Open Risk API

Improved financial risk management through open data, open source and web technologies

Philippos Papadopoulos, OpenRisk

TopQuants/DNB Autumn Meeting
Amsterdam, Nov 18, 2015

Work supported in part by an EU FIWARE-FINODEX Grant

Risk IT Spending

IDC Financial Insights on *Risk IT* spend (2015):

"Investments in risk technologies account for an average of **17.1%** of overall IT spending (~**$485 billion**)".

Trick Question: What is the industry getting in return for **80bln** annual spend on Risk IT?

BCBS 239

"One of the most significant lessons learned from the global financial crisis that began in 2007 was that banks' **information technology (IT) and data architectures were inadequate** to support the broad management of financial risks"

Did anybody have "a satisfaction guaranteed or you money back" contract?

Einstein's Insanity Quote

{ "Insanity”: “Doing the same thing over and over again and expecting different results” }

What exactly are we doing *differently* concerning risk data and risk models 8 years after the crisis?

Structure of this Talk

- Risk Data
  - Data Formats and Data Standards
  - SDMX
  - REST and Linked Data
- Risk Models
  - Linked Models
  - Description of a Model (DOAM)
- Example
  - Open Risk Dashboard

DATA

CSV

All metadata is hardcoded in column header strings

Date	Open	High	Low	Close	Volume	Adj Close
2015-11-06	2098.600098	2101.909912	2083.73999	2099.199951	4369020000	2099.199951
2015-11-05	2101.679932	2108.780029	2090.409912	2099.929932	4051890000	2099.929932
2015-11-04	2110.600098	2114.590088	2096.97998	2102.310059	4078870000	2102.310059
2015-11-03	2102.629883	2116.47998	2097.51001	2109.790039	4272060000	2109.790039
2015-11-02	2080.76001	2106.199951	2080.76001	2104.050049	3760020000	2104.050049
2015-10-30	2090	2094.320068	2079.340088	2079.360107	4256200000	2079.360107
2015-10-29	2088.350098	2092.52002	2082.629883	2089.409912	4008940000	2089.409912
2015-10-28	2066.47998	2090.350098	2063.110107	2090.350098	4698110000	2090.350098

What is the currency?
Who or what collected the data?
What does “Open” mean?
Is the accuracy spurious?
...

XML (Extensible Markup Language)

```
<?xml version="1.0" encoding="ISO-8859-1"?>

<datafeed>

<quote f6="17.95" f8="2895611" f10="17.94" f11="17.93"
f2="18.09" f3="17.76" f4="17.84" f1="17.31" f5="157943943"
f12="476300" f13="100" f25="Bank of America Corp."
f24="1.35" f42="N/A" f21="18.48" f22="14.6" f14="0.64"
f15="3.6973" f104="51976217.45" f98="1.7962" f99="1.114"
f97="N/A" f100="N/A" f101="13.2963" f102="N/A" f103="N/A"
f39="US0605051046" f26="0.2">06 Nov 21:00, GMT</quote>

</datafeed>
```

XML very powerful but it is difficult to work with

JSON (Javascript Object Notation)

```
{ "quote” : {
        "symbol":"BHP.AX",
        "Ask":"22.75",
        "AverageDailyVolume":"9521200",
        "Bid":"22.65",
        "BookValue":"12.17"
    },
  “Names” : [“BHP”, “Biliton”]
}
```

Arbitrary nesting of objects { }

Array Type [ ]

Data Standards and Data Formats

Data **Standards** have two components:

- Defining *what is being encoded* (taxonomies, vocabularies) - “metadata”
- Prescriptions of what constitutes a *valid instance*

Data **Formats** (e.g. File Formats) are the means to create valid instances of Data Standards - “actual data files”

Data problems are more due to lack of standards than due to poor data formats...

Data Standards and Formats

There are no Data Standards specific to financial risk management.

Some near misses:

- XBRL (eXtensible Business Reporting Language)
  - Focused on financial reporting
- SDMX (Statistical Data and Metadata eXchange)
  - Focused on statistical data exchange
- FpML, FIX, SWIFT etc are for product and transaction data exchange

SDMX could be used for some purposes (timeseries of historical data)

JSON-LD (Linked Data Format)


```
[
  {
    "@id": "http://127.0.0.1:5012/models/hhi",
    "@type": ["http://openriskplatform.org/ns/doam#model"],
    "http://openriskplatform.org/ns/doam#name": [ { "@value": "HHI" } ],
    "http://xmlns.com/foaf/0.1/name": [ { "@value": "Hirschman" } ],
    "http://xmlns.com/foaf/0.1/nick": [ { "@language": "en", "@value": "hhi" } ]
  },
  {
    "@id": "http://127.0.0.1:5012/models/shannon",
    "@type": [ "http://openriskplatform.org/ns/doam#model" ],
    "http://openriskplatform.org/ns/doam#name": [ { "@value": "Shannon" } ],
    "http://xmlns.com/foaf/0.1/name": [ { "@value": "Shannon" } ],
    "http://xmlns.com/foaf/0.1/nick": [ { "@language": "en", "@value": "sha" } ]
  }
]
```

SDMX

SDMX Information Model

- Data Set: a collection of similar data, sharing a structure, which covers a fixed period of time.
- metadata set is information regarding the formal SDMX view of statistical
- exchange: it may describe the statistical approach; the maintainers of data or data structures;
- data release calendar; the quality of data, etc
- Concept Schema
- Data Structure Definition (DSD) The DSD describes the structure of a particular set of data through a list of descriptor concepts.
- It defines which concepts are dimensions (identification and description - for example: Frequency, country, variable/topic, time period), and which are attributes (just description / qualification - for example the unit of measure, confidentiality flag or the data status flag).
- In addition it determines the attachment level for each of these concepts, based on the packaging structure (Data Set, Group, Series/Section or Observation) as well as if they are mandatory or conditional

SDMX Example

- USD / EUR Exchange Rate

SDMX Example

<message:GenericData
xsi:schemaLocation="http://www.sdmx.org/resources/sdmxml/schemas/v2_1/message
https://sdw-wsrest.ecb.europa.eu:443/vocabulary/sdmx/2_1/SDMXMessage.xsd
http://www.sdmx.org/resources/sdmxml/schemas/v2_1/common
https://sdw-wsrest.ecb.europa.eu:443/vocabulary/sdmx/2_1/SDMXCommon.xsd
http://www.sdmx.org/resources/sdmxml/schemas/v2_1/data/generic
https://sdw-wsrest.ecb.europa.eu:443/vocabulary/sdmx/2_1/SDMXDataGeneric.xsd">
<message:Header></message:Header>
<message:DataSet action="Replace" validFromDate="2015-11-06T11:49:36.745+01:00"
structureRef="ECB_EXR1"></message:DataSet>
</message:GenericData>

REST API's and Linked Data

REST Principles (1)

Separate Client from Server:

- Clients can be browsers, mobile or desktop apps, anything that asks for information
- Excel client retrieves data from a database
- Excel client retrieves data from its own datasheets

Layered system:

- The client cannot tell if it is connected directly to the *ultimate* server holding the data or to some intermediary.

REST Principles (2)

Uniform Interface

- Identification of Resources: **Each data point has a URL**. The client can get the data from the URL in different formats (XML, JSON)
- The client can manipulate the data through the URL (CRUD operations). Most current work on API's is about accessing databases via REST methods
- Self-descriptive messages (Media type MIME)
- Hypermedia Links: **I can find what to do with the data by following links embedded in the metadata returned from the server**

REST Principles (3)

Stateless:

- Client context is not stored in the server. All the information for a request is send every time from the client. This provides resilience. If the connection drops, you just start again!

Cacheable:

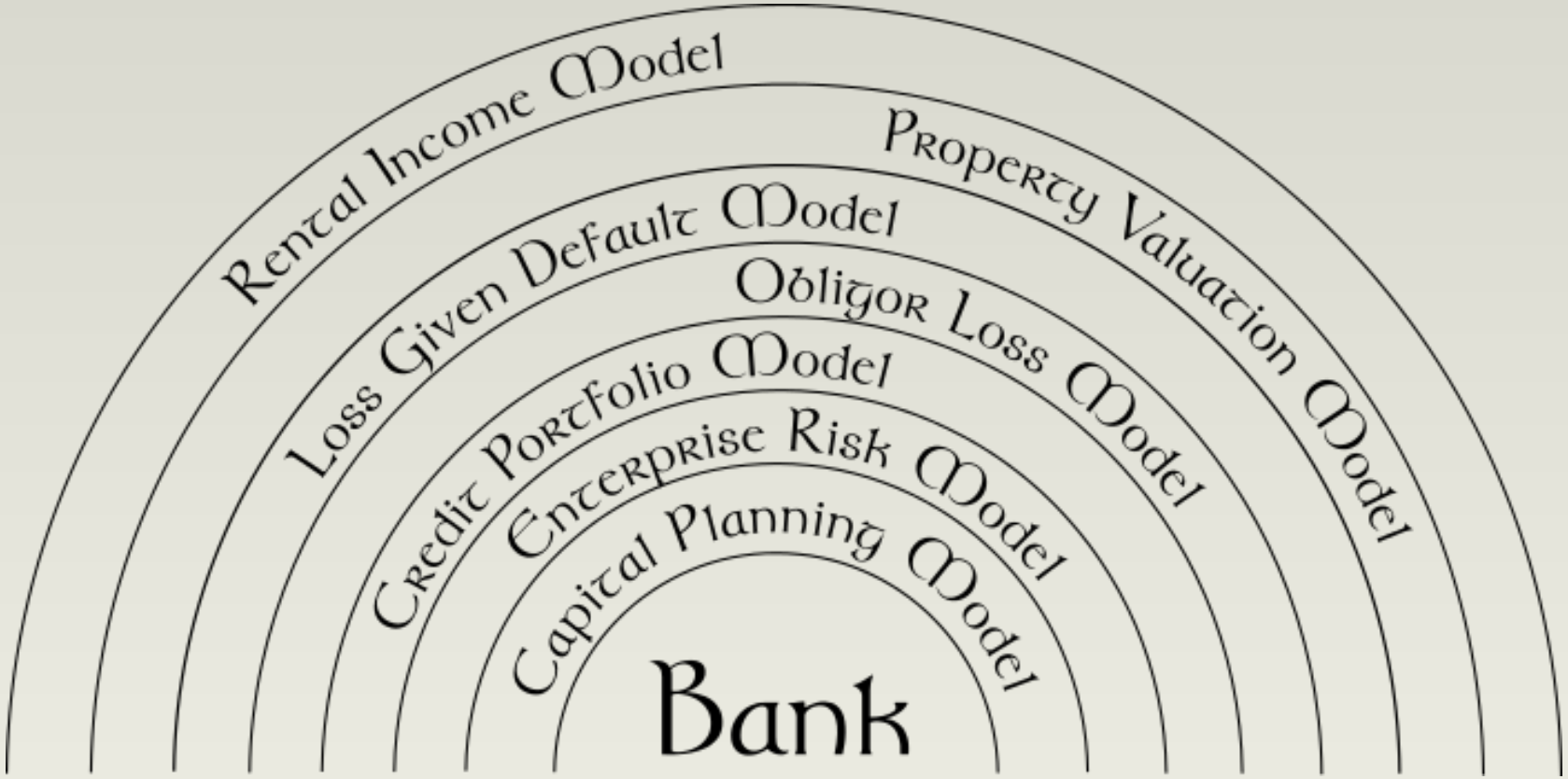
- Server responses must define themselves as cacheable or not

Code on demand:

- If the client needs code to present the data, this can be send by the server together with the data. **All modern web dev is based on this!**

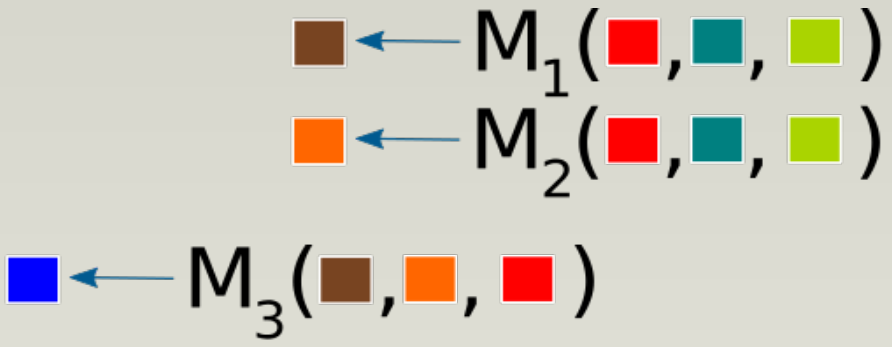
Linked Models
(Open Risk API)

Seven Heavens of Model Complexity

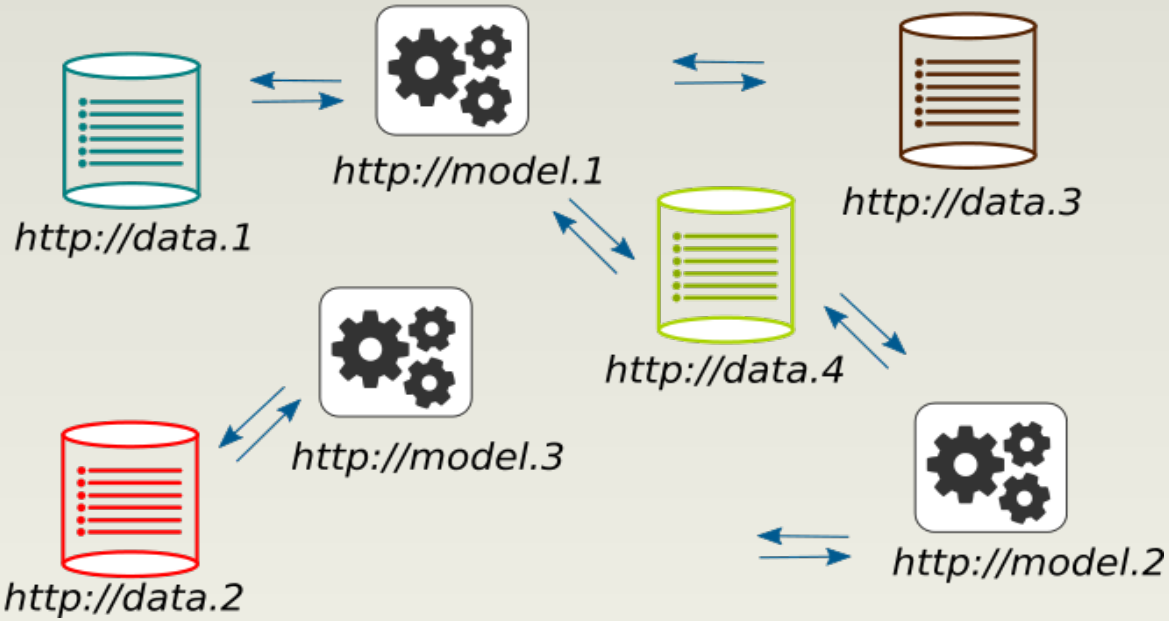


Linked Models

Mathematical Representation



Network Representation



API core elements

A resource is either: a REST data service, or a REST model server

- A Data service can hold input risk data, output risk data etc
- A Model server is where model instances can be accessed

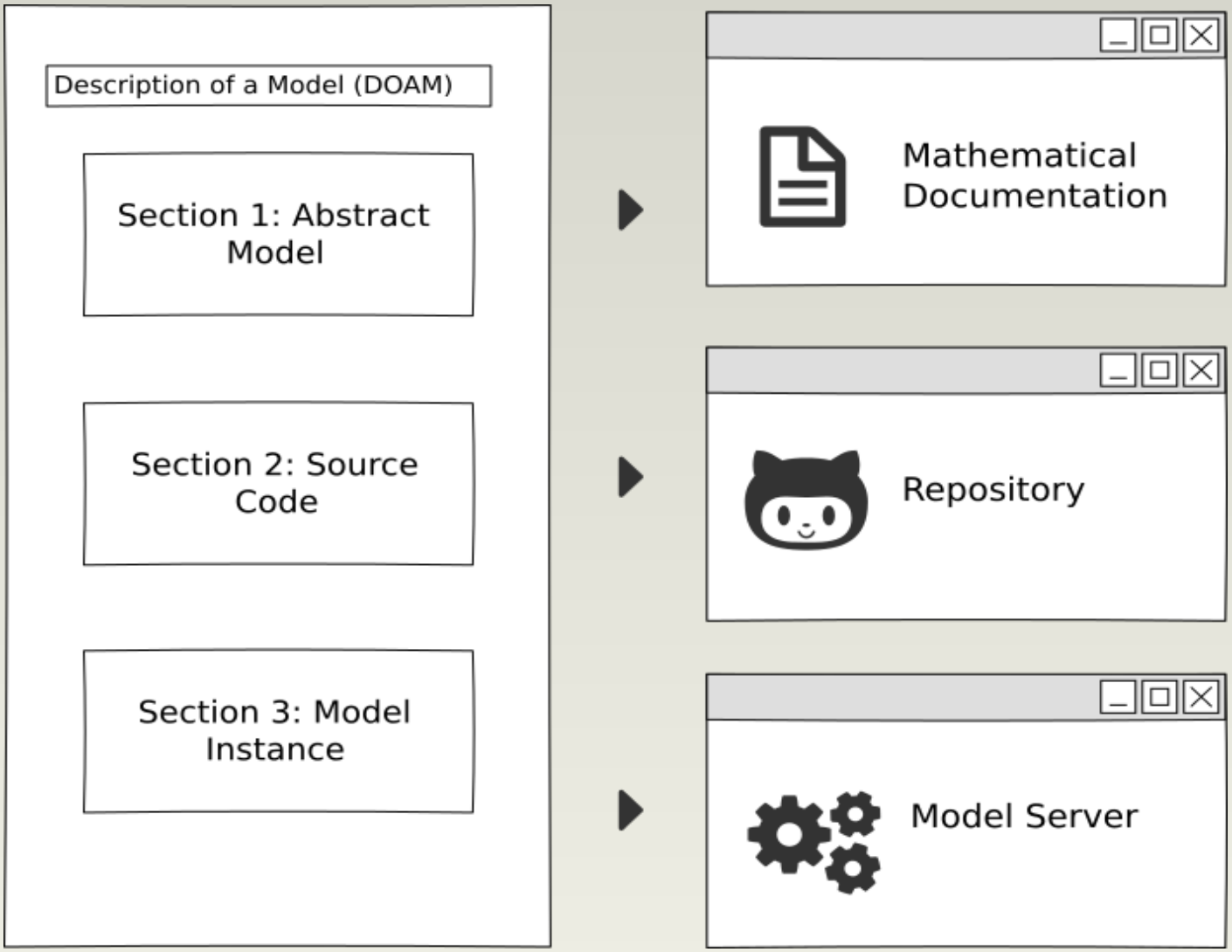
Workflows: Sequences of operations with documented inputs, outputs and models

Layering and Hypermedia:

- Each model knows about its direct data inputs / outputs
- Each data point knows about its direct producer

Description of a Model (DOAM)

Three Meanings of “Model”



Description of a Model (Key Details)

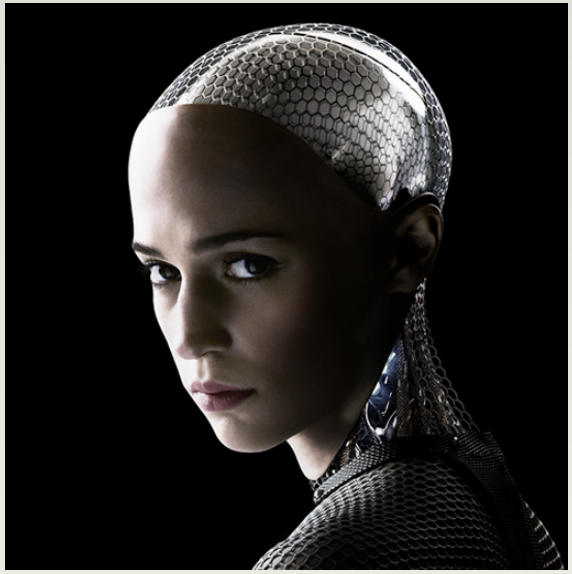
All Three Components are linked to URL's

But: Abstract Model Documentation and Source Code are a type of *static* metadata

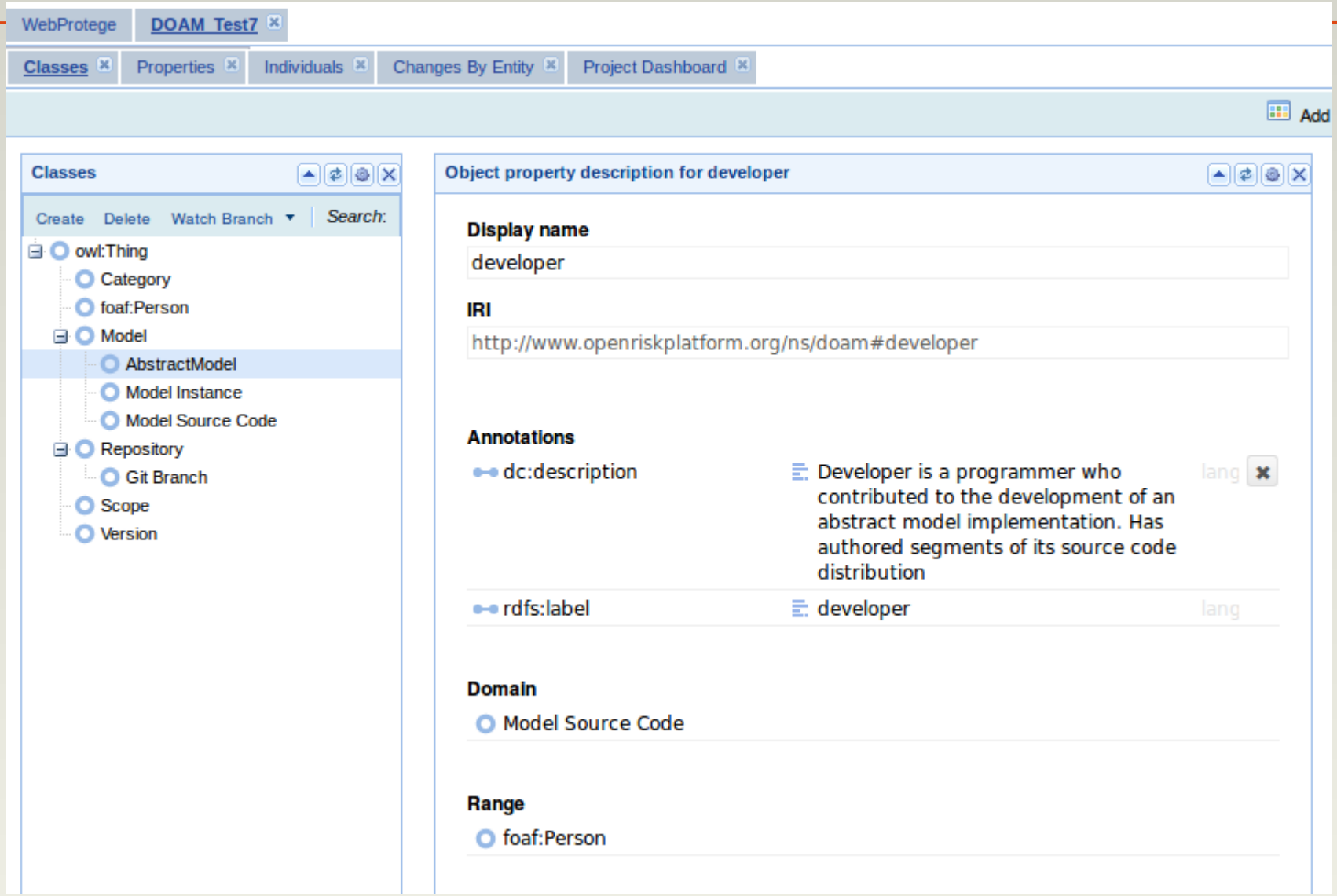
The *active* component is the Model Instance.

You can ask the Model Instance:

- where are you *documented*?
- where is the *code* that your are executing?



DOAM on WebProtege (Stanford Ontology Editor)



Example API Messages


```
[
  {
    "@id": "http://127.0.0.1:5012/models/shannon",
    "@type":  ["http://openriskplatform.org/ns/doam#model"],
    "http://openriskplatform.org/ns/doam#hasInput":
         [{ "@id": "http://127.0.0.1:5011/obligors/" }],
    "http://openriskplatform.org/ns/doam#hasOutput":
         [{ "@id": "http://127.0.0.1:5010/results/" }],
    "http://openriskplatform.org/ns/doam#name":
         [{ "@value": "shannon" }],
    "http://xmlns.com/foaf/0.1/mbox":
         [ {"@id": "mailto:models_r_us@example.org" }],
    "http://xmlns.com/foaf/0.1/name":
         [ {"@value": "shannon" }]
  }
]
```

Example:
Open Risk Dashboard

Open Risk Dashboard

Interactive, visual, online resource that supports financial risk management tasks for users across the EU area.

Open data retrieved from the European Central Bank Statistical Data Warehouse

Open source FIWARE software technologies (big data backends, security and much more)

The application retrieves on demand statistical data from the warehouse and provides value enhancing visualization and risk modeling layers

It can be combined with private data for portfolio risk analysis

Currently in development, demo based on alpha version

FI-WARE
Open APIs for Open Minds

Remote source: https://sdw-wsrest.ecb.europa.eu/

EUROPEAN CENTRAL BANK

EUROSYSTEM

ECB SDMX 2.1 RESTful web service

Home | Overview | Data | Metadata | Schemas | Content negotiation | Status codes | Useful tips

sdmx

Statistical Data and Metadata eXchange

The ECB SDMX 2.1 RESTful web service offers programmatic access to the statistical data and metadata disseminated via the ECB Statistical Data Warehouse.

It offers two modes of operation:

- **Data retrieval**: You know the data you want to retrieve (e.g.: daily exchange rates of the Japanese yen against the euro).
- **Data discovery**: Using a metadata-driven approach, you want to discover the data exposed by the web service.

Data Web EndPoints Directory

Data EndPoint Manager

Select from the list of stored Data EndPoints or add a new entry

Show 10 **entries** **Search:**

DataFlow ID	DataFlow Name	Action 1	Action 2	Action 3	Action 4
1	EONIA Rate	Explore	Edit	Delete	Clone
2	Austria GDP	Explore	Edit	Delete	Clone
3	Belgium GDP	Explore	Edit	Delete	Clone
5	Bulgaria GDP	Explore	Edit	Delete	Clone
6	Croatia GDP	Explore	Edit	Delete	Clone
7	Cyprus GDP	Explore	Edit	Delete	Clone
8	Czech GDP	Explore	Edit	Delete	Clone
9	Denmark GDP	Explore	Edit	Delete	Clone
10	Estonia GDP	Explore	Edit	Delete	Clone
11	Finland GDP	Explore	Edit	Delete	Clone

Showing 1 to 10 of 32 entries

First Previous 1 2 3 4 Next Last

Create New Data EndPoint

HTTP GET request:

Fetch Data from EndPoint

Name: Austria GDP

URL: http://sdw-wsrest.ecb.europa.eu/service/data/MNA/Q.Y.AT.W2.S1.S1.B.B1GQ._Z._Z._Z.XDC.LR.GY

Local Storage: MNA.Q.Y.AT.W2.S1.S1.B.B1GQ._Z._Z._Z.XDC.LR.GY/2015-11-09-13:08:07

Result: {"Dates": ["1997-Q1", "1997-Q2", "1997-Q3", "1997-Q4", "1998-Q1", "1998-Q2", "1998-Q3", "1998-Q4", "1999-Q1", "1999-Q2", "1999-Q3", "1999-Q4", "2000-Q1", "2000-Q2", "2000-Q3", "2000-Q4", "2001-Q1", "2001-Q2", "2001-Q3", "2001-Q4", "2002-Q1", "2002-Q2", "2002-Q3", "2002-Q4", "2003-Q1", "2003-Q2", "2003-Q3", "2003-Q4", "2004-Q1", "2004-Q2", "2004-Q3", "2004-Q4", "2005-Q1", "2005-Q2", "2005-Q3", "2005-Q4", "2006-Q1", "2006-Q2", "2006-Q3", "2006-Q4", "2007-Q1", "2007-Q2", "2007-Q3", "2007-Q4", "2008-Q1", "2008-Q2", "2008-Q3", "2008-Q4", "2009-Q1", "2009-Q2", "2009-Q3", "2009-Q4", "2010-Q1", "2010-Q2", "2010-Q3", "2010-Q4", "2011-Q1", "2011-Q2", "2011-Q3", "2011-Q4", "2012-Q1", "2012-Q2", "2012-Q3", "2012-Q4", "2013-Q1", "2013-Q2", "2013-Q3", "2013-Q4", "2014-Q1", "2014-Q2", "2014-Q3", "2014-Q4", "2015-Q1", "2015-Q2", "2015-Q3"], "Values": [2.217957415010541, 1.4952646670360448, 2.603915391251621, 3.4844420557526234, 3.8711806975562757, 4.816067706322502, 3.2601678103644227, 2.2466063440008677, 2.7618061150113116, 2.582388491133547, 3.7910056900912315, 4.529280841148431, 3.762514607046197, 4.077648080309015, 3.3164700637610567, 3.7271910080763915, 3.1857890383549137, 1.4701240521921344, 0.7509293902191683, -0.2630640558786479, 0.8503299737073533, 1.651648796543892, 1.9188824231800572, 1.4907087041831746, 0.9738746651605235, 0.5395810265205636, 0.8221220610957536, 1.0615049098469553, 1.8212636717391595, 2.551019608043159, 2.805661175943941, 2.43766755916599, 1.8116357307202957, 2.2896292873054724,

Filter: http://.../sdw_data/?where={“identifier”:”MNA.Q.Y.AT...”}

p://192.168.2.4/data/sdw_data/?where={"identifier":"MNA.Q.Y.AT.W2.S1.S1.B.B1GQ._Z._Z._Z.XDC.LR.GY/2015-11-09-16:4

Explore

Outputs & Graphs

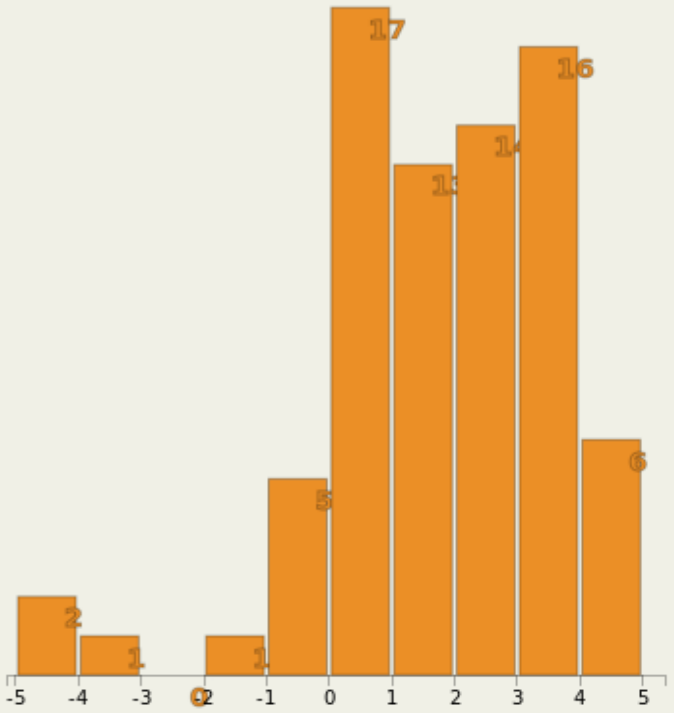
Data

Show 10 **entries** **Search:**

Date	Value
1997-Q1	2.217957415010541
1997-Q2	1.4952646670360448
1997-Q3	2.603915391251621
1997-Q4	3.4844420557526234
1998-Q1	3.8711806975562757
1998-Q2	4.816067706322502
1998-Q3	3.2601678103644227
1998-Q4	2.2466063440008677
1999-Q1	2.7618061150113116
1999-Q2	2.582388491133547

Showing 1 to 10 of 75 entries Previous12345...8Next

Histogram



Portfolio Explorer

Enter a Portfolio ID to explore its data (e.g., http://192.168.2.4/data/obligors?page=20)

http://192.168.2.4/data/obligors?page=20

Explore

Portfolio Data

Show 10 **entries** **Search:**

Obligor ID	EAD	PD	LGD
55585d88a738bd73a9622eaf	0.7329055040609092	0.4729960784316063	0.37852625572122633
55585d88a738bd73a9622eb0	0.689321905374527	0.1245687345508486	0.6736927803140134
55585d88a738bd73a9622eb1	0.18958246265538037	0.15601618238724768	0.911496989428997
55585d88a738bd73a9622eb2	0.27868858398869634	0.8853766347747296	0.8967847297899425
55585d88a738bd73a9622eb3	0.41706046648323536	0.9074420728720725	0.04495105077512562
55585d88a738bd73a9622eb4	0.483097686432302	0.9074658222962171	0.17864870862104
55585d88a738bd73a9622eb5	0.8165187507402152	0.30357093061320484	0.7600250744726509
55585d88a738bd73a9622eb6	0.22635430982336402	0.5878760376945138	0.6005772473290563
55585d88a738bd73a9622eb7	0.40399777819402516	0.7092708968557417	0.415051968768239
55585d88a738bd73a9622eb8	0.2595019252039492	0.26533142081461847	0.0812034762930125

Showing 1 to 10 of 25 entries

First Previous 1 2 3 Next Last

Model Directory

Select an API EndPoint to explore its model catalog

Testing 1: http://192.168.2.4/api

Explore

Show 10 entries

Search:

Model Name	Description	Model URL
Concentration Ratio	Calculation of the Concentration Ratio	http://192.168.2.4/api/cr
Correlation	Calculation of Standard Correlation between two timeseries	http://192.168.2.4/api/correlation
Credit Simulation	Calculation of Portfolio Credit Loss	http://192.168.2.4/api/creditsim
Gini Calculator	Calculation of the Gini Concentration Index	http://192.168.2.4/api/gini
Hannah Kay Calculator	Calculation of the Hannah Kay Concentration Index	http://192.168.2.4/api/hk
HHI Calculator	Calculation of the HHI Concentration Index	http://192.168.2.4/api/hhi
Shannon Calculator	Calculation of the Shannon Concentration Index	http://192.168.2.4/api/shannon

Model Explorer

Enter a Linked Model URL to explore its properties

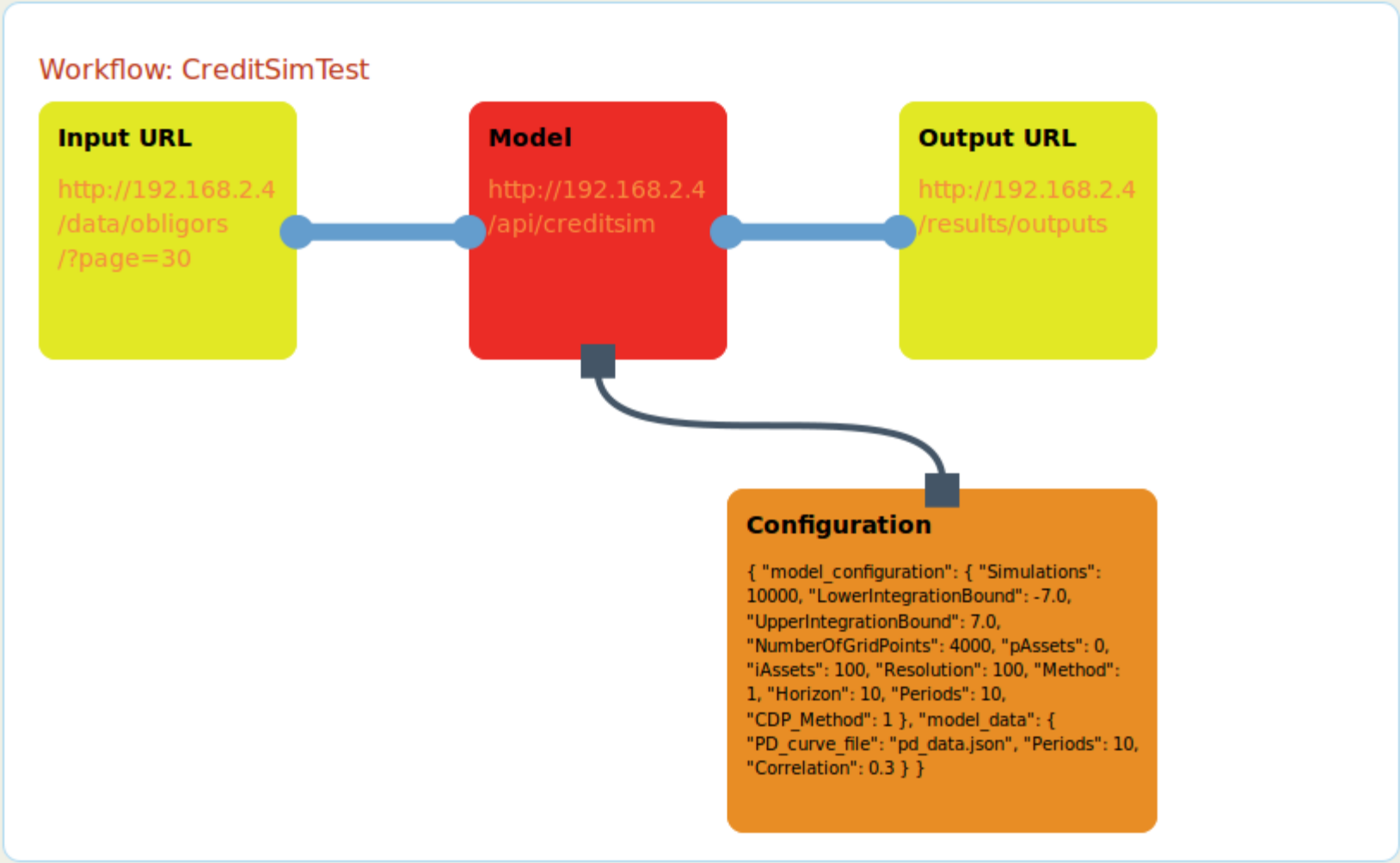
http://192.168.2.4/api/creditsim

Explore

Model Data


```
{
        "@context": {
                "doam": "http://www.openriskplatform.org/ns/doam#"
        },
        "@graph": [
                {
                        "@id": "_:Nfa324b06c54c4f5fbfff0604b4e164c4",
                        "@type": "doam:Version",
                        "doam:created": "2015-01-01",
                        "doam:name": "unstable",
                        "doam:revision": "0.0.1"
                },
                {
                        "@id": "_:N2d182e734e8c472f8ccdb1ba9772d2ec",
                        "@type": "http://xmlns.com/foaf/0.1/Person",
                        "http://xmlns.com/foaf/0.1/homepage": {
                                "@id": "https://www.openrisk.eu"
                        },
```

Workflows (Credit Portfolio Simulation)



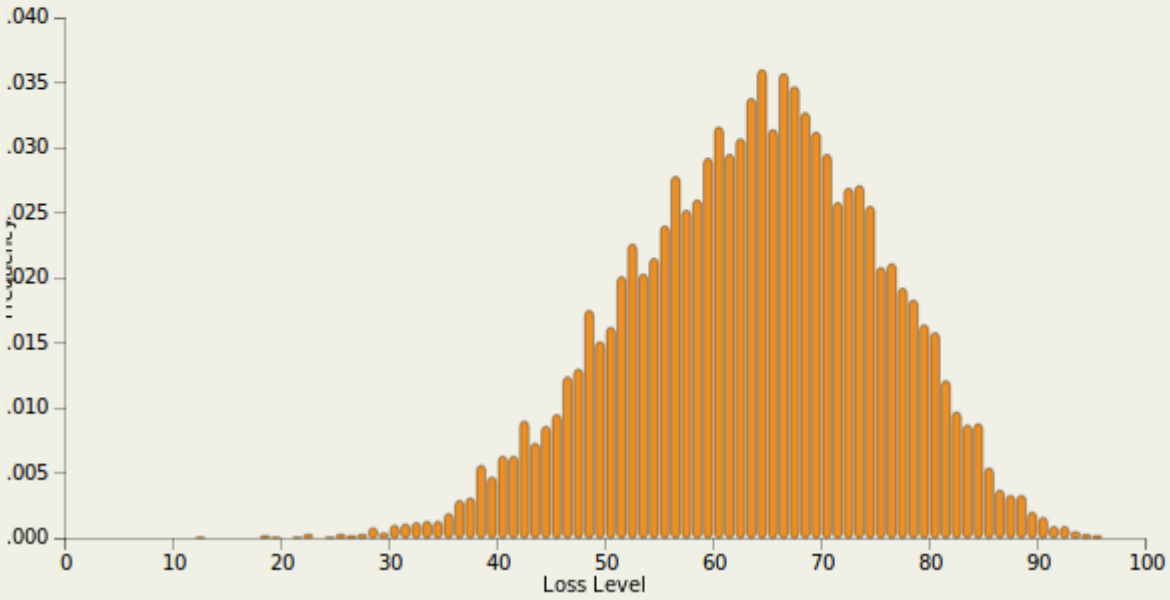
Results (Credit Loss Distribution)

http://192.168.2.4/results/outputs/56421118e9d780148a529dd5

Explore

Outputs & Graphs

Histogram



Summary

- Data formats are evolving rapidly (json, json-ld)
- Powerful Semantic Web tools facilitate data standard development (rdf graphs, owl and ontology editors)
- Best practice IT architectures coalesce around loosely coupled client/server and REST API's
- Risk management post crisis requires fixing both the risk data AND the risk model mess
- The Open Risk API integrates linked risk data and linked risk models in a well documented, *future ready,* framework

Open Risk Dashboard: We demonstrated retrieving and processing macro data and calculating portfolio risk measures using web components conforming to the API

Thanks for the attention!

OpenRisk

An independent provider of training and risk analysis tools to the broader financial services community

Learn more about our mission at: www.openrisk.eu