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
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 80bIn annual spend on Risk IT?
“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
<table>
<thead>
<tr>
<th>Date</th>
<th>Open</th>
<th>High</th>
<th>Low</th>
<th>Close</th>
<th>Volume</th>
<th>Adj Close</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015-11-06</td>
<td>2098.600098</td>
<td>2101.909912</td>
<td>2083.73999</td>
<td>2099.199951</td>
<td>4369020000</td>
<td>2099.199951</td>
</tr>
<tr>
<td>2015-11-05</td>
<td>2101.679932</td>
<td>2108.780029</td>
<td>2090.409912</td>
<td>2099.929932</td>
<td>4051890000</td>
<td>2099.929932</td>
</tr>
<tr>
<td>2015-11-04</td>
<td>2110.600098</td>
<td>2114.590088</td>
<td>2096.97998</td>
<td>2102.310059</td>
<td>4078870000</td>
<td>2102.310059</td>
</tr>
<tr>
<td>2015-11-03</td>
<td>2102.629883</td>
<td>2116.47998</td>
<td>2097.51001</td>
<td>2109.790039</td>
<td>4272060000</td>
<td>2109.790039</td>
</tr>
<tr>
<td>2015-11-02</td>
<td>2080.76001</td>
<td>2106.199951</td>
<td>2080.76001</td>
<td>2104.050049</td>
<td>3760020000</td>
<td>2104.050049</td>
</tr>
<tr>
<td>2015-10-30</td>
<td>2090</td>
<td>2094.320068</td>
<td>2079.340088</td>
<td>2079.360107</td>
<td>4256200000</td>
<td>2079.360107</td>
</tr>
<tr>
<td>2015-10-29</td>
<td>2088.350098</td>
<td>2092.52002</td>
<td>2082.629883</td>
<td>2089.409912</td>
<td>4008940000</td>
<td>2089.409912</td>
</tr>
<tr>
<td>2015-10-28</td>
<td>2066.47998</td>
<td>2090.350098</td>
<td>2063.110107</td>
<td>2090.350098</td>
<td>4698110000</td>
<td>2090.350098</td>
</tr>
</tbody>
</table>

What is the currency?
Who or what collected the data?
What does “Open” mean?
Is the accuracy spurious?
...
<?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
{
    "quote": {
        "symbol": "BHP.AX",
        "Ask": "22.75",
        "AverageDailyVolume": "9521200",
        "Bid": "22.65",
        "BookValue": "12.17"
    },
    "Names": ["BHP", "Biliton"]
}
```
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)


[ 

  {
    "@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
<message:GenericData
xsi:schemaLocation="http://www.sdmx.org/resources/sdmxml/schemas/v2_1/message
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:DataSet action="Replace" validFromDateTime="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.
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

$M_1$, $M_2$, $M_3$

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”

Section 1: Abstract Model

Section 2: Source Code

Section 3: Model Instance

Mathematical Documentation

Repository

Model Server
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)

[Image: A screenshot of the WebProtege interface showing the classes and object property description for the developer concept. The display name is 'developer', the IRI is http://www.openriskplatform.org/ns/doam#developer. The annotations include a description of the developer as a programmer who contributed to the development of an abstract model implementation and has authored segments of its source code distribution. The domain is Model Source Code, and the range is foaf:Person.]
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
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

![Data EndPoint Manager Table]

<table>
<thead>
<tr>
<th>DataFlow ID</th>
<th>DataFlow Name</th>
<th>Action 1</th>
<th>Action 2</th>
<th>Action 3</th>
<th>Action 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>EONIA Rate</td>
<td>Explore</td>
<td>Edit</td>
<td>Delete</td>
<td>Clone</td>
</tr>
<tr>
<td>2</td>
<td>Austria GDP</td>
<td>Explore</td>
<td>Edit</td>
<td>Delete</td>
<td>Clone</td>
</tr>
<tr>
<td>3</td>
<td>Belgium GDP</td>
<td>Explore</td>
<td>Edit</td>
<td>Delete</td>
<td>Clone</td>
</tr>
<tr>
<td>5</td>
<td>Bulgaria GDP</td>
<td>Explore</td>
<td>Edit</td>
<td>Delete</td>
<td>Clone</td>
</tr>
<tr>
<td>6</td>
<td>Croatia GDP</td>
<td>Explore</td>
<td>Edit</td>
<td>Delete</td>
<td>Clone</td>
</tr>
<tr>
<td>7</td>
<td>Cyprus GDP</td>
<td>Explore</td>
<td>Edit</td>
<td>Delete</td>
<td>Clone</td>
</tr>
<tr>
<td>8</td>
<td>Czech GDP</td>
<td>Explore</td>
<td>Edit</td>
<td>Delete</td>
<td>Clone</td>
</tr>
<tr>
<td>9</td>
<td>Denmark GDP</td>
<td>Explore</td>
<td>Edit</td>
<td>Delete</td>
<td>Clone</td>
</tr>
<tr>
<td>10</td>
<td>Estonia GDP</td>
<td>Explore</td>
<td>Edit</td>
<td>Delete</td>
<td>Clone</td>
</tr>
<tr>
<td>11</td>
<td>Finland GDP</td>
<td>Explore</td>
<td>Edit</td>
<td>Delete</td>
<td>Clone</td>
</tr>
</tbody>
</table>

Showing 1 to 10 of 32 entries

[Create New Data EndPoint]
HTTP GET request:

Fetch Data from EndPoint

Name: Austria GDP

**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

**Portfolio Data**

<table>
<thead>
<tr>
<th>Obligor ID</th>
<th>EAD</th>
<th>PD</th>
<th>LGD</th>
</tr>
</thead>
<tbody>
<tr>
<td>55585d88a738bd73a9622ead</td>
<td>0.7329055040609092</td>
<td>0.4729960784316063</td>
<td>0.37852625572122633</td>
</tr>
<tr>
<td>55585d88a738bd73a9622eb0</td>
<td>0.689321905374527</td>
<td>0.1245687345508486</td>
<td>0.6736927803140134</td>
</tr>
<tr>
<td>55585d88a738bd73a9622eb1</td>
<td>0.18958246265538037</td>
<td>0.15601618238724768</td>
<td>0.911496989428997</td>
</tr>
<tr>
<td>55585d88a738bd73a9622eb2</td>
<td>0.27868858398869634</td>
<td>0.8853766347747296</td>
<td>0.8967847297899425</td>
</tr>
<tr>
<td>55585d88a738bd73a9622eb3</td>
<td>0.41706046648323536</td>
<td>0.9074420728720725</td>
<td>0.04495105077512562</td>
</tr>
<tr>
<td>55585d88a738bd73a9622eb4</td>
<td>0.483097686432302</td>
<td>0.907465822962171</td>
<td>0.17864870862104</td>
</tr>
<tr>
<td>55585d88a738bd73a9622eb5</td>
<td>0.8165187507402152</td>
<td>0.30357093061320484</td>
<td>0.7600250744726509</td>
</tr>
<tr>
<td>55585d88a738bd73a9622eb6</td>
<td>0.22635430982336402</td>
<td>0.5878760376945138</td>
<td>0.6005772473290563</td>
</tr>
<tr>
<td>55585d88a738bd73a9622eb7</td>
<td>0.40399777819402516</td>
<td>0.7092708968557417</td>
<td>0.415051968768239</td>
</tr>
<tr>
<td>55585d88a738bd73a9622eb8</td>
<td>0.2595019252039492</td>
<td>0.26533142081461847</td>
<td>0.0812034762930125</td>
</tr>
</tbody>
</table>
## Model Directory

Select an API EndPoint to explore its model catalog

Testing 1: http://192.168.2.4/api

---

<table>
<thead>
<tr>
<th>Model Name</th>
<th>Description</th>
<th>Model URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentration Ratio</td>
<td>Calculation of the Concentration Ratio</td>
<td><a href="http://192.168.2.4/api/cr">http://192.168.2.4/api/cr</a></td>
</tr>
<tr>
<td>Correlation</td>
<td>Calculation of Standard Correlation between two timeseries</td>
<td><a href="http://192.168.2.4/api/correlation">http://192.168.2.4/api/correlation</a></td>
</tr>
<tr>
<td>Credit Simulation</td>
<td>Calculation of Portfolio Credit Loss</td>
<td><a href="http://192.168.2.4/api/creditsim">http://192.168.2.4/api/creditsim</a></td>
</tr>
<tr>
<td>Gini Calculator</td>
<td>Calculation of the Gini Concentration Index</td>
<td><a href="http://192.168.2.4/api/gini">http://192.168.2.4/api/gini</a></td>
</tr>
<tr>
<td>Hannah Kay Calculator</td>
<td>Calculation of the Hannah Kay Concentration Index</td>
<td><a href="http://192.168.2.4/api/hk">http://192.168.2.4/api/hk</a></td>
</tr>
<tr>
<td>HHI Calculator</td>
<td>Calculation of the HHI Concentration Index</td>
<td><a href="http://192.168.2.4/api/hhi">http://192.168.2.4/api/hhi</a></td>
</tr>
<tr>
<td>Shannon Calculator</td>
<td>Calculation of the Shannon Concentration Index</td>
<td><a href="http://192.168.2.4/api/shannon">http://192.168.2.4/api/shannon</a></td>
</tr>
</tbody>
</table>
Model Explorer

Enter a Linked Model URL to explore its properties

http://192.168.2.4/api/creditsim

Explore

Model Data

```json
{
   "@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"
         }
      }
   ],
   "http://xmlns.com/foaf/0.1/name": "OpenRisk"
}
```
Workflow: CreditSimTest

Input URL:
http://192.168.2.4/data/obligors/?page=30

Model:
http://192.168.2.4/api/creditsim

Output URL:
http://192.168.2.4/results/outputs

Configuration:

```json
{
    "model_configuration": {
        "Simulations": 10000,
        "LowerIntegrationBound": -7.0,
        "UpperIntegrationBound": 7.0,
        "NumberOfGridPoints": 4000,
        "pAssets": 0,
        "iAssets": 100,
        "Resolution": 100,
        "Method": 1,
        "Horizon": 10,
        "Periods": 10,
        "CDP_Method": 1,
        "model_data": {
            "PD_curve_file": "pd_data.json",
            "Periods": 10,
            "Correlation": 0.3
        }
    }
}
```

Calculate
Results (Credit Loss Distribution)

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