

## **EDM Council Memorandum**

TO: Financial Conduct Authority  
RegTech & Advanced Analytics (regtech@fca.org.uk)

RE: **Using Technology to Achieve Smarter Regulatory Reporting**

FR: John Bottega, Executive Director, EDM Council  
Michael Atkin, Strategic Advisor, EDM Council

DA: June 4, 2018

The EDM Council is a non-profit trade association founded by the financial industry to elevate the practice of data management as a business and operational priority. Our 200+ members represent financial institutions, data vendors, technology companies, regulators and market authorities across the global industry. The Council is involved in the development of data content standards as well as the implementation of data management best practices.

The Council is a strong advocate for the modernization of regulatory reporting. We support this “smarter regulatory reporting” initiative and applaud the FCA for taking leadership in this critical area. The Council is the originator and steward of the Financial Industry Business Ontology (FIBO). FIBO is an open-source, business conceptual model of the structure of financial instruments, business entities and financial processes. It is expressed in the Web Ontology Language (OWL) for machine readable inference processing as well as in glossaries, UML diagrams and spreadsheets for evaluation by financial professionals.

The Council is an original member of the Financial Research Advisory Committee of the US Treasury Office of Financial Research. We have been a long-time participant in discussions with regulators, market authorities and financial institutions on the scope of challenges associated with data harmonization across the global financial industry. We are members of ISO and active participants in TC68/SC9/WG1 (semantic model for ISO 20022).

The Council is currently participating in a semantic proof-of-concept with the Commodity Futures Trading Commission on interest rate swaps. We have been collaborating with the University College Cork and the Governance, Risk & Compliance Technology Centre on the alignment of FIBO with the Financial Industry Regulatory Ontology (FIRO). It is from these perspectives that we offer our commentary.

### **Summary of EDM Council Position**

The EDM Council agrees and fully supports the FCA strategy for the 2017 “Model Driven Machine Executable Regulatory Reporting” TechSprint. We are particularly aligned with the importance of advocacy by FCA to mandate the adoption of both W3C Web standards and semantic precision using ontologies. Regulatory engagement is essential to help overcome industry inertia and facilitate

adoption. In addition to the promulgation of standards, regulators and market authorities have an onward opportunity to ensure precision of meaning (interpretation) of regulatory rule language and to publish the FCA/PRA Handbook in standard RDF. The Council believes the long-term value associated with the adoption of these standards is to change the nature of regulatory oversight by shifting from “report-based” to “meaning-based” (granular) reporting.

We are skeptical of the ability to standardize the meaning of aggregated and derived concepts because many of these concepts are contextual and dependent on the role performed, the strategy adopted and the risk profile of the firm. We agree that the underlying data and business rule concepts can be standardized (aligned to precise meaning). It might be more useful to think about promoting full transparency as the objective of how the reports are aggregated. We believe standardized data in RDF/OWL along with transparent calculations of how the business concepts are derived would provide enhanced flexibility to regulators in performing their supervisory and market oversight functions.

### **Response to Individual FCA Questions**

**Q1:** *Are there more efficient ways to achieve machine executable reporting? Are there better ways to achieve the desired output at each step?*

**EDMC:** We agree with the FCA approach of ontology (verified in conjunction with industry SMEs) using precise semantics published in RDF/OWL. Business rules should be machine executable. There are many ways to implement business rules (i.e. SBVR, RuleML, SHACL). IT professionals will likely not agree on which specific implementation syntax is best. Multiple technology stacks are inevitable and OK if they are logically and structurally interoperable. Regulatory rules should align with the industry-standard ontology to avoid misinterpretation and should be published in executable (i.e. RDF) form.

**Q2:** *What technologies exist that would mean that the manual mapping work in Step 4 could be automated?*

**EDMC:** Mapping from relational databases to RDF is mostly an automated process using the R2RML standard from the W3C. Human mapping is required for ensuring that ontology concepts are precisely implemented. IT experts estimate that machine learning algorithms can be applied to achieve upwards of 80% semi-automation from data elements in RDBMS tables to ontology concepts.

**Q3:** *What is the most effective mechanism for collaboration by firms for addressing potential regulatory reporting ambiguity?*

**EDMC:** The most critical component is the alignment of regulatory rules to precise meaning. This will help avoid the problem of common terms that have different meaning, common meaning that use different terms and vague definitions that don’t capture critical nuances between front and back office applications. We advocate public/private partnership in conjunction with industry trade groups and standards bodies as the most efficient mechanisms for collaboration. The EDM Council has adopted

technical collaboration via GitHub combined with standard products for issue management and testing. We have implemented a slightly modified version of the “[Build, Test, Deploy, Maintain](#)” methodology adapted from the US Department of Defense as the mechanism for ensuring technical governance.

**Q4:** *Are there particular regulatory reporting requirements that could most easily be adapted to machine executable reporting? For example, is a natural starting point to focus on existing requirements that apply to a small set of firms or to a large group of firms? Would a new reporting requirement or an ad hoc data request be more appropriate?*

**EDMC:** We suggest the natural starting point is existing reporting requirements for derivatives transparency. Significant work has already been performed including verification of the ontology and mapping to regulatory reporting requirements. We also encourage the FCA to ensure involvement of the complete supply chain in the initiative (i.e. derivatives firms → ISDA/Common Domain Model → Swap Execution Facilities → Swaps Data Repositories → and regulators).

**Q5:** *Are there any regulatory rules or policies that could be introduced to help implement machine executable reporting?*

**EDMC:** The most important is the requirement that all regulatory reporting must use structured standards. Standard language via the ontology is required to ensure that content is both machine and human readable. All standards-based formats for structured data (i.e. RDF, OWL, XML) should be permitted.

**Q6:** *Are there any specific regulatory rules or policies that could act as a barrier to implementing machine executable reporting?*

**EDMC:** The most significant barrier would be to allow firms to avoid the adoption of standards that is frequently the result of the phrase “use if available” in regulatory rules or policies. We would also caution against segmenting out players by size or any other criteria. There are no significant cost barriers associated with the adoption of data content standards and bifurcating compliance would impede industry-wide adoption. The standards that are adopted need to be truly open source. Coordination across regulatory regimes should be viewed as essential to address any potential interpretation conflicts or semantic inconsistencies.

**Q7:** *What are the opportunities in developing an open source rather than a commercial solution? How can we best use open standards and open collaboration to agree and implement the underlying architecture and approach?*

**EDMC:** The ontology (and its expression in RDF/OWL) must be open source if it is to achieve broad, industry-wide and global adoption. The EDM Council uses and advocates the [MIT open source license](#) as the model for ensuring a level playing field. The adoption of open source standards would not be a disincentive for commercial enterprises to participate in the process. In an open source model,

competition is based on customer support and enhanced functionality. Open source standards form the backbone behind the creation of a broad variety of implementable commercial products that (because of open standards) would be interchangeable. The movement to machine executable reporting using open standards represents an extraordinary opportunity for an expanded partnership between the public and private sectors. We encourage the FCA to use this process as a mechanism to bring all sides of the industry together to validate the various ontologies, clarify any semantic inconsistencies and to define the most efficient approaches to RDF/OWL operational implementation.

**Q8:** *Do you agree with our view of the potential benefits of machine executable reporting?*

**EDMC:** Without question. The adoption of machine executable reporting (i.e. ontology, OWL and SBVR) would eliminate interpretation conflicts that accompany the translation of regulatory language into human understandable business rules and machine executable code. The efficiency gains would also be significant. Relational systems have many unique physical schemas for data. As we exchange data between various systems, the number of physical concepts increases with the number of systems. And for large, global financial institutions the number of physical systems is overwhelming. But from a logical perspective, the number of concepts needed to describe the spectrum of financial contracts is finite. Adopting a shared and reusable model (ontology and OWL) where the baseline concepts have been aligned to explicit meaning, makes it easier to process and integrate data resulting in fewer errors, enhanced data quality validation and significant cost savings. In addition to these clarity and efficiency gains, the adoption of machine executable reporting would significantly increase analytical flexibility for both firms and regulators. The separation of meaning from the structure of data and the linking of concepts without having to reconfigure columns and rows would facilitate scenario-based analysis in response to changing circumstances.

**Q9:** *How do we ensure that the potential benefits and costs are appropriately shared across the industry?*

**EDMC:** The adoption of standards for precision of meaning would reduce costs and improve utility for the entire industry. It would also align with the direction of the financial institutions beyond regulatory reporting. This means that the investment required to comply with regulatory obligations would be extensible for internal operations as well as applicable to new initiatives such as distributed ledger technology and for the adoption of machine learning algorithms.

**Q10:** *Can you provide indicative costs of the current expense of regulatory reporting to your firm? It would be helpful if you are able to separate these costs by specific reporting requirements.*

**EDMC:** Not applicable - we do not have access to this information.

**Q11:** *Which aspects of the current system (interpreting reporting requirements, changes to systems and processes, ongoing data submission, compliance and legal oversight) result in the most significant costs for firms?*

**EDMC:** In our experience, there are three core areas that create significant costs. The first is the maintenance of systems and processes tailored specifically for regulatory reporting. This is often due to the nature of complex reporting requirements from multiple regulatory bodies as well as to the fact that the requirements for regulatory compliance are not synchronized with the requirements for internal operations. The lack of synchronization highlights definitional mismatches and process reconciliations associated with unravelling derived calculations and for contextual data aggregation. The second is the risk of error and misinterpretation of regulatory reporting requirements. When something is identified as erroneous, the cost of forensics (unraveling the complex data meaning/compounding processes) along with the fines associated with enforcement actions can be higher than the costs of reporting. We believe, however, that the most substantial cost is one related to the inability of both the firm and the regulator to fully analyze or gain new insight from regulatory reporting data. Data that is not appropriately defined and harmonized cannot be effectively aligned to critical business and risk operations and not put to productive use. All three of these essential challenges are caused by the reliance on multiple, non-integrated, point solutions – none of which can be reconciled without a shared data “lingua franca.”

**Q12:** *What role would it be most useful for us to play in the progress of this work? Who should take the lead: FCA, industry or a combined approach?*

**EDMC:** Leadership by the Financial Conduct Authority (and regulators in general) on the adoption and mandatory usage of unique identifiers and data content standards would be of great service to the industry. We support the creation of public/private partnerships leveraging global trade associations and international standards bodies as the best mechanism for facilitating progress in the adoption of machine executable reporting. We support the development of global regulatory consortia to reduce redundancy and reconcile both legal and definitional differences.

**Q13:** *Are there existing models of collaboration between industry and regulators, both within and outside financial services that could be adopted?*

**EDMC:** The process developed for the creation of the Global LEI Foundation is an excellent example of collaboration between regulators and industry. We believe it can and should be replicated for the goal of machine executable reporting. We believe the governance infrastructure within ISO, GS1, the World Wide Web Consortia (W3C) and Schema.org provide a reasonable framework to reference.

**Q14:** *Do you have a view on what kind of funding model would be the most appropriate to progress the further development of the initial prototype design?*

**EDMC:** Not currently.

**Q15:** *Can you provide detail on the business case potential of a move toward machine executable reporting for your firm or for firms of varied sizes in general?*

**EDMC:** Not applicable.

**Q16:** *Are there any potential legal or other unintended consequences associated with a move toward machine executable reporting?*

**EDMC:** None that we can identify. Transparency and precision are good for all stakeholders.

**Q17:** *What is the most appropriate model to fund the development and ongoing run costs of machine executable reporting?*

**EDMC:** Every stakeholder would be responsible for their own operational infrastructure (supported by the marketplace of suppliers and expertise). There would be a small initial investment in basic “knowledge graph” architecture including the platform, ontology editor, query editor and storage – but the total cost of implementation is relatively small. A modest operating budget would be required to support the governance mechanism. A progressive assessment (based on size of institution) would be the most equitable funding model.

**Q18:** *How can we ensure that the development of this proof of concept benefits from collaboration with international regulatory counterparts?*

**EDMC:** We recommend the use of established international collaboration mechanisms such as the Financial Stability Board and the International Organization of Securities Commissions to coordinate across regulatory regimes.

**Q19:** *What kind of standards would assist the implementation of machine executable reporting? For example, would a common data model need to be established?*

**EDMC:** Four standards types would be required: (1) standard identifiers for instruments and entities, (2) standard ontology which would obviate the need for a traditional data model, (3) standard business rule mechanism, and (4) standard messaging via JSON-LD for validation.