CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Understanding RegTech for digital regulatory compliance
Authors	Butler, Tom;O'Brien, Leona
Publication date	2018-12-07
Original Citation	Butler T., O'Brien L. (2019) Understanding RegTech for Digital Regulatory Compliance. In: Lynn T., Mooney J., Rosati P., Cummins M. (eds) Disrupting Finance. Palgrave Studies in Digital Business & Enabling Technologies. Palgrave Pivot, Cham. pp. 85-102 DOI: 10.1007/978-3-030-02330-0_6
Type of publication	Book chapter
Link to publisher's version	https://link.springer.com/content/pdf/10.1007%2F978-3-030-02330-0_6.pdf - 10.1007/978-3-030-02330-0_6
Rights	© 2019 The Authors. This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it. The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. - https://creativecommons.org/licenses/by/4.0/
Download date	2026-11-04 04:38:22
Item downloaded from	https://hdl.handle.net/10468/8086



University College Cork, Ireland
Coláiste na hOllscoile Corcaigh



CHAPTER 6

Understanding RegTech for Digital Regulatory Compliance

Tom Butler and Leona O'Brien

Abstract This chapter explores the promise and potential of Regulatory Technologies (RegTech), a new and vital dimension to FinTech. It draws on the findings and outcomes of a five-year research programme to highlight the role that RegTech can play in making regulatory compliance more efficient and effective. The chapter presents research on the Bank of England/Financial Conduct Authority (FCA) RegTech Sprint initiative, whose objective was to demonstrate how straight-through processing of regulations and regulatory compliance reporting using semantically enabled applications can be made possible by RegTech. The chapter notes that the full benefits of RegTech will only materialise if the pitfalls of a fragmented Tower of Babel approach are avoided. Semantic standards, we argue, are the key to all this.

Keywords RegTech · FinTech · Semantic technologies · Standards

T. Butler (✉) · L. O'Brien
BIS, University College Cork, Cork, Ireland
e-mail: TButler@ucc.ie

© The Author(s) 2019

T. Lynn et al. (eds.), *Disrupting Finance*, Palgrave Studies in Digital Business & Enabling Technologies,
https://doi.org/10.1007/978-3-030-02330-0_6

6.1 Introduction

RegTech is information technology (IT) that (a) helps firms manage regulatory requirements and compliance imperatives by identifying the impacts of regulatory provisions on business models, products and services, functional activities, policies, operational procedures and controls; (b) enables compliant business systems and data; (c) helps control and manage regulatory, financial and non-financial risks; and (d) performs regulatory compliance reporting. In reference to the previous generation of RegTech (and FinTech) Law Professor Kenneth Bamberger points out that:

> While these technology systems offer powerful compliance tools, they also pose real perils. They permit computer programmers to interpret legal requirements; they mask the uncertainty of the very hazards with which policy makers are concerned; they skew decision-making through an "automation bias" that privileges personal self-interest over sound judgment; and their lack of transparency thwarts oversight and accountability. These phenomena played a critical role in the recent financial crisis. (Bamberger 2009, pp. 669–670)

As Bamberger notes, in the rush to embrace new technologies, the downside risks are either ignored, played down, or transferred. This is apparently so with RegTech, as Patrick Armstrong, Senior Officer for Financial Innovation at the European Securities and Markets Authority (ESMA), warned "it is not without risks" and financial institutions must take care to "delegate tasks, not responsibility…for compliance and risk management activities. Instead the ultimate responsibility remains with the regulated financial institution" (McNulty 2017). There is a dearth of IS research in this new important area of study; however, research in other disciplines is also deficient. Few academic researchers have adopted a critical stance (Packin 2018), while others fail to question the risks associated with this new paradigm (Arner et al. 2016), including the real possibility of creating a digital *Tower of Babel* (Butler 2017). This is, perhaps, the major issue confronting the financial industry.

The industry has, according to Andrew Haldane, created a *Tower of Babel*, which refers to the absence of a "common language" in the financial industry, and the existence of heterogeneous terms and concepts to describe similar business objects, processes, and products (Haldane 2012). This problem permeates the industry down to individual financial

institutions, where products, concepts, and terms have different meanings in and across business functions and communities of practice. The emergence of FinTech and RegTech will do little to solve fundamental problems if the industry ends up with a digital *Tower of Babel* that simply digitises the status quo. The essential issue of arriving at shared business and regulatory terminological dictionaries, thesauri, and taxonomies is a huge challenge and a significant obstacle for RegTech. Only then will the issue of *semantic interoperability* be addressed—that is, the capacity of information systems to exchange data with unambiguous, shared meaning. "Semantic interoperability ensures that these exchanges make sense—that the requester and the provider have a common understanding of the "meanings" of the requested services and data" (Heiler 1995, p. 271). However, even if a "common language" did exist in a financial institution, we are still left with what Bamberger (2009) termed the problem of "translation"—that is the gap between the meanings accorded to business concepts by business professionals and the intended behaviours of automated processes, and the meanings recorded by systems analysts and software engineers in data stores and the behaviours embedded in software code. Bamberger demonstrated how the "translation problem" resulted in financial, risk and compliance applications and systems that masked risk, led users into a false sense of security, and provided erroneous signals to business managers resulting in poor decision-making on key issues.

The remainder of this paper is structured as follows. First, we discuss the business drivers for the adoption of RegTech, which incorporates a regulatory perspective. Then we present our empirical case study. We finish by discussing the implications of our research and offer concluding comments.

6.2 Business Drivers of RegTech

The costs of regulatory compliance for the financial industry represent the primary drivers for RegTech adoption. Research published by *The Trade* indicates that banks spent over $100 billion on regulatory compliance in 2016 alone and this cost is rising (McDowell 2017). Bain & Co. estimates that governance risk and compliance (GRC) expenditure accounts for 15–20% of "run the bank cost" and 40% of "change the bank costs" (Memminger et al. 2016). Looking at specific regulations, Dodd Frank has cost $36 billion to date, while MiFID II has cost a €2.5 billion

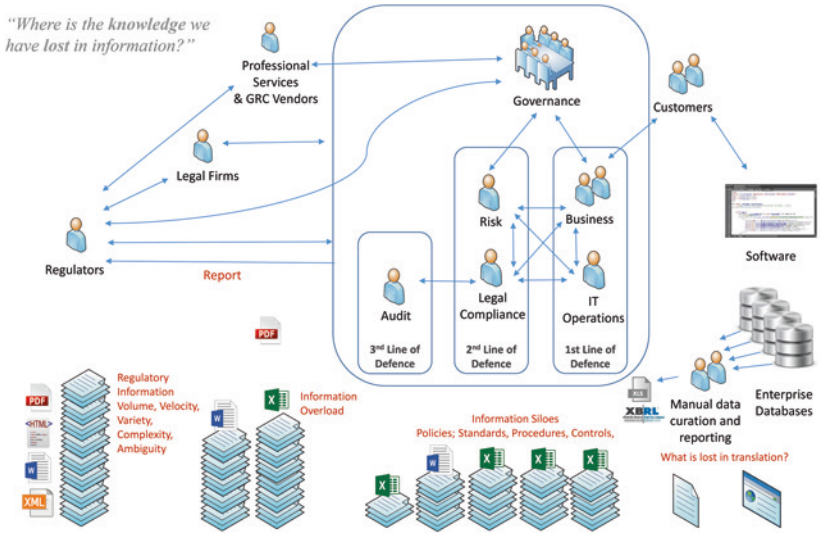


Fig. 6.1 Information overload, complexity, silos, and loss

to date—the latter cost is sure to rise significantly. Thus, given the existing trend, it is estimated that the cost of regulatory compliance will rise from 4 to 10% of revenue in financial institutions by 2022.

At a practical level, it is the volume of regulations that is driving costs. Take, for example, that over 50,000 regulations were published between 2009 and 2012 in the G20. Furthermore, over 50,000 regulatory updates were published in 2015 alone, 100% up on 2012. The scale of the paper mountain for firms is breath-taking: The FCA Handbook stands over 2 metres high; the US Dodd Frank Act has generated over 22,000 pages of provisions; the EU's MiFID II has approximately 30,000 pages of related texts in approximately 1.5 metres paragraphs. Each week sees an average of 45 new regulatory-related documents issued (JWG 2017). Thus, the volume, variety, velocity, and complexity of regulation is challenging firms. Figure 6.1 captures graphically the current approach to regulatory compliance and reporting. Information overload, multiple communication paths, information siloes, and manual data curation all increase the risk of information loss and reduced empirical fidelity while driving significant cost.

Another driver is the relative complexity of financial institutions in terms of their business models, legal entity structures, processes, products, services, and markets served. Monitoring, interpreting, and complying with the current and planned regulations is a challenge, even for the largest banks. For smaller firms the costs and complexity may become prohibitive (Walker 2018). The move from people-based solutions to RegTech is reflected in the fact that the aforementioned costs are being expended on consultants, professional services, and IT vendors (Marenzi 2017).

The final drivers are information and data related. Unpublished research from a 5-year-long empirical study by the authors on the issue, involving UK and US regulators and financial institutions, indicates that financial institutions are challenged by their inability to understand:

1. Regulatory requirements and compliance imperatives;
2. The impacts of such regulations on functional activities, policies, and procedures;
3. The changes that are required to business processes and activities;
4. The risks associated with financial products and related business models;
5. The implications for IT systems in terms of data governance and analytics; and
6. How to meet consumer needs, while protecting their rights.

It was in this context that RegTech was first identified as a separate, but emerging industry sector in the financial industry, in the UK Treasury's 2015 Budget Report and subsequently explored comprehensively in the UK's Government Organisation for Science (Walport 2015). UK regulators took note of this new perspective on technology-based solutions for the myriad of problems facing the financial industry and have responded accordingly. In 2016, for example, and in the context of the FCA's Innovation Hub and its Project Innovate, the FCA identified a number of candidate FinTech and RegTech solutions and how they might be used.

Given the significant challenges facing both regulators and the regulated, Christopher Woolard, Director of Strategy and Competition at the FCA, identified several use cases for and capabilities of RegTech (Woolard 2016) viz.

1. First, making the business of complying with reporting requirements simpler.
2. Second, technology that drives efficiencies in regulatory compliance by seeking to close the gap between the intention of regulatory requirements and the subsequent interpretation and implementation within firms.
3. Third, technology that simplifies and assists firms in managing and exploiting their existing data, supporting better decision-making and finding those who are not playing by the rules easier.
4. Finally, technologies and innovations that allow regulation and compliance processes to be delivered differently and more efficiently.

The above conceptualisation indicates a role for several new technologies including artificial intelligence (AI), blockchain/distributed ledger technologies (DLT), machine learning, natural language processing (NLP), and data analytics. However, it was clear from several presentations at the European Central Bank Data Standards for Granular Data Conference 2017 that the *sine qua non* for the success of RegTech would be the use of open standards and semantic technologies (Palmer 2017). A standards-based approach would be necessary to address the core issues of the translation and *Tower of Babel* problems (Bamberger 2009; Haldane 2012; Butler 2017).

6.3 RegTech in Focus: Digital Regulatory Reporting

The need for an industry-wide standards-based approach to regulatory compliance and reporting, articulated in the position paper of Butler (2017) on open standards for RegTech, found purchase with both UK regulators and the financial industry. The Bank of England and the FCA subsequently hypothesised that standards-based RegTech could help automate, and make more efficient and cost-effective, the task of regulatory reporting by financial institutions. To confirm their hypothesis, the UK regulators instituted a Technology Sprint—in this case the RegTech Sprint.

The remainder of this chapter presents a short case study of the RegTech Sprint. This was undertaken by the Bank of England and the

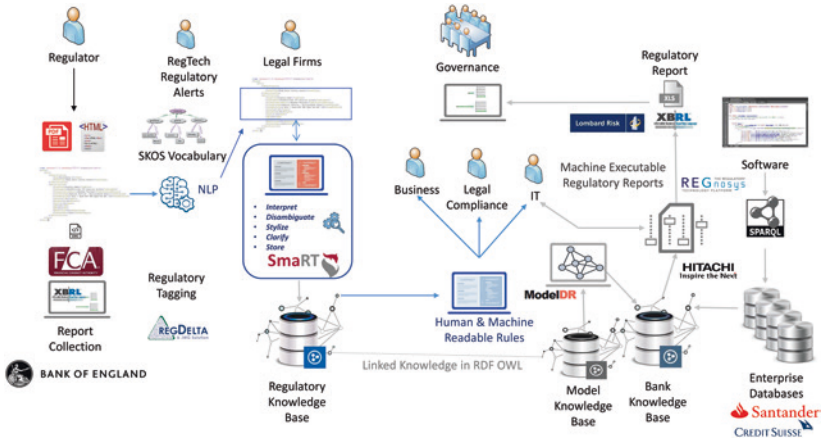


Fig. 6.2 Digital Regulatory Reporting

FCA in conjunction with over 50 participants from across the financial industry in the UK. The Sprint took place during the last two weeks of November 2017. Participants included regulators from the FCA, Bank of England, industry firms Santander, Credit Suisse, Hitachi Vantara, Lombard Risk, Model Drivers, Regnosys, JWG-RegDelta, Governor Software, law firms Linklaters and Burgess Salmon, academics from Yale, and the GRC Technology Centre at University College Cork.

The primary objective of the RegTech Sprint was to provide a Proof of Concept (PoC) that demonstrated the feasibility of the straight-through processing of regulations and semi-automated regulatory reporting—this process is termed Digital Regulatory Reporting. Figure 6.2 presents the key activities that realised this objective.

The first step in Digital Regulatory Reporting is to digitise the regulatory provisions. As indicated in Fig. 6.2, the FCA currently publishes its Handbook of Regulations in HTML and PDF formats. Key concepts are linked using Hypertext. In the case of the FCA, specific FCA concepts are defined in the Handbook Glossary. In its current form, the Handbook provisions and rules are not readily machine-readable or machine-executable. The first part of the PoC was to investigate how AI could be employed to process regulatory provisions and provide Regulatory Alerts.

6.3.1 Phase 1: Digital Regulatory Alerts

RegTech vendor RegDelta has developed taxonomies of regulatory topics using the World Wide Web Consortium (W3C)'s Simple Knowledge Organisation System (SKOS)[1] and AI to semantically tag regulatory provisions to indicate their scope and application so alerts could be generated. SKOS is based on the W3C's Resource Description Framework (RDF),[2] and is one of the three foundational Semantic Web technologies, the other two being SPARQL and the Web Ontology Language (OWL).[3] SKOS enables organisation to transcend the limitations of business glossaries to create taxonomies and thesauri that are both human- and machine-readable. Semantic tagging of regulations is but the first step in the regulatory compliance process and the SKOS namespace helps address the *Tower of Babel* problem. This is an example of the straightforward application of AI and semantic technologies to help manage the volume and complexity of regulations by having a machine answer the *what* and *which* questions. That is, *what* are the themes in regulatory provisions and *which* activities and products do they target. While this process helps partially digitalise regulatory provisions, Sprint participants recognised that the source regulations would have to be redrafted and captured in RDF, if the objective was to be achieved. This task fells to the SmaRT application.

6.3.2 Phase 2: Making Regulations Digital

The core semantic technologies in SmaRT are based on W3C and industry standard semantic technologies. SmaRT applies the Semantics of Business Vocabulary and Business Rules (SBVR) standard proposed by Object Management Group. SBVR enables business subject matter experts to capture and express their vocabularies and rules in a systematic way according to the precepts of first-order deontic-alethic logic. The SmaRT application applies its Mercury implementation of SBVR Structured English to capture the meaning of the tagged regulatory provisions in vocabularies (*alethic logic*, indicating a necessity, possibility, or impossibility) and rules (*deontic logic*, i.e. what should be, specifying

[1] https://www.w3.org/2004/02/skos/.

[2] https://www.w3.org/RDF/.

[3] https://www.w3.org/2003/Talks/0617-Munchen-IH/25.html.

obligations, prohibitions, permissions, etc.). This approach helps formalise regulatory and legal knowledge in the form of vocabularies (a la SKOS) and rules explicitly expressed to identify whether a regulatory provision prohibits certain behaviours, processes or products, or permits them, or obliges firms to behave in a certain way towards customers, and so on. SmaRT regulatory knowledge is persisted in an RDF Knowledge Base to make it machine-readable and machine-computable. AI-based inference and reasoning capabilities are provided by the SmaRT Ontology expressed in OWL.[4] An OWL ontology permits granular semantic querying (using SPARQL) and reasoning to identify new or consolidate existing knowledge—it helps bring semantic interoperability to traditional data systems.

In order to achieve the objective of the Sprint, the Bank of England and the FCA chose to use the Supervision Reporting Requirements provisions in the FCA Handbook (Sup 16.12) along with supplementary definitions supplied by regulators from the Bank of England. Sup 16.12 instructs financial institutions how to report relevant financial data to the UK supervisory authorities—the Bank of England and the Prudential Regulatory Authority (PRA) in this case. The problem with the FCA Handbook, and indeed regulatory rules in general, is that they tend to be ambiguous, drafted in legalese, with links to parent and related legislation, references to financial and technical standards, and links to relevant guidelines.

During the Sprint, the SmaRT application enabled legal and financial industry experts to transform complex legislation in Sup 16.12, related regulatory rules, and other texts containing standards and guidelines, into a human-readable regulatory natural language (RNL) This standards-based RNL is captured by SmaRT using a combination of human knowledge and expertise, augmented by AI and Machine Learning technologies, such as those in RegDelta, and presented to users in an HTML- and XML-based web interface.

Thus, Sprint participants encoded the regulatory provisions into SmaRT's vocabulary and rules in a human-readable format. However, they are also automatically persisted in a machine-readable format in the SmaRT Regulatory Knowledge Base in RDF. As indicated, SmaRT captures the relationships between concepts in RDF triples (e.g. investments

[4] https://www.w3.org/OWL/.

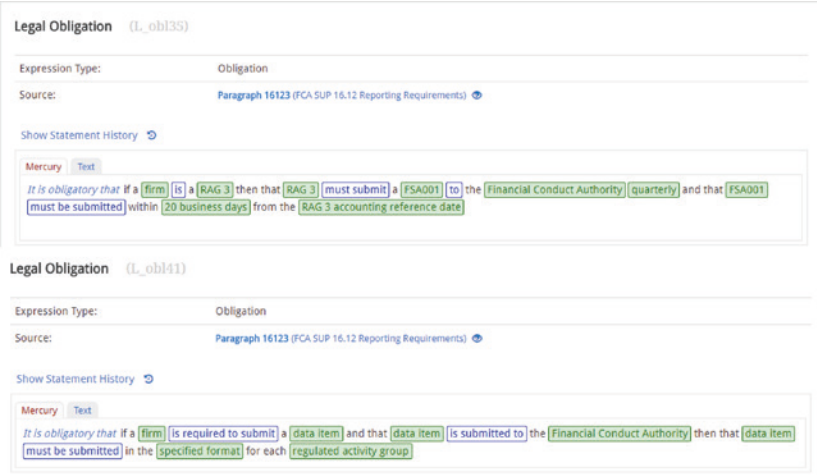


Fig. 6.3 SmaRT rules from Sup 16.12

firm manufactures financial products). The SKOS namespace is limited to expressing taxonomic and related categorical relationships. SmaRT's namespace is much richer. In Fig. 6.3, the relationship RAG must submit FSA001 is captured as an RDF triple. Triples are declarative axioms and the building blocks of SmaRT rules. SmaRT rules are represented as RDF/Turtle graph pattern.[5] Thus expressed, data can be checked for consistency or compliance with such rules. Standard W3C technologies such as SPARQL,[6] SPIN,[7] and SHCL[8] are employed for this purpose. SPARQL (SPARQL Protocol and RDF query language), is, as its name indicates, the W3C query language for the Semantic Web and siloed and distributed networked systems. For example, SPARQL can be used to enable querying and integration of siloed financial and risk data for regulatory reporting and risk management. Rules such as those present in Fig. 6.3 are the constituent elements of complex regulatory provisions expressed in a human-readable format in HTML and in a

[5] https://www.w3.org/2001/sw/DataAccess/rq23/#BasicGraphPatternMatching.
[6] https://www.w3.org/TR/rdf-sparql-query/.
[7] https://www.w3.org/Submission/2011/SUBM-spin-overview-20110222/.
[8] https://www.w3.org/TR/shacl/.

machine-readable format as an RDF graph pattern. The closest analogue to this approach is the relationship between Wikipedia and DBpedia.[9] The human-readable Wikipedia content is expressed in HTML and rendered into web page in a browser; however, DBpedia captures Wikipedia content and data in RDF, with concepts modelled in an ontology. Thus, information contained in related/linked Wikipedia pages can be queried (using SPARQL), extracted, federated, integrated, relationships uncovered or inferred, and new knowledge created.

6.3.3 Phase 3: Performing Digital Regulatory Reporting

The knowledge embedded in the SmaRT RDF-based vocabulary was used by software engineers from Hitachi, Regnosys, and Lombard Risk to map firm-specific data concepts in the anonymised customer account data supplied by Santander to equivalent concepts in the Regulatory Knowledge Base. The RDF-based rule graph patterns were employed to create SPARQL queries to extract compliant data on retail customer accounts. Using the SPARQL Inference Notation (SPIN) framework, rules can be graphed and executed. A software application was created to automate this process. This was then used to extract the required data, transform and load it, and then perform the required calculations and populate relevant cells in the FSA 001 Balance Sheet form for submission to the Bank of England.

The major achievement in executing the PoC came when the rule governing customer account reporting was changed. Once the rule change was captured in SmaRT and expressed in RDF, the software application executed over the changed rule (as an RDF graph pattern) and then populated the appropriate fields in FSA 001 form with the required data. No change in the software algorithm was required. This was a major development and provided the PoC.

6.3.4 Phase 4: Creating Meta-Data Models for Semantic Interoperability

Referring back to Fig. 6.2, Model Drivers (ModelDR) demonstrated a key benefit of the above approach to solving the semantic interoperability problem discussed earlier. ModelDR semantic modelling application

[9] https://en.wikipedia.org/wiki/DBpedia.

was able to create ontology-based meta-data models, based on SmaRT vocabularies and rules, that will help scale up the findings and make Digital Regulatory Reporting a reality in the Enterprise. During the Sprint, the ModelDR application was integrated with SmaRT in order to demonstrate how SMEs could capture domain knowledge (here on regulatory provisions) and use this as an input to semantic models expressed in OWL—such models may be used for data virtualisation (see Kontchakov et al. 2013). Such models are currently being built at great cost by major banks. The ability to have business professionals participate in this process is argued to make this process more efficient and help address the aforementioned translation problem and make semantic interoperability possible. This approach also helps scale up the PoC to a working enterprise-wide solution.

6.4 Discussion and Implications

While Know Your Customer (KYC), Anti-money Laundering (AML), and the financial crime problem spaces are, perhaps, the most mature areas in the application of RegTech, Enterprise Data Management is, perhaps, the most important area for the application of RegTech and the lens through which all RegTech solutions should be viewed. The rationale for this assertion is simple and straightforward—financial enterprises have become more or less fully digitised and almost all people, business objects and processes are represented in and through data, be it structured or unstructured. Regulations themselves are unstructured data—although, regulators are seeking to bring structure to regulatory provisions and rules. In the area of conduct risk, for example, predictive analytics and machine learning are being used to identify insider (cyber) threats, suspicious activity (fraud and financial crime), insider trading, and employee misconduct, all based on data captured from phone calls, emails, business transactions, and so on. Unfortunately, the approaches being taken by multiple vendors using proprietary approaches, and applying technologies as diverse as AI, machine learning, NLP, DLT, biometrics, cryptography, cloud computing, and open APIs, may result in a digital *Tower of Babel*, as *semantic interoperability* a major issue for the industry.

The traditional approach employed in Bamberger's *Technologies of Compliance*, whether GRC or RegTech, is to transform and map regulatory provisions, compliance imperatives and rules into software code.

This approach is evident in early RegTech solutions in the AML/KYC/Financial Crime domains. Thus, financial institutions adopting such technologies face a "black box" solution, with an attendant regulatory risk that a client will, for example, be on-boarded in breach of governing regulations. This might happen if a regulatory rule is not properly encoded or if all permutations and combinations are not accounted or tested for. Depending on the gravity of the breach, a financial institution could find itself with a hefty fine or risk of being put out of business. In the course of our research, lawyers critical of RegTech put this argument forward and criticised the often unquestioning acceptance of vendor claims by financial institutions as to the abilities of their software applications to automate or support decision-making around KYC and client on-boarding. RegTech vendors cite Intellectual Property (IP) considerations for their unwillingness to disclose what is in their black boxes (Tyler 2017). What is required here by financial institutions and regulators is provenance between the original provisions/rules and their instantiation in computer algorithms. The RegTech Sprint PoC provides evidence of the utility of a human- and machine-readable intermediate language in ensuring faithful translation between the source regulations, the interpreted provisions, and their representation in the software that underpins technologies of compliance.

Echoing Kenneth Bamberger's argument presented at the beginning of this chapter, Packin (2018, p. 194) warns of the downside to RegTech and argues that its adoption "requires a carefully tailored design of the technology, a joint effort of the regulators and the private sector and some shifts in corporate thinking." Evidence from the RegTech Sprint indicates that this is underway. However, the RegTech genie is out of the bottle and a major problem facing the industry is the growing number of proprietary RegTech solutions from multiple vendors, none of which are aligned around a common model or infrastructure. It has been brought to our attention by executives from Globally Systemic Important Banks (GSIBs or GSIFIs) that the last thing they want is to have multiple FinTech and RegTech solutions, from multiple vendors, adding to the proliferation of applications across their institutions and to the burgeoning "spaghetti pots" of code and data.

The FCA is advocating the adoption of open source technologies and open semantic standards, such as those developed by the W3C, to link and make machine-readable and machine-executable structured and unstructured data across heterogeneous sources. Hence, its focus is on

XML/RDF (and also Turtle and JSON-LD), ontologies (in OWL), and related standards, such as the SBVR, and its extensions, to express regulatory vocabularies and rules in order to underpin the semantic interoperability of systems.

It is also clear from our experience that large financial institutions are beginning to address the core problems of data governance and data virtualisation using semantic technologies that enable interoperability between systems. SBVR is being used by major banks to help map regulatory concepts on to business concepts. Ontologies are being used as meta-data models hosted in RDF triple stores as knowledge bases for data extracted from heterogeneous relational data stores and other sources. This semantic approach to data virtualisation uses SPARQL to field federated queries over the distributed meta-data/data in relational data stores. The operational data stays where it is, with the data of interest returned from multiple data sources, integrated using the ontologies (as meta-data models), with further analysis and processing performed in an RDF triple store. This approach takes on an AI dimension when inferencing engines or reasoners are employed to add knowledge to a knowledge base. A semantic reasoner or rules engine consists of algorithms that infer logical conclusions from a set of asserted axioms or facts expressed in RDF/OWL. From a data perspective, previously unknown or unrecognised relationships across heterogeneous data sets can be asserted, thus adding more knowledge. Successful ontology-based solutions already exist in a wide variety of domains from defence and intelligence, to capital markets, to regulatory compliance (Palantir 2018).

From a business perspective, this approach enables regulatory semantics (vocabularies and rules underpinning regulatory provisions and compliance imperatives) to be mapped to business semantics (vocabularies and rules expressed in business policies, operational standards, controls through to meta-data repositories/data dictionaries). It also permits the disambiguation, extraction, and integration of heterogeneous firm-specific data for regulatory compliance reporting and risk management. Capturing business semantics in a knowledge base is the *sine qua non* of good data governance. An example here is the Model Knowledge Base referred to in Fig. 6.1. Linking a model knowledge base like this with a regulatory knowledge base and integrating both with a business knowledge base can enable straight-through processing of regulations and automated regulatory compliance reporting of both financial data and non-financial data. Interestingly, industry bodies like the IFRS

Foundation/International Accounting Standards Board are using semantic technologies and ontologies to make their standards and reporting XBRL taxonomies both human- and machine-readable. So too are the ISO20022/SWIFT initiatives in their efforts to enhance its financial messaging standard. Thus, a point of convergence is not too far off, and firms across the financial industry need to ensure that they are in a position to capitalise on the very real benefits offered by semantic technologies for FinTech and RegTech.

6.5 Conclusion

This chapter draws on recent field research to demonstrate the promise and potential of RegTech. However, it adds a cautionary note that the full benefits of RegTech will only materialise if the pitfalls of a fragmented *Tower of Babel* approach are avoided (Butler 2017). Semantic standards are the key to all this.

It is clear that the benefits of RegTech go well beyond straight-through processing of compliance reporting of financial data, such as balance sheet reporting and the quantification of organisational or systemic financial risk. We have argued that one of the benefits of the application of semantic standards is data governance, through the ability to create machine-readable meta-models that enable data virtualisation across heterogeneous data stores. This approach may make the enterprise data warehouse a thing of the past. However, it also enables an enhanced data-driven approach to the management of non-financial risk and associated regulatory compliance reporting. Here, data from siloed, heterogeneous databases can be virtualised and ontologies and/or predictive analytics/machine learning algorithms and AI applied to identify insider or cyber threats, suspicious activities, financial fraud by customers/clients, and a wealth of other applications. Readers will begin to realise that in this context RegTech can be applied across industry sectors, and not just the financial industry, as non-financial risks, such as operational risk and employee misconduct, consumer protection are not industry specific.

The financial industry spends more on IT and data than any other, over $360 billion per annum according to Gartner. Given the fundamental problems, it faces with regulatory compliance, which is costing the industry over $100 billion per annum, and the persistent problem of data governance, management, and analytics, it seems absurd to see financial institutions chasing will-o'-the-wisp solutions or technologies which

may turn out to nothing more be fads, with little practical application or impact. Our previous research revealed that as of late 2016, the industry Chief Data Officers had yet to go beyond CDO 1.0 (Governance) to reach CDO 2.0 (Analytics) (Butler 2017). The point here being that the industry still finds difficulty in realising the benefits of data analytics, and this has major implications for RegTech, as it still has not solved the problems of data governance. Likewise, while it is clear that AI, machine learning, and robotics have significant implications for FinTech and, particularly, RegTech domains, the real benefits of AI, in terms of unsupervised learning, are still some way off. Nevertheless, it is clear that ontologies, machine learning, and NLP technologies are being used effectively. That said, it is clear that AI will do little to address the fundamental issue of "natural stupidity" in financial engineers and quants (Wilmott and Orrell 2017) or in financial experts and senior managers, whether in banks or general business organisations (Kahneman 2012), or in those responsible for systematic misconduct and fraud in the banking system (Vaughan and Finch 2016).

RegTech holds much promise for regulators and firms in the financial industry to fully benefit from the power that digitalisation offers—to solve a big problem for a relatively small effort. However, a considered, collaborative approach by all stakeholders is required if that promise is to become a reality. As societal stakeholders, IS researchers have a role to play in this, in that there is much the discipline can offer in helping practitioners develop and express a "common language" in human- and machine-readable formats.

Acknowledgements SmaRT was developed under the Enterprise Ireland Commercialisation Fund, co-funded by the Irish Government and European Regional Development Fund (ERDF) under Ireland's European Structural and Investment Funds Programme 2014–2020.

References

Arner, D. W., Barberis, J., & Buckey, R. P. (2016). FinTech, RegTech, and the reconceptualization of financial regulation. *Northwestern Journal of International Law & Business, 37,* 371–414.

Bamberger, K. A. (2009). Technologies of compliance: Risk and regulation in a digital age. *Texas Law Review, 88,* 669–739.

Butler, T. (2017). Towards a standards-based technology architecture for RegTech. *Journal of Financial Transformation, 45,* 49–59.

Haldane, A. G. (2012). *Towards a common financial language*. Securities Industry and Financial Markets Association (SIFMA) "Building a Global Legal Entity Identifier Framework" Symposium, New York. Available at: http://www.bis.org/review/r120315g.pdf. Last accessed 17 August 2018.

Heiler, S. (1995). Semantic interoperability. *ACM Computing Surveys (CSUR), 27*(2), 271–273.

JWG. (2017). *RegDelta: Part of our MiFID II solution*. Available at: https://jwg-it.eu/insight/mifid-programme-planner/. Last accessed 25 October 2017.

Kahneman, D. (2012). *Thinking, fast and slow*. England: Penguin Books.

Kontchakov, R., Rodriguez-Muro, M., & Zakharyaschev, M. (2013). Ontology-based data access with databases: A short course. In *Reasoning web: Semantic technologies for intelligent data access* (pp. 194–229). Berlin and Heidelberg: Springer.

Marenzi, O. (2017). *Capital markets and investment banking 2017–2018 forecast*. Available at: http://www.opimas.com/research/193/detail/. Last accessed 25 October 2017.

McDowell, H. (2017). Banks spent close to $100 billion on compliance last year. *The Trade News*. Available at: https://www.thetradenews.com/Sell-side/Banks-spent-close-to-$100-billion-on-compliance-last-year/. Last accessed 17 August 2018.

McNulty, L. (2017). Top regulator: City firms must bear responsibility for RegTech risks. *Financial News*. Available at: https://www.fnlondon.com/articles/city-firms-must-bear-responsibility-for-regtech-risk-20170516. Last accessed 25 October 2017.

Memminger, M., Baxter, M., & Lin, E. (2016). *Banking RegTechs to the rescue?* Available at: http://www.bain.com/publications/articles/banking-regtechs-to-the-rescue.aspx. Last accessed 17 August 2018.

Packin, N. G. (2018). RegTech, compliance and technology judgment rule. *Chicago-Kent Law Review, 93*, 193–220.

Palantir. (2018). *Our solutions*. Available at: https://www.palantir.com/solutions/. Last accessed 17 August 2018.

Palmer, J. (2017). *The data journey: Finding and fixing the bumps & holes in the road*. Data Standards for Granular Data Conference, European Central Bank.

Tyler, T. (2017). *RegTech and FinTech's impact on the regulated sector*. Available at: http://www.aidcompliance.com/regtech-fintechs-impact-regulated-sector/. Last accessed 17 August 2018.

Vaughan, L., & Finch, G. (2016). *The fix: How bankers lied, cheated and colluded to rig the world's most important number*. New York: Wiley.

Walker, O. (2018). M&A in asset management sector climbs to 8-year high. *The Financial Times*. Available at: https://www.ft.com/content/2f1e77f2-f80c-11e7-88f7-5465a6ce1a00. Last accessed 17 August 2018.

Walport, M. (2015). *FinTech futures: The UK as a world leader in financial technologies*. London: UK Government Office for Science.

Wilmott, P., & Orrell, D. (2017). *The money formula: Dodgy finance, pseudo science, and how mathematicians took over the markets*. Wiley.

Woolard, C. (2016). *Innovation in RegTech*. Available at: https://www.fca.org.uk/news/speeches/london-fintech-week-2016-innovation-regtech. Last accessed 17 August 2018.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

