing Science Institute, 2014). McKinsey Global Institute (MGI) finds that twice as much potential value lies in using social tools to enhance communications, knowledge sharing, and collaboration within and across enterprises (Chui et al., 2012). Social media is where the distribution changes drastically. About 73 percent of participants are planning to increase their budget, with another 14 percent planning to keep theirs the same. That's 87 percent of people using it to some degree of satisfaction, with only 4 percent either refusing to adopt it or actively planning to decrease their budget (AudienceBloom, 2016).
Open and Linked Data in conceptual Disrupting Marketing Mix 3.0 Model

The core of the proposed conceptual Disruptive Marketing Mix 3.0 model is the impact of the use of Open Data, Linked Data and web 3.0 technologies and tools on business. On one hand the use of Open Data sets can leads to a lot of benefits, but the ability to extract meaningful information from it. On the other hand, the adoption of the Linked Data best practices can lead to the extension of the Web with a global data space connecting data from diverse domains. On Figure 1 are shown the twelve areas of Marketing Mix 3.0 and the relation with Open and Linked Data.

Figure 1. Conceptual Disruptive Marketing Mix 3.0 Model

The use of web of data enables new types of business and marketing. The implementation of the proposed model can impact the business performance and to improve the marketing processes. The key characteristic of the proposed model is the prioritizing the data. Similarly, the use of the diversity of Web 3.0 and social applications can allow organization to gain competitive and operational advantage by leveraging information about the connections between people, together with discrete information about individuals, to facilitate collaboration and flow of information, and predict behavior.
Conclusion
Business model innovation is becoming a key topic these days as business leaders increasingly recognize that disruptive innovation requires not only innovative products, but also fundamentally new business approach. Open Data is seen as an asset and putting specific values on widely shared data is a worthwhile goal for business organizations. At the same time, the process of using particular Open Datasets requires a specific approach. As the business needs and data sources change, the IT solutions no longer work and new solutions must be built. In this research we propose a model for conceptual Disruptive Marketing Mix 3.0 which is based on Disruptive technologies and Open Linked Data and how the data can be explored. The 12 key areas of the model are a base for future research and analyses to identify what it the impact of the use of Open Data and Linked Data in practice and how this model can be boost towards improvements of efficiency and efficacy in business. The future research steps will be to discover the extent to which the business is using Open Linked Data in everyday operations and what are the priorities of the Open Datasets.
References


