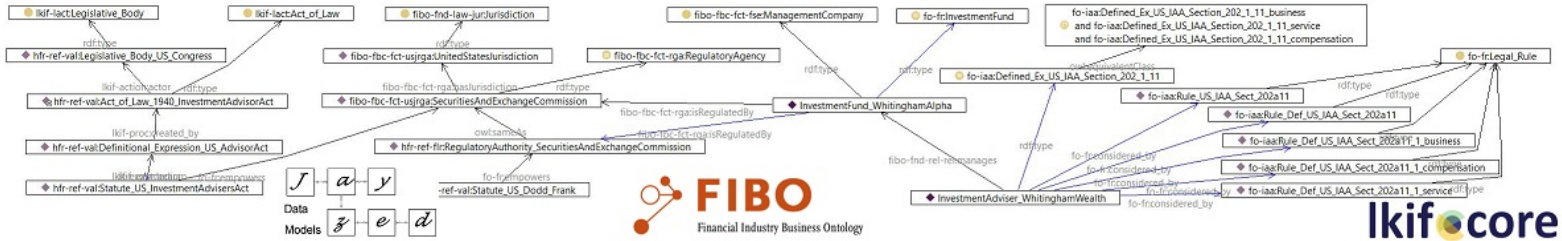
Hedge Fund Regulation Ontology

Prototype v1.0 – Filing the US Investor Advisor Act (IAA)

Jurgen Ziemer, Jayzed Data Models Inc., www.hedgefundontology.com



Agenda

Data Models

I. Introduction
   1. Use Case
   2. OWL primer
   3. Reference ontologies, FIBO and LIKIF

II. Implementing SEC, IAA rules
   1. Loading the regulation: CFR-2012-title17-vol3-part275
   2. Loading and mapping: Public Disclosure Compilation Reports
   3. Processing form Private Fund (PF)

III. Conclusion
   1. Next versions and applications
   2. References and recommended reading

Introduction

The regulatory regime for Hedge Funds has tightened worldwide. European countries implemented the Alternative Investment Manager Directive. The Dodd Frank Act brought Hedge Funds under supervision or the Securities & Exchange Commission as Investment Advisers. Form PF (private fund) requires volumes of heterogeneous data.
Conventional compliance architecture is entangled in numerous systems, transformations, and mappings. Neither the business nor regulators have trust and proof that the code accurately implements the logic of the rules, and that reported numbers accurately trace back to their data sources.

This document proposes a Semantic Web approach to regulatory compliance. Core is the Ontology Web Language (OWL). A W3C standard proven scalability, handling complexity in Bio and Medical field. The core value proposition is that Everything is a Triple. Data, schema, mapping, transformations, rules, … everything is stored a uniform cells of subject-predicate-object.

The Hedge Fund Regulation (HFR) ontology is based on two industry standards:

- FIBO **Financial** Industry Business Ontology for funds, clients, securities, derivatives, positions etc.
- LKIF **Legal** Knowledge Interchange Format for SEC rules, forms, submissions and responses.

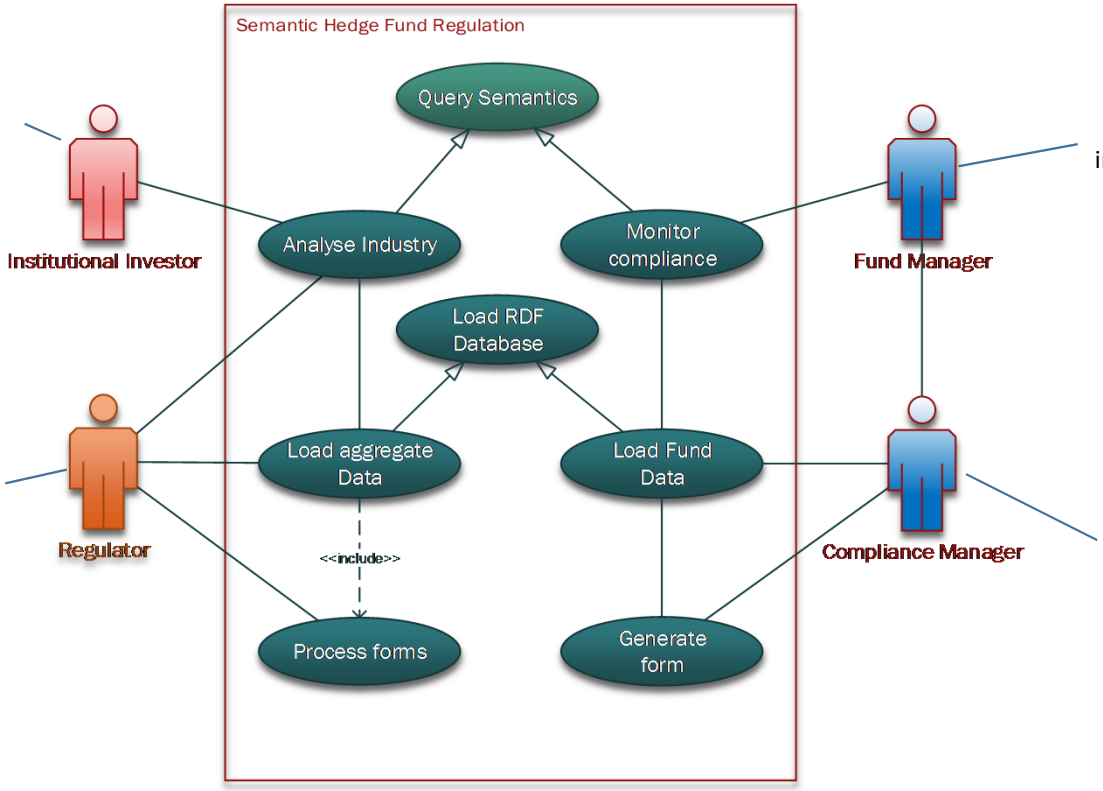
The introduction continues with Uses Case for Regulators and Compliance Managers and a non-technical primer on OWL. We present structure and content of the two reference ontologies and show how HFR integrates and extends them.
The prototype details how the ontology computes a fund's IAA filing requirements. (Code of Federal Regulations, title 17 vol3).

We conclude with an outlook of the next HFR releases and application in other financial areas.

Use Case

The SEC provides the investors with Advisor data in XML. Publication in OWL on a semantic endpoint would serve the public even better and facilitate queries across other information sources

Process IAA and PF forms via ontology alignment.
The aggregated data can be stored in an RDF database. Regulators will be able run SPARQL queries against the database.



The fund manager monitors compliance and can assess the impact of changes to fund structure, investments and client base.

The fund's compliance officer can populate forms using inference and SPARQL rules. The RDF database of the fund holds internal financial data, positions, investors, structure.

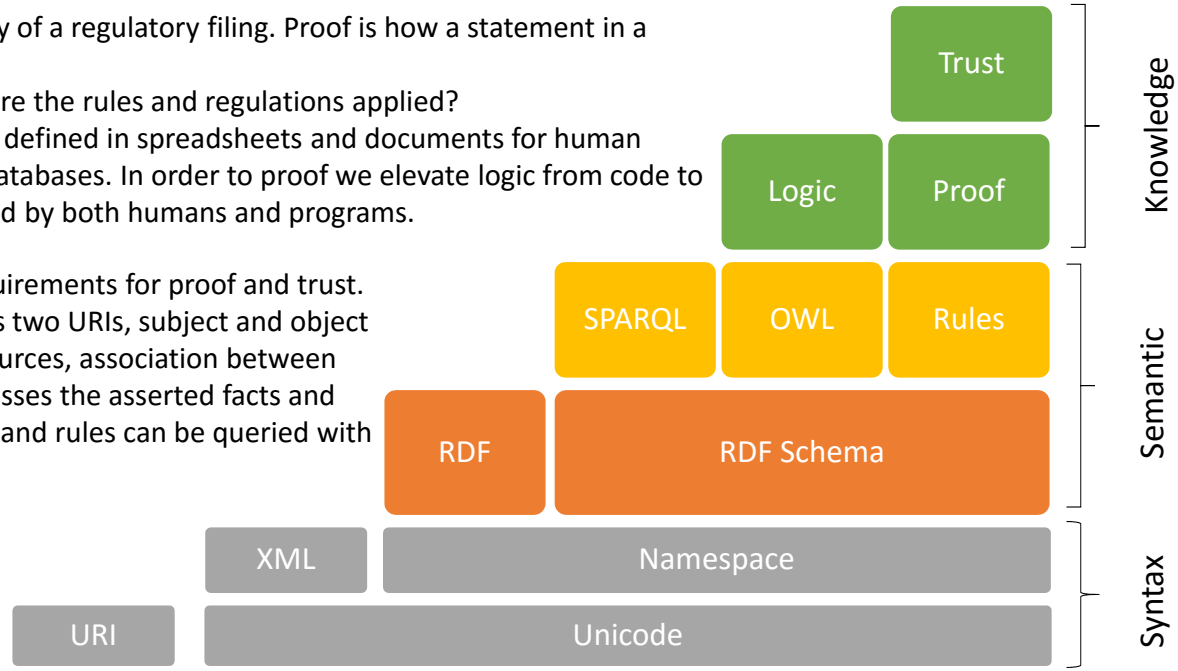
OWL – Primer: Semantic Web Layers

"The Semantic Web is an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation." Sir Tim Berners-Lee, director or the World Wide Web Consortium (W3C)

Trust is the firm belief in the integrity, reliability as accuracy of a regulatory filing. Proof is how a statement in a regulatory filing as been derived.
What are the underlying sources of Financial Data? What are the rules and regulations applied?
In traditional approaches to regulations Business Rules are defined in spreadsheets and documents for human readers. They get mapped and encoded in programs and databases. In order to proof we elevate logic from code to schema. So that a single artifact is understood and validated by both humans and programs.

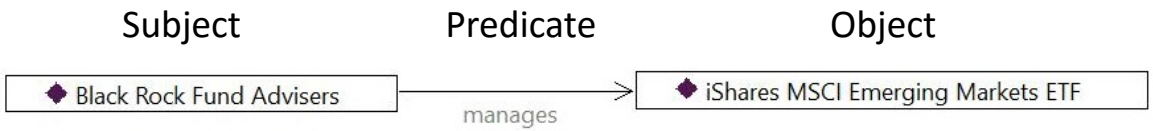
The Semantic layer build upon the foundation to fulfill requirements for proof and trust. RDF removes ambiguity from XML. Unequivocal it connects two URIs, subject and object with a predicate. RDFS and OWL define classes of RDF resources, association between classes and class restrictions. The ontology Reasoner processes the asserted facts and infers new information. All information in including results and rules can be queried with SPARQL selects.

Today's traditional web provides the basic building blocks to encode text and Uniform Resource Identifiers. XML provides machine readable syntax. Namespaces facility XML-Schema definition, XSD.



Everything is a Triple

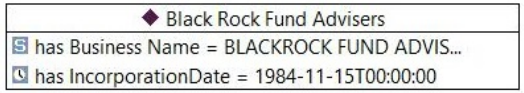
The Resource Description Framework (RDF) uses an elementary grammar to define a cell of information.



Black Rock manages the iShares MSCI Emerging Markets ETF. Both are web resources identified by a URI.
The **Object Property** *manages* connects the two resources.
This is similar to creating

- an association between objects. E.g. associating two Java instances.
- a Foreign Key reference between two database records.

The **Data Property** assigns a value to the subject resource.
Here is the Ontology Web Language (OWL) code. The colon separates the Namespace from the resource name.
rdfs:label - a build in RDF data property
fibo-fnd-rel-rel:manage – an object property defined in the Financial Services Business Ontology (FIBO).



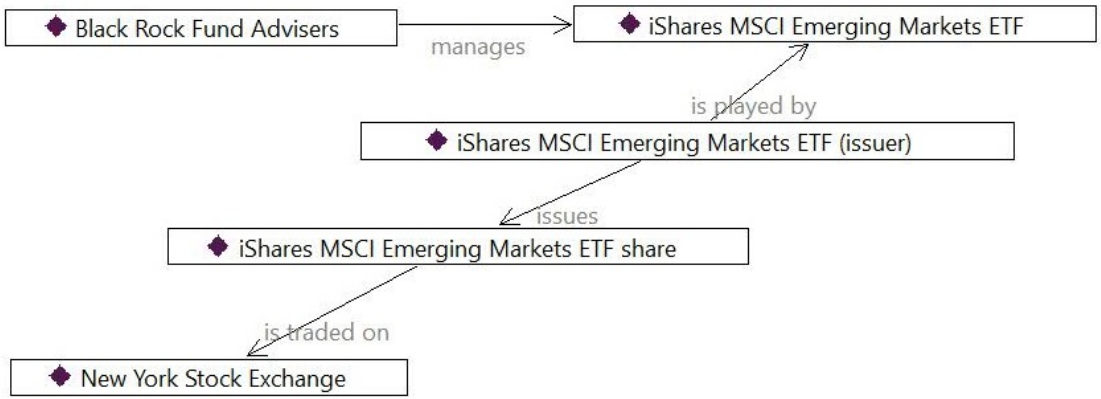

```
:Black_Rock_Fund_Advisers
  rdf:type fo-fr:InvestmentAdviser ;
  :hasBusinessName "BLACKROCK FUND ADVISORS"^^xsd:string ;
  :hasIncorporationDate "1984-11-15T00:00:00"^^xsd:dateTime ;
  fibo-fnd-rel-rel:manages :iShares_MSCI_Emerging_Markets_ETF ;
  rdfs:label "Black Rock Fund Advisers"^^xsd:string ;
```

Data is a triple

The iShares ETF (the legal entity of the fund is an issuer of shares. The shares are traded on the NYSE. All depicted instances may have more data properties and object properties connecting to other instances. We get a web of information called a Graph.

SPARQL, the Protocol and RDF Query Language is SQL for RDF/OWL ontologies. The ontology can be file based or an RDF database aka a Triple Store.

In order to list the exchanges, where Black Rock ETFs are traded, we traverse the graph chaining all object properties.

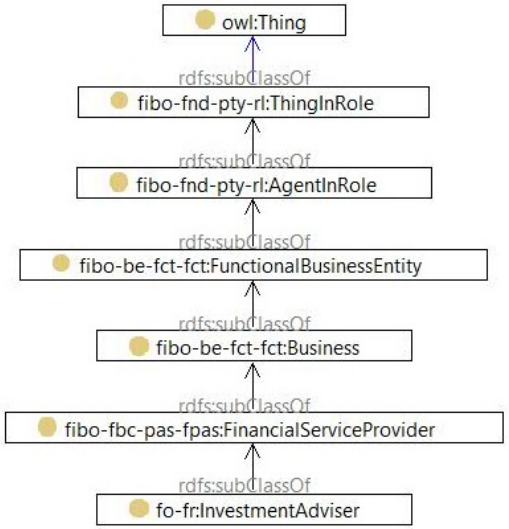


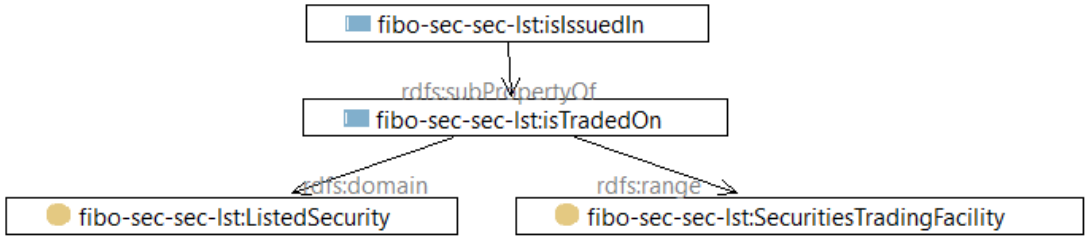

```
SELECT ?investmentadviser ?securitiesexchange
WHERE {
   ?investmentadviser fibo-fnd-rel-rel:manages ?managementcompany .
   ?exchangetradedsecurity fibo-sec-sec-lst:isTradedOn ?securitiesexchange .
   ?issuer fibo-fbc-fct-rga:issues ?exchangetradedsecurity .
   ?issuer fibo-fnd-pty-rl:isPlayedBy ?managementcompany .
}
```

The schema is in triples

RDF Schema (RDFS) extends RDF to allow describing taxonomies of classes and properties.

OWL classes are interpreted as sets that contain individuals aka instances. The RDF object property rdf:type asserts that the individual is an instance of the class.





We can define data and object properties of the class. The RDFS property's Domain restricts the type of the subject, Range restricts the type of the object.

The RDF-schema property subClassOf enables rich hierarchies of classes. The subPropertyOf build hierarchies of class association.
The core difference to object models is the notion of the set rather then class objects.
In Java we must explicitly assign Black Rock to be an Investment Adviser and assign it to be a Financial Service Provider. In RDFS/OWL an instance of Investment Advisor is automatically an instance of all the super classes.

Logic described in triples

RDFS provides constructs similar to UML class / object and data modeling Entity Relationship models.
Modeling tools provide model transformations, import and export between OWL, ER and UML.

- Data modelers still need to derive the Physical model from the Logical Model.
- Object modelers generate code from UML

The ontology already contains schema and data expressed in triples and we can use metadata in our SPARQL queries.

Ontology
- Class
- Sub Class
- Data Property
- Object Property

Entity Relationship
- Entity
- Subtype
- Attribute
- Relationship

UML Class
- Class
- Generalization
- Attribute
- Association

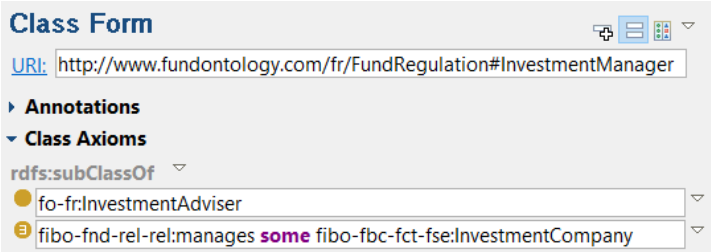
```
SELECT ?investmentadviser ?managementcompany ?FS_provider ?ETF_class
WHERE {
   ?investmentadviser fibo-fnd-rel-rel:manages ?managementcompany .
   ?investmentadviser rdf:type ?FS_provider .
   ?managementcompany rdf:type ?class .
   ?managementcompany rdf:type ?ETF_class .
   ?FS_provider rdfs:subClassOf fibo-fbc-pas-fpas:FinancialServiceProvider .
}
```

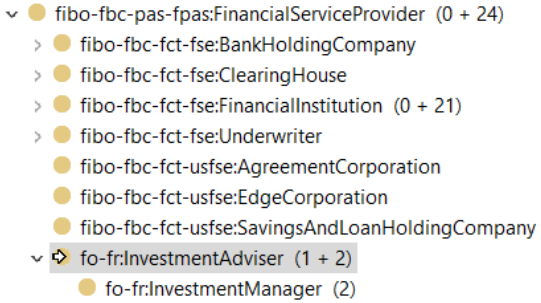
[investmentadviser]	Black_Rock_Fund_Advisers
[managementcompany]	iShares_MSCI_Emerging_Markets_ETF
FS_provider	fo-fr:InvestmentAdviser
ETF_class	fibo-fbc-fct-fse:ManagementCompany

Class restrictions

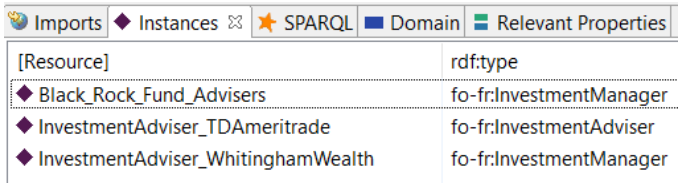
A restriction describes a class of individuals based on the properties that instances of the class participate in. In other words a restriction is a kind of class, just like a named class. Typically, we apply restrictions to narrow down the set of instances from class to subclass.

An Investment Manager is an Investment Adviser, who manages some Investment Company. In the ontology editor (Topbraid Composer) we define the **existential restriction (some)** on the object property. The **universal restriction (only)** would mean that the adviser manages nothing else but Investment Companies. We can place restrictions on Quantifiers, Cardinalities and even Values. An example of a Value restriction is a Prime Bond with hasRating **only** 'AAA'.





The class hierarchy browser shows the number of direct class instances plus total of instances in the subclasses. Our example has Black Rock and Whitingham Wealth as Investment Adviser. TDAmeritrade doesn't manage funds.



Defined Class and the Reasoner

So far, the class restrictions although more refined than database constraints still depend on the implementation to assert that Black Rock is an Investment Manager and TDAmeritrade is only an Investment Advisor.
Changing Investment Manager to a **Defined Class** will have the Resoner infer this new knowledge from existing asserted information.

We move the class restriction from rdfs:subTypeOf to owl:equivalent class. Now the Manager is defined as the intersection of all Investment Advisers **and** all things that manage an Investment Company.

```
owl:equivalentClass
fo-fr:InvestmentAdviser
  and (fibo-fnd-rel-rel:manages some fibo-fbc-fct-fse:InvestmentCompany)
```

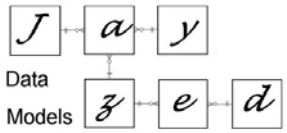
The Reasoner tool processes class definitions and infers that a particular individual is a member of the class.

- fibo-fbc-pas-fpas:FinancialServiceProvider (0 + 53)
  - fibo-fbc-fct-fse:BankHoldingCompany
  - fibo-fbc-fct-fse:ClearingHouse
  - fibo-fbc-fct-fse:FinancialInstitution (0 + 37)
  - fibo-fbc-fct-fse:Underwriter
  - fibo-fbc-fct-usfse:AgreementCorporation
  - fibo-fbc-fct-usfse:EdgeCorporation
  - fibo-fbc-fct-usfse:SavingsAndLoanHoldingCompany
  - fo-fr:InvestmentAdviser (4 + 6)
    - fo-fr:InvestmentManager (2)
    - fo-fr-iaa:Defined_Ex_US_IAA_Section_202_1_11_business (2 + 2)

The class hierarchy now shows the defined classes in blue and we see two inferred instances: Black Rock and Whitingham Wealth.
Checking for class **subsumption** is a key task of the reasoner key differentiator to non-semantic technologies.

We formulate complex IAA rules as Defined Classes from class restriction building blocks. The Reasoner places the matching Funds, Advisers, Assets as inferred members of the rule. –> Semantic Regulation Compliance

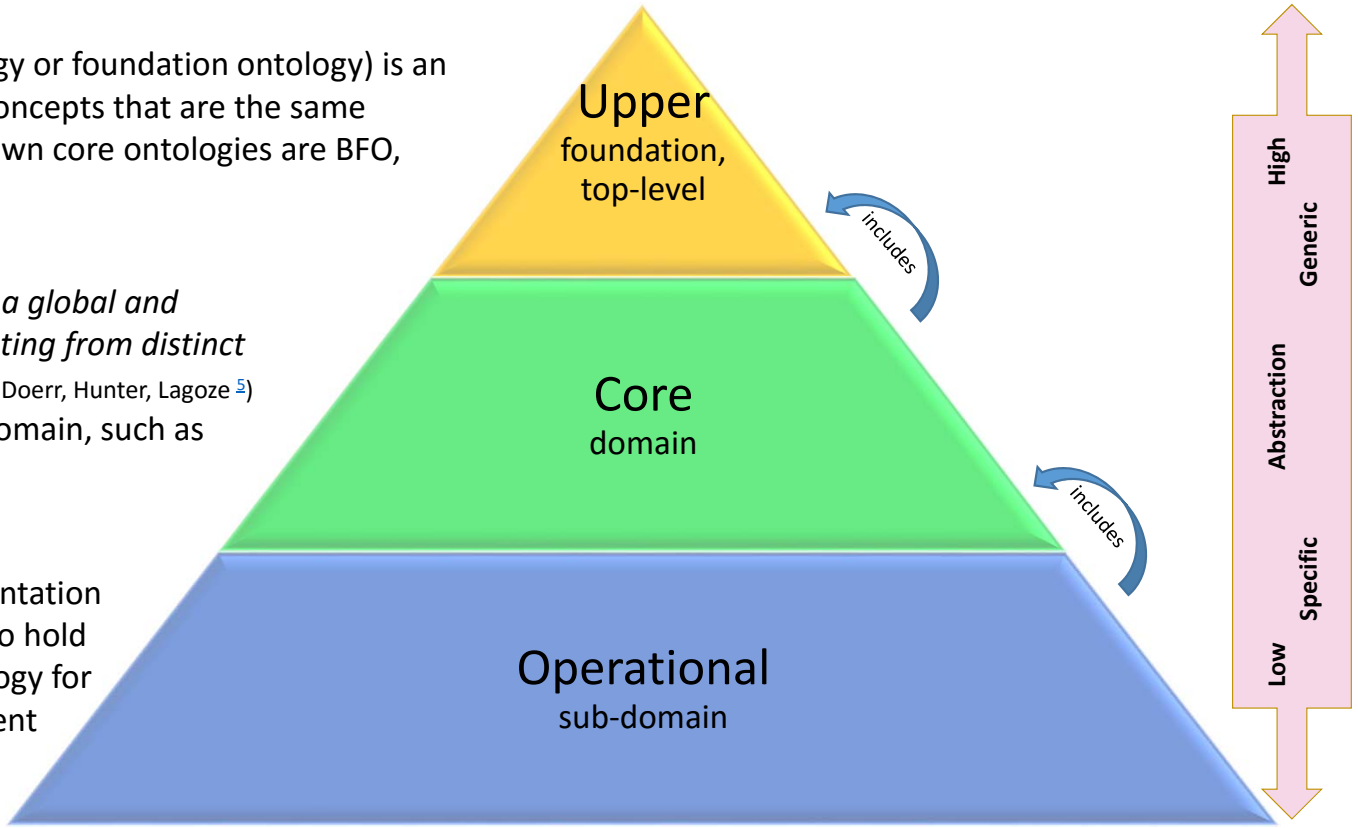
Reference Ontologies - levels



An upper ontology (aka top-level ontology or foundation ontology) is an ontology which describes very general concepts that are the same across all knowledge domains.[6] Well known core ontologies are BFO, GFO, DOLCE, SUMO and Dublin Core.

The goal of a core ontology is to provide a global and extensible model into which data originating from distinct sources can be mapped and integrated. (Doerr, Hunter, Lagoze [5])
The core ontology applies to a specific domain, such as biology, medical, Legal and Finance.

The Operational Ontology is an implementation of a core ontology. It is specific enough to hold source data. HFR is an operational ontology for the sub-domain of (alternative) investment management.

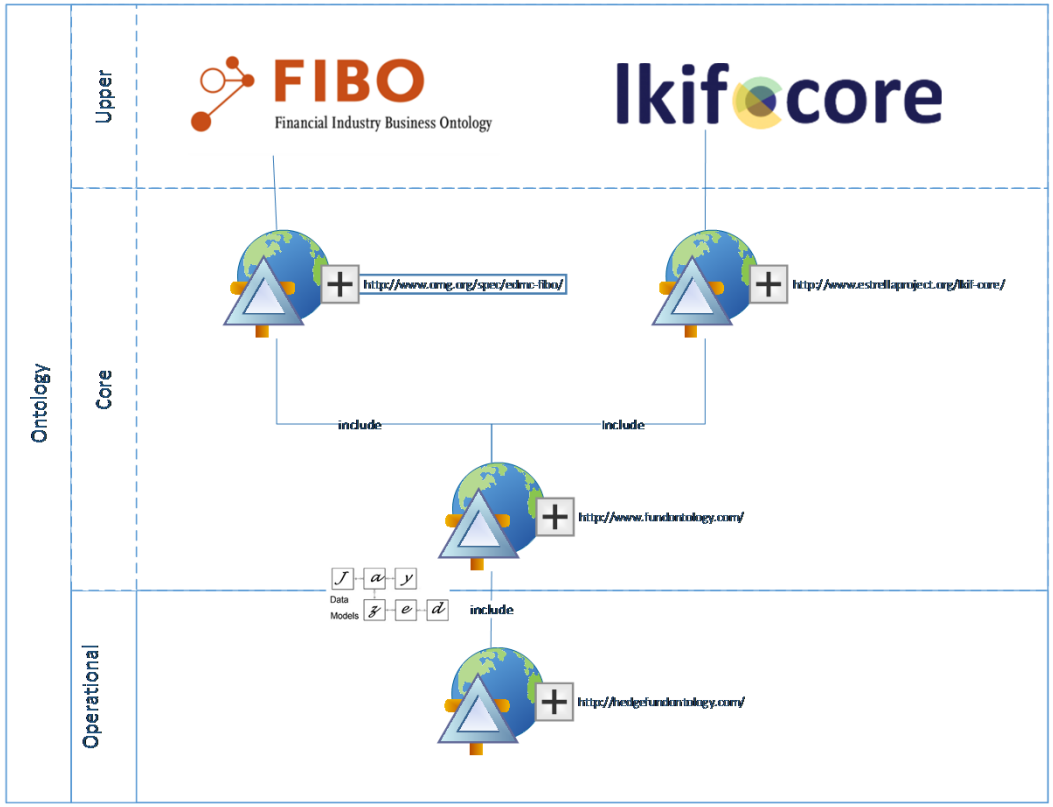


Reference ontologies for HFR

It is a standard practice and often mandated for Data Modelers at Financial Institutions to utilize Reference Models*. For a Semantic Web solution, integration with Industry Standards is even more important.

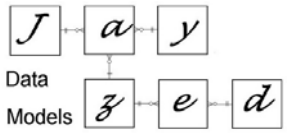
Hedge Fund Regulation (HFR) is an operational ontology integrating legal and financial information.
http://hedgefundontology.com/

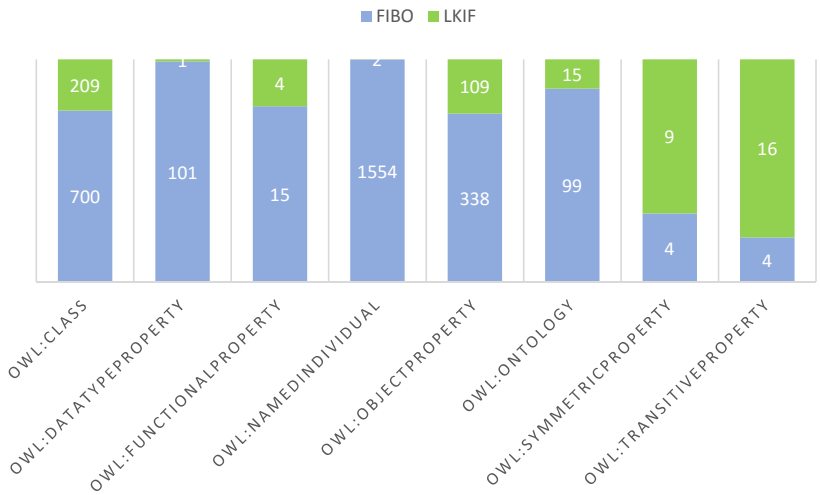
- HFR includes Fund Regulation (FR). A core ontology with concepts common for all investment vehicles. http://fundontology.com/
- FR includes the Financial Industry Business Ontology (FIBO). http://www.omg.org/spec/EDMC-FIBO/
  FR includes the Legal Knowledge Interchange Format (LKIF). A core ontology for Legal domain. http://www.estrellaproject.org/
- Unfortunately neither FIBO nor LKIF utilize an upper ontology for generic concepts.



* Over 150 banks for example have licensed the IBM Banking Data Warehouse model [7].

FIBO / LKIF statistics





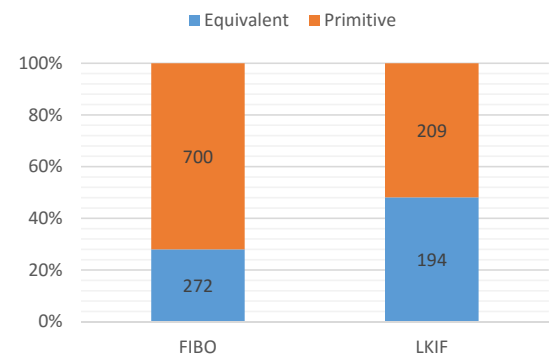
The statistics show the number of elements in the two reference ontologies.

- Class – Together we have over 900 classes to utilize and leverage. LKIF provides over 200 classes defining legal concepts. FIBO already has 700 classes and is still growing. (see following detail slides)
- Note the absence of Data Properties and Named Individuals in LKIF. FIBO has some operational elements, data properties and individuals for Currencies, Countries and even the 12 Federal Reserve district banks.
- The OWL:Ontology bar shows the actual number of OWL files.
- The high number of symmetric and transitive properties show the high ontological commitment in LKIF.

Three-layer ontology design pattern proposes

- *Primitive layer consists of classes/properties forming taxonomic trees in which a single parent may be asserted.*
- *Complex layer refines the primitive layer by imposing restrictions such as necessary or necessary and sufficient conditions beyond the asserted subsumption.*
- *Application restriction layer applies highly restrictive constraints and may be used for the purposes of document validation and application interoperability. (Dumontier, Villanueva-Rosales* [8])

The ratio of Primitive to Equivalent classes show that LKIF has larger Complex layer than FIBO.



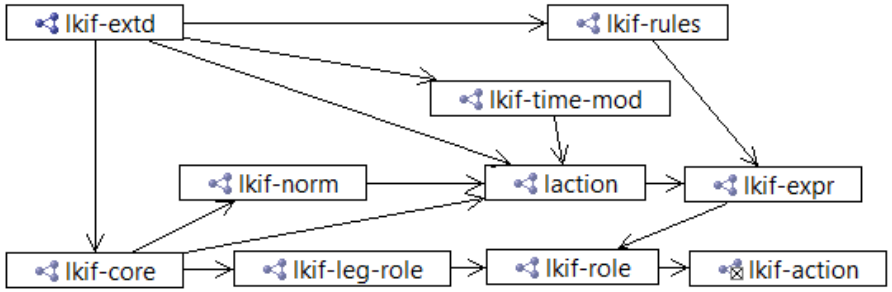
Legal Knowledge Interchange Format (LKIF)

The **E**uropean project for **S**tandardized **T**ransparent **R**epresentations in order to **E**xtend **L**egal **A**ccessibility (Estrella, IST-2004-027655) aims to develop and validate an open, standards-based platform allowing public administrations to develop and deploy comprehensive legal knowledge management solutions, without becoming dependent on proprietary products of particular vendors.[5]

LKIF is intended to model legal rules of the kind found in legislation and regulations.[6]
It is the main deliverable of the ESTRELLA project. LKIF is an Upper and Core ontology. The lead architect, Rinke Hoekstra made the OWL files are available on GitHub: https://github.com/RinkeHoekstra/lkif-core

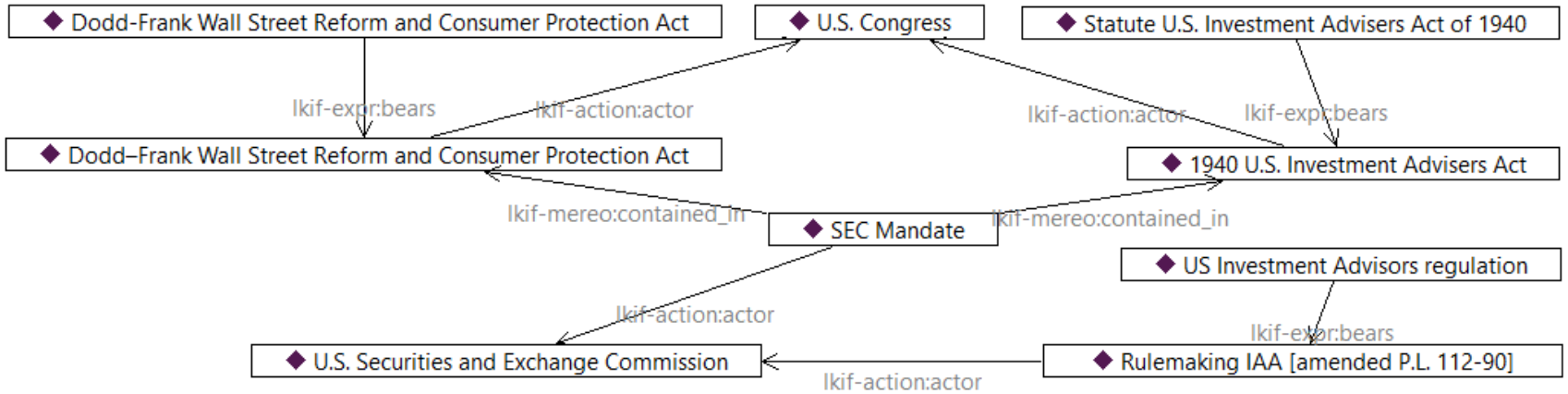
The import graph shows the major OWL files. The Jayzed approach to reference ontologies is to extend – not to change the domain standard. We did add namespace prefix "lkif" to differentiate from "fibo" and "fo-fr" (Fund Ontology – Fund Regulation).

Core is the base ontology module. It contains abstract and basic concepts of an Upper ontology. LKIF **Extended** adds rules and the basic concept of time modification.
Role is a basic concept, corresponding to role in the FIBO. **Legal Role is** a domain concept specializing Role. HFR extends the **Norm** to hold Laws and Regulations for the Investor Advisor Act. The **Action** module defines Agents and acts they perform. **Expression** models propositions, qualifications, statements and legal documents.



US Legal Framework instances

The diagram shows LKIF instances to define semantics of the IAA legal background. Congress enacted (**actor**) the 1940 Investor Adviser Act and the Dodd-Frank Act. There are two instances each: The Act and the Statute. The Statute **bears** the text of the Act of Law passed in congress.

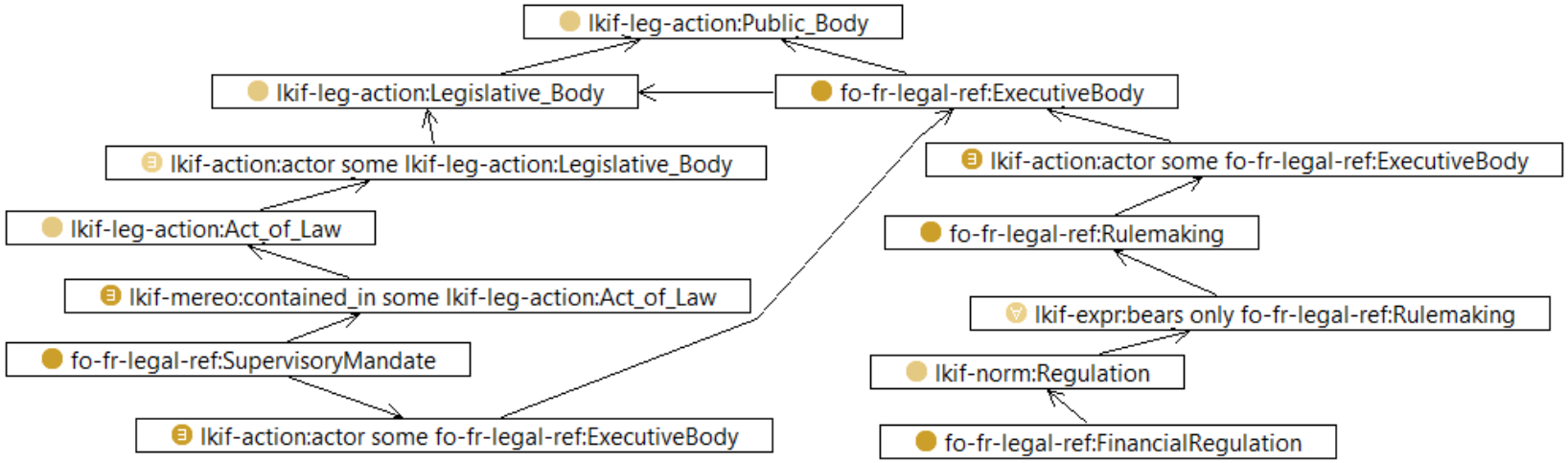


The laws **contain** provisions that give the Security and Exchange Commission a **mandate** to supervise Investment Advisers. The SEC **enacted** a Rulemaking (IAA amended P.L. 112-90). In other words, the SEC announced the final version the regulation. The US Investment Advisors regulation **bears** the text of the rules (e.g. CFR-2012-title 17-vol3 part 275).

US Legal Framework Classes

The graph shows some of the LKIF and HFR classes for the instances. Solid yellow dots represent primitive classes.

Existential 'some' restriction universal 'only' restriction



Legislative Body is the class for US Congress. The **Executive Body** holds the SEC and other Regulators. An existential restriction ties the Legislative Body to the **Act of Law**. The US IIA and Dodd-Frank are an acts of law. The Act of Law contains a **Supervisory Mandate**. A class restriction refers to the Executive Body that got the mandate. Corresponding to the Act of Law we have a class for the **Rulemaking**. The **Regulation** class anchor for rules. Part II will explain more LKIF concepts and their implementation.

Financial Industry Business Ontology (FIBO)

FIBO is a collaboration between the Enterprise Data Management Council (EDMC) and the Object Management Group. The EDMC leads design in collaboration with major Financial Institutions. OMG provides governance and publishes FIBO as a formal standard.

The EDM Council is a 501(c)(6) non-profit trade association founded by the financial industry to elevate the practice of data management as a business and operational priority. The Council is a leading advocate for the development and implementation of data content standards and the publication of data management best practices.[7]

The Object Management Group® (OMG®) is an international, open membership, not-for-profit technology standards consortium, founded in 1989. OMG standards are driven by vendors, end-users, academic institutions and government agencies. OMG Task Forces develop enterprise integration standards for a wide range of technologies and an even wider range of industries.[8]

FIBO™ is a business conceptual ontology standard providing a description of the structure and contractual obligations of financial instruments, legal entities, market data and financial processes.
The primary application of the business conceptual ontology is for data harmonization and for the unambiguous sharing of meaning across data repositories. This common language (or Rosetta stone) for the financial industry supports business process automation and facilitates risk analysis.[9]

II. Implementing SEC IAA rules

Data Models

Identify and review Legal Sources

- US Congress
  - Investment Advisers Act of 1940
  - Dodd-Frank Act – title IV Hedge & Private Equity Funds
- Securities and Exchange Commission
  - Code of Federal Regulations (CFR): Title17-vol3
- Government Publishing Office
  - GPO publishes and XML version of the CFR.

scope →

Load CFR into ontology

- Reverse engineer XML/XSD files into OWL CFR classes
- Identify and extend Legal Knowledge Interchange Format (LIKIF) ontology classes: *Legal Source, Legal Document, Regulation, Statute, Code Federal of Regulations*
- Map CFR to Fund Ontology and create mapping rules
- Run Reasoner to execute mapping rules. That will move the data from CFR to Fund Regulation Ontology.

load →

Define Fund Regulation in OWL classes and axioms.

- Identify and extend Financial Industry Business Ontology (FIBO) classes for Fund data.
- Identify and extend LKIF classes for Rules: Qualified, Normatively Qualified
- US IAA, Ex. US IAA Section 202-1-11
  - Defined Ex. US IAA Section 202-1-11
  - Excluded Ex US IAA Section-202-1-11
- Run Reasoner to compute inferences.
  That will classify a Fund that fulfills the rule axioms as a member of the US IAA Section 202-1-11 class

infer →

Examine classified funds, reasons for classification and linage to Legal source

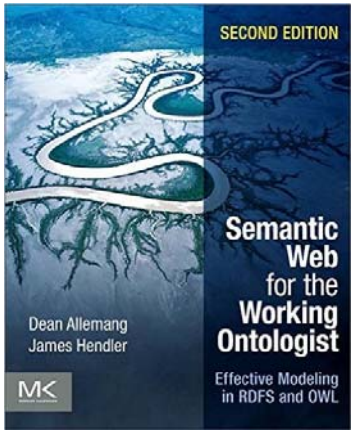
- Review funds classified as Mandated to file
- US IAA, Ex. US IAA Section 202-1-11 list "Whitingham Wealth". It does not list any other excluded Funds.
  - Defined Ex. US IAA Section 202-1-11 Whitingham Wealth is included in all the subclasses.
  - Excluded Ex US IAA Section-202-1-11 Whitingham Wealth is not excluded.
- Traverse the Graph to browse Linage, Legal application, and Financial Information.
- Run SPARQL queries and export results

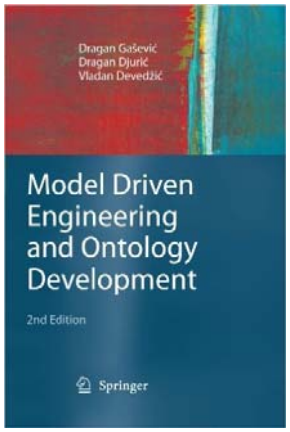
Links and References

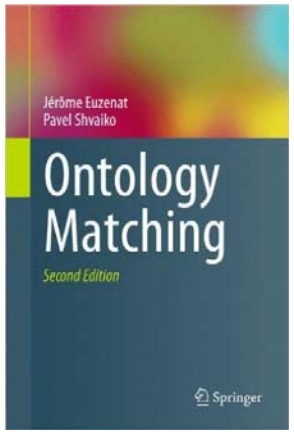
Legal Knowledge Interchange Format (LKIF), http://www.estrellaproject.org/

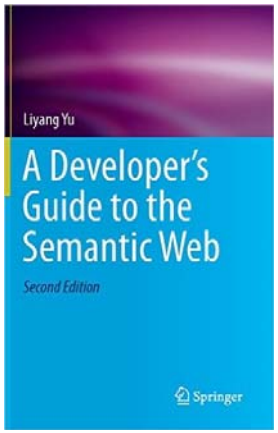
Financial Industry Business Ontology (FIBO)
- Enterprise Data Management Council http://www.edmcouncil.org/financialbusiness
- Object Management Group http://www.omg.org/hot-topics/finance.htm

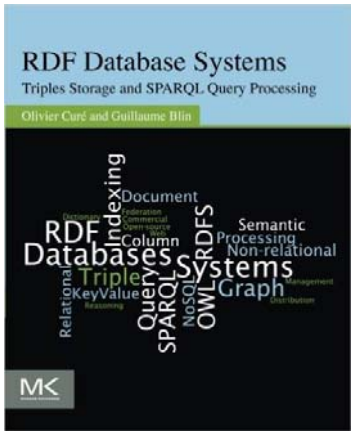
Jayzed Data Models http://www.jayzed.com
www.fundontology.com www.hedgefundontology.com











References

1. Towards a core ontology for information integration
   M Doerr, J Hunter, C Lagoze - Journal of Digital information, 2006 - journals.tdl.org
2. Upper ontology
   Wikipedia - https://en.wikipedia.org/wiki/Upper_ontology
3. IBM Banking Data Warehouse – product page.
   http://www-03.ibm.com/software/products/en/banking/
4. Three-Layer OWL Ontology Design
   Michel Dumontier, Natalia Villanueva-Rosales1 - School of Computer Science, 2Department of Biology, Carleton University, 2007
5. ESTRELLA project website: http://www.estrellaproject.org/
6. ESTRELLA Deliverable N°: 4.1 -The Legal Knowledge Interchange Format (LKIF)
   http://www.estrellaproject.org/doc/Estrella-D4.1.pdf
7. Enterprise Data Management Council website: http://www.edmcouncil.org/
8. Object Management Group website: http://www.omg.org/
9. Financial Industry Business Ontology (FIBO) project website: http://www.edmcouncil.org/financialbusiness