Sharing Data, Hiding Complexity with RDF and Clojure

> Malcolm Sparks James Henderson

TechMesh 2012

◆□▶ ◆□▶ ◆臣▶ ◆臣▶ 臣 のへで

Financial, regulatory pressures in banking are presenting IT with new challenges.

(日本)(同本)(日本)(日本)(日本)

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

Financial, regulatory pressures in banking are presenting IT with new challenges.

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ - 三 - のへで

How can we process data :-

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

Financial, regulatory pressures in banking are presenting IT with new challenges.

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ - 三 - のへで

How can we process data :-

Faster (as always)

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

Financial, regulatory pressures in banking are presenting IT with new challenges.

◆□▶ ◆□▶ ◆□▶ ◆□▶ ●□

How can we process data :-

- Faster (as always)
- At greater volumes (risk sensitivities HistSim)

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

Financial, regulatory pressures in banking are presenting IT with new challenges.

How can we process data :-

- Faster (as always)
- At greater volumes (risk sensitivities HistSim)
- More precisely (finer granualarity Dodd Frank initial margin)

(日本)(同本)(日本)(日本)(日本)

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

"Organizations which design systems ... are constrained to produce designs which are copies of the communication structures of these organizations"

Mel Conway (1968)

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

Conway in action



Figure : An example communication structure

Q: Do Bob and Alice get the same view of the data?

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

What does it all mean?

CommonReference Narrative Numericld LegalEntity ManagementArea GLRegion LocalBankingTrading AggregationLevel SettlementEngineSource ISIN CUSIP PrivatePlacement OtherSecurityId FinancialProductType ReportingCurve InternalExternalFlag BuySell Fas133Flag SpvId MtnId LinkageReference LinkageReference2 NodeRef ReasonCode YNFlag

Observation: Data doesn't travel well

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

What went wrong?



Conclusion: We need our data models to span across our systems.

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

As we scale up :-

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

・ロト・西ト・ヨト・ヨー もんの

As we scale up :-

Complexity increases

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

・ロト・西ト・ヨト・ヨー うくぐ

As we scale up :-

- Complexity increases
- Deviations increase

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

▲□▶ ▲□▶ ▲三▶ ▲三▶ 三三 のへで

As we scale up :-

- Complexity increases
- Deviations increase
- Concepts become more abstract

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

・ロト・日本・日本・日本・日本・日本

As we scale up :-

- Complexity increases
- Deviations increase
- Concepts become more abstract
- Understanding decreases

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

As we scale up :-

- Complexity increases
- Deviations increase
- Concepts become more abstract
- Understanding decreases

As data escapes its domain model :-

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

As we scale up :-

- Complexity increases
- Deviations increase
- Concepts become more abstract
- Understanding decreases

As data escapes its domain model :-

It loses assumptions (meaning)

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

As we scale up :-

- Complexity increases
- Deviations increase
- Concepts become more abstract
- Understanding decreases

As data escapes its domain model :-

- It loses assumptions (meaning)
- Local identifiers fail to become global ones.

(日本)(同本)(日本)(日本)(日本)

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

Implications

 Our primary tool for flexible data processing (data modelling) does not scale! Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

◆□▶ ◆□▶ ◆臣▶ ◆臣▶ 臣 のへぐ

Implications

- Our primary tool for flexible data processing (data modelling) does not scale!
- We are fire-fighting this problem with a code hosepipe!

・ロト ・ 日 ・ ・ 日 ・ ・ 日 ・ ・ つ へ ()

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

Introduction to RDF



```
http://animals.com/ginger
http://actions.com/satOn http://tech-mesh-hotel.co.uk/mat ;
rdf:type http://animals.com/Cat ;
```

```
http://actions.com/satOn
    rdfs:label "sat on" ;
    rdfs:comment "The action of sitting on a given object." ;
```

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

RDF features

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

・ロト・日本・日本・日本・日本・日本

RDF features

URIs - globally unique identification and transparency

▲□▶ ▲□▶ ▲三▶ ▲三▶ 三三 のへで

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

URIs - globally unique identification and transparency

・ロト ・ 日 ・ ・ 日 ・ ・ 日 ・ ・ つ へ ()

Namespaces (for federated governance)

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

RDF features

URIs - globally unique identification and transparency

・ロト ・ 日 ・ ・ 日 ・ ・ 日 ・ ・ つ へ ()

- Namespaces (for federated governance)
- Storage and transfer neutrality

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

RDF features

URIs - globally unique identification and transparency

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ - 三 - のへで

- Namespaces (for federated governance)
- Storage and transfer neutrality
- Metadata

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

Observations

RDF has the power to break silos

RDF is an non-disruptive way of applying federated data governance

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

・ロト ・ 日下 ・ 日下 ・ 日下 ・ 今日 ト

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

ata Loses Fidelity s You Move It

RDF and the RAP

The Power of Graphs

・ロト・日本・日本・日本・日本・日本

Exposing the data in the 'ODS'

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

▲□▶ ▲□▶ ▲ □▶ ▲ □▶ ▲□ ● ● ● ●

- Exposing the data in the 'ODS'
- The tables in the ODS are built from a meta-model

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

・ロト ・日 ・ ・ ヨ ・ ・ ヨ ・ うへぐ

- Exposing the data in the 'ODS'
- The tables in the ODS are built from a meta-model

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

We get the meta-model in XML format

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

So why do we transform it to RDF?

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

So why do we transform it to RDF?

Disambiguation of terms

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

・ロト ・日 ・ ・ ヨ ・ ・ ヨ ・ うへぐ

So why do we transform it to RDF?

- Disambiguation of terms
- Merging different data sources

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

・ロト・日本・日本・日本・日本・日本

Exposing a data dictionary



Dictionaries

RAP Dictionary: 'Legal Entity'

Description:

A special type of Organisation that represents an entity that can be sued if it does not meet its obligations

Attributes:

None

Out-bound relationships:

GRC Currency (to Currency) LFC Currency (to Currency)

In-bound relationships:

Legal Entity (from Trader Book) Legal Entity (from Party)

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

◆□▶ ◆□▶ ◆∃▶ ◆∃▶ → ヨ → のへぐ

Forming a View

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

ata Loses Fidelity s You Move It

RDF and the RAP

The Power of Graphs

・ロト・日本・日本・日本・日本・日本

Forming a View

A View in graph form:



Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

Forming a View

RDF storage of a view:

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

・ロト ・ 日・ ・ 田・ ・ 日・ ・ 日・

Delivering the View



Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

・ロト・西下・日下・日下・日下

The Power of Graphs

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

ata Loses Fidelity s You Move It

RDF and the RAP

The Power of Graphs

くりゃく 聞き ふぼや ふぼや ふりゃく

Storing config in RDF

<hosts/p21> <instance> <instances/PROD1> .

<hosts/u11> <instance> <instances/UAT2> .

<hosts/d30> <instance> <instances/SIT1> .

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

Storing config in RDF

<hosts/p21> cmdb:hostName "p21.db.com" .

<instances/PROD1> <service> <services/PROD1/http> .

・ロト ・ 日 ・ ・ 日 ・ ・ 日 ・ ・ つ へ ()

<services/PROD1/http> cmdb:port 8080 .

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

◆□▶ ◆□▶ ◆目▶ ◆目▶ ◆□▶

- Disambiguation
- Meta-data

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

)ata Loses Fidelity Is You Move It

RDF and the RAP

The Power of Graphs

・ロト・日本・日本・日本・日本・日本

- Disambiguation
- Meta-data
- Querying the config

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

・ロト・日本・日本・日本・日本・日本

- Disambiguation
- Meta-data
- Querying the config
- Multiple config sources

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

(日本)(同本)(日本)(日本)(日本)

Graph-Zip: https://github.com/james-henderson/graph-zip

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

・ロト ・ 日 ・ ・ 日 ・ ・ 日 ・ ・ つ へ ()

An example config graph:



Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

▲□▶ ▲□▶ ▲三▶ ▲三▶ 三三 のへの

Importing the graph:

(def config-map (build-in-memory-graph [{:subject "<hosts/p21>" :property :hostname :object "p21.db.com"} {:subject "<hosts/p21>" :property :instance :object "<instances/PROD1>"}

{:subject "<instances/PROD1>" :property :service :object "<services/PROD1/swank>"} {:subject "<services/PROD1/swank>" :property :port :object "4005"} {:subject "<services/PROD1/swank>" :property :rdf/type :object "<SwankService>"}

```
{:subject "<instances/PROD1>" :property :service :object "<services/PROD1/http>"}
{:subject "<services/PROD1/http>" :property :port :object "8080"}
{:subject "<services/PROD1/http>" :property :binds-to :object "127.0.0.1"}
{:subject "<services/PROD1/http>" :property :df/type :object "<fttpService>"}
```

```
{:subject "<instances/PROD1>" :property :jvm :object "<jvms/PROD1>"}
{:subject "<jvms/PROD1>" :property :jvmMaxHeap :object "2048m"}]))
```

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

An example config graph:



Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

Building the zipper:

(def prod-host-zipper (graph-zip config-map "<hosts/p21>"))

▲□▶ ▲□▶ ▲三▶ ▲三▶ 三三 のへで

An example config graph:



イロト 不得 トイヨト イヨト

3

Navigating the graph:

;; -> ("<services/PROD1/swank>" "<services/PROD1/http>")

Sharing Data, Hiding Compl<u>exity</u>

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

An example config graph:



イロト 不得 トイヨト イヨト

-

Finding out the Swank port:

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

Graph-Zip: https://github.com/james-henderson/graph-zip Neo-Zip: https://github.com/james-henderson/neo-zip Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

◆□▶ ◆□▶ ◆目▶ ◆目▶ ◆□▶

In XML:

<hosts> <host> <hostname>p21.db.com</hostname> <instances> <instance> <label>PROD1</label> <services> <service> <type>swank</type> <port>4005</port> <bindsTo>127.0.0.1</bindsTo> </service> <service> <type>http</type> <port>8080</port> </service> </services> </instance> </instances> </host>

...

</hosts>

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

・ロト ・ 日・ ・ 田・ ・ 日・ ・ 日・

Using traditional zip syntax:

```
(for [host (xml-> doc :host)]
  {:hostname (xml1-> host :hostname text)
    :instances (for [instance (xml-> host :instances :instance)]
        {:label (xml1-> instance :label text)
            :services (for [service (xml-> instance :services :service)]
            {:type (xml1-> service :type text)
            :port (xml1-> service :port text)
            :bindsTo (xml1-> service istindsTo text)}})})
```

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

Using a datalog syntax:

```
(let [db (d/db conn)]
 (for [[host hostname] (d/q '{:find [?h ?name]
                               :where [[?h :type :host]
                                        [?h :hostname ?name]]}
                             db)]
    f:hostname hostname
     :instances (for [[instance label]
                      (d/g '{:find [?inst ?label]
                             :in [$ ?host]
                             :where [[?host :instance ?inst]
                                      [?inst :label ?label]]}
                           db host)]
                  {:label label
                   :services (for [[service type port binds-to]
                                    (d/q '{:find [?service ?type ?port ?binds-to]
                                           :in [$ ?instance]
                                           :where [[?instance :service ?service]
                                                   [?service :port ?port]
                                                   [?service :type ?type]
                                                   [?service :binds-to ?binds-to]]}
                                        db instance)]
                               {:type type
                                :port port
                                 :binds-to binds-to})})))
```

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

An enhanced zipper query:

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

Combining Zippers with RAP Views An enhanced zipper query:

Resulting data structure:

Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP



Sharing Data, Hiding Complexity

Malcolm Sparks James Henderson

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

・ロト・(四ト・(川下・(日下・(日下)))

James Henderson @jhenderson_89 https://github.com/james-henderson

Malcolm Sparks

@malcolmsparks
https://github.com/malcolmsparks

Sharing Data, Hiding Complexity

Malcolm Sparks James Hende<u>rson</u>

Data Loses Fidelity As You Move It

RDF and the RAP

The Power of Graphs

▲□▶ ▲□▶ ▲三▶ ▲三▶ 三三 のへで