



Towards Quality-Driven Integration of Spatial Data

Omar Boucelma

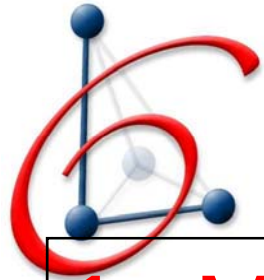
LSIS, Université Paul Cézanne, Aix-Marseille

Credit:

F.-M. Colonna, M. Essid, Y. Lassoued, M. Quafafou, M. Kadiri

Kick-off Meeting STIC-AMSUD

Recife, July 22nd 2008



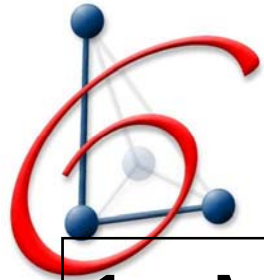
Outline

1. Motivation
2. Starting point: the VirGIS project
3. Reasoning with (Data) Quality
4. Quality-Aware Query Processing
5. Service-Based Quality Approach
6. Conclusion



Motivation

- Build a GIS integration system
 - Leverage of current database/XML technologies
 - Complies with Open GIS consortium recommendations
 - Take into account quality information (metadata)



Outline

1. Motivation
2. Starting point: the VirGIS project
3. Reasoning with (Data) Quality
4. Quality-Aware Query Processing
5. Service-Based Quality Approach
6. Conclusion



Integration Approach

- Mediator/Wrappers
- Global Local Approach View
- Requires two models (mainly)
 - Integration Model
 - Query Model



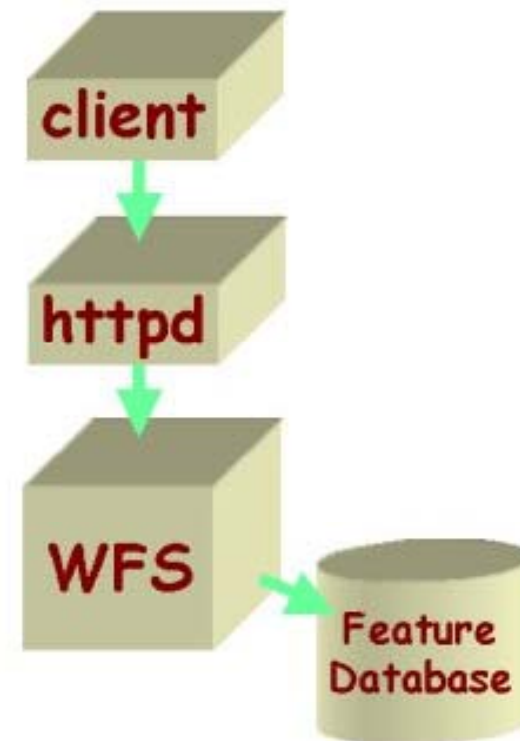
Using OpenGIS Standards

- GML: Geography Markup Language
 - XML encoding of geography
 - Enables one to leverage the whole world of XML technologies.
 - Provides for vector mapping in a standard browser.
 - Enables complex features & feature associations
- Integration choices
 - “GML” as an integration model
 - Which query model ?



WFS: Web Feature Specification

- Direct access to Geographic Features in remote, distributed Datasets
- Standard format - vendor independent
- Queries are XML encoding of SQL_like





WFS Queries (XML/KVP)

- `http://www.myserver.com/wfs.cgi?service=WFS&version=1.0.0&request=GetFeature&typename=ROAD`
- ```
<?xmlversion="1.0"?>
<GetFeature version="1.0.0" service="WFS">
 <Query typeName="ROAD"/>
</GetFeature>
```

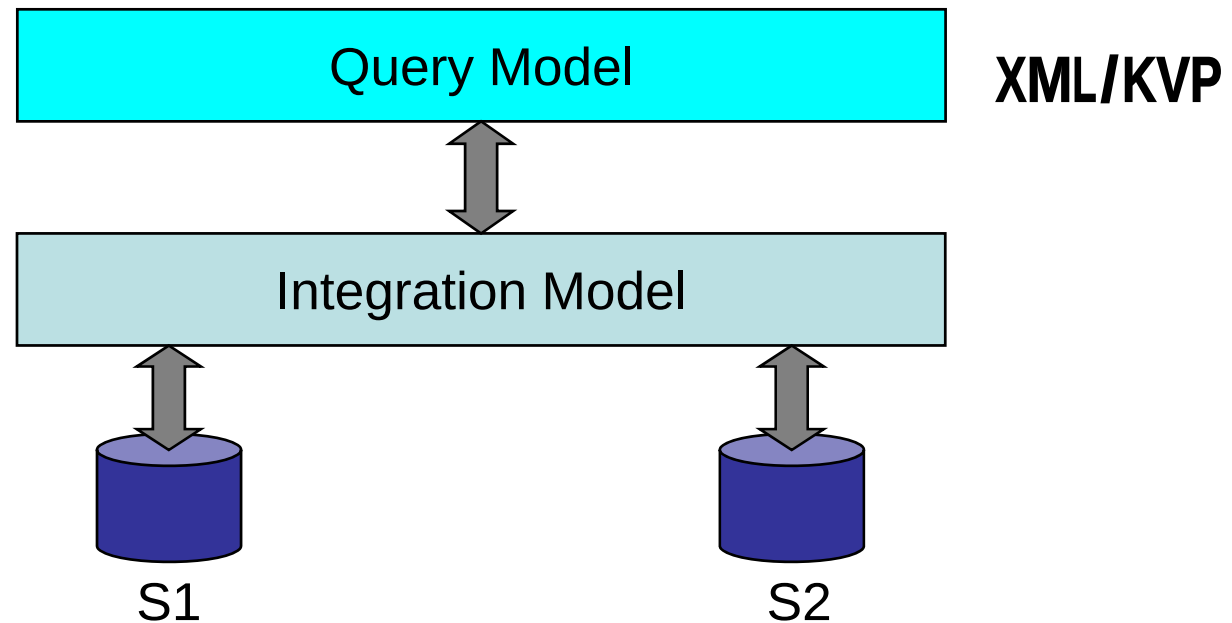




# Integration Approach

**GQuery**

**GML**



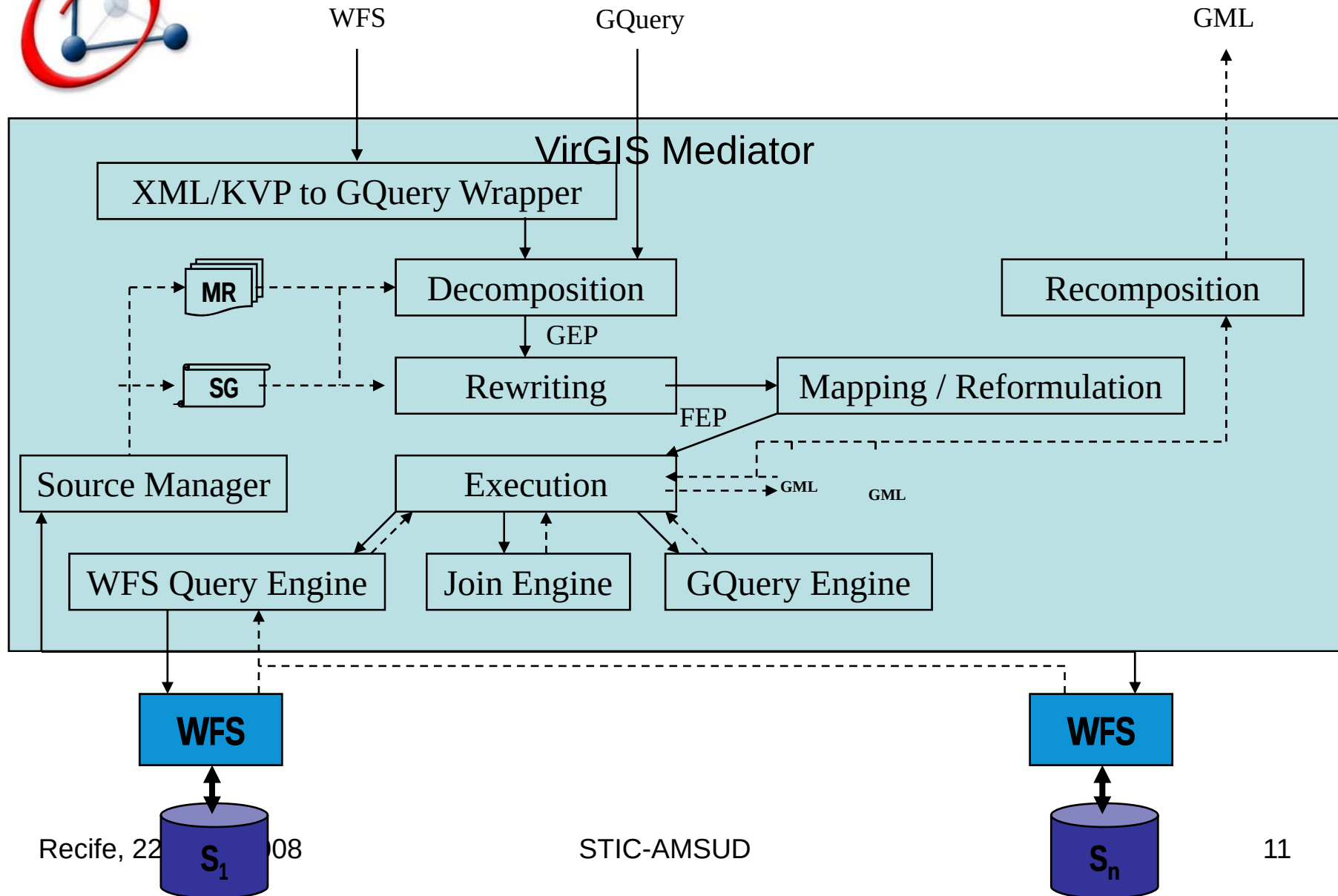


# GQuery

- Query Language for GML
  - XQuery Application
  - Resources defined as functions in a functional language
- Performance issues
  - Relies on existing XQuery engines
  - Enhancement with a local join module (SortJoin)



# VirGIS Architecture



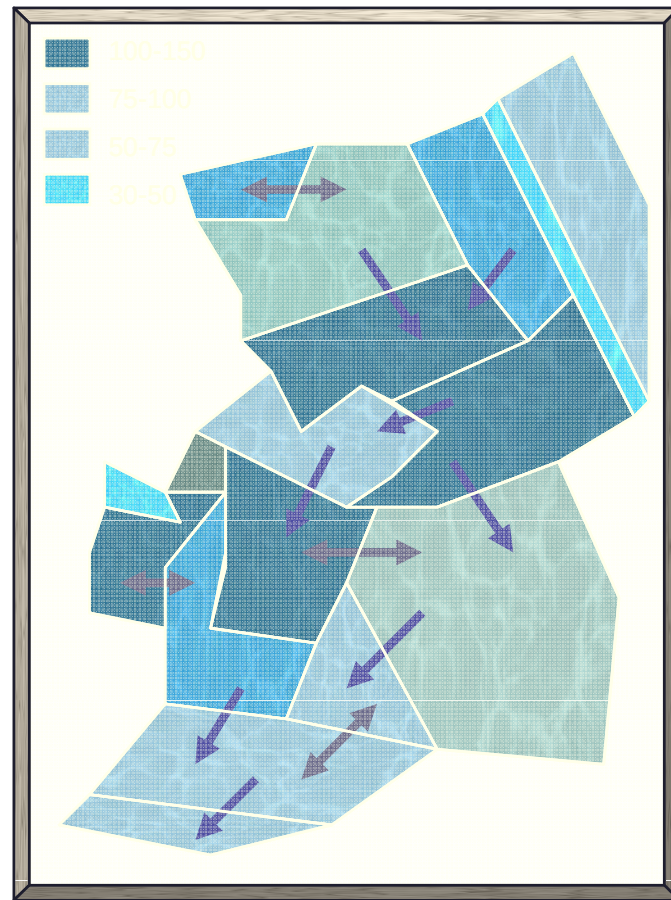


# Outline

1. Motivation
2. Starting point: the VirGIS project
3. Reasoning with (Data) Quality
4. Quality-Aware Query Processing
5. Service-Based Quality Approach
6. Conclusion



# The Flooding Case Study



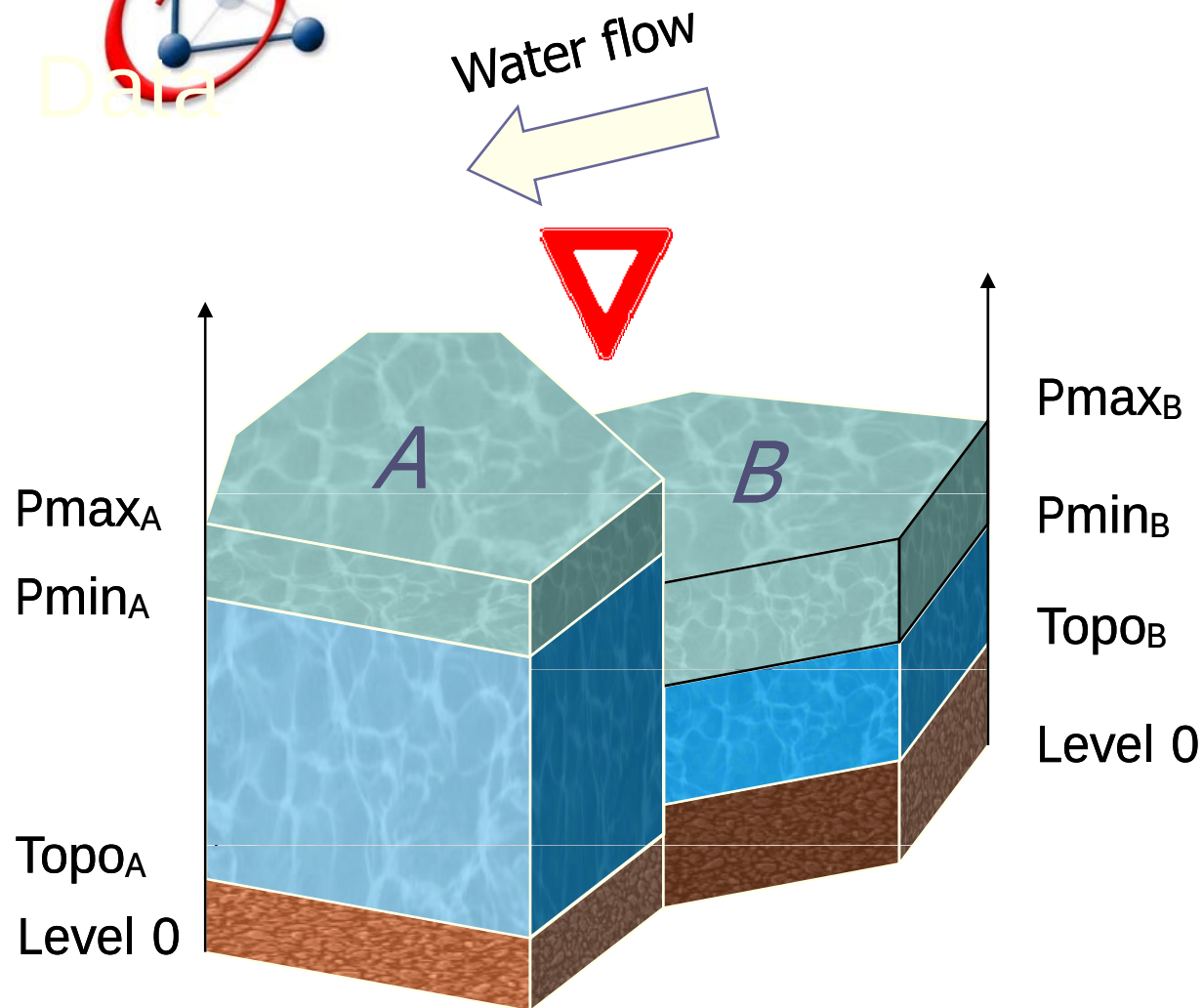
A flooded area divided into parcels.

Two kinds of information:

1. Water levels in parcels
2. Water flows between parcels



# The Flooding Case Study



Quality  
Parameters

Spatial cover

Validity date

Provider

Spatial accuracy

Deficit rate

Precision

Accuracy

Consistency



# Quality Parameters

	<u>parcel</u>	Geom	Topo	Height	<u>flow</u>	<u>Parcel-flow</u>
Cover	<i>Region</i>				<i>Region</i>	
Validity	<i>Date</i>	<i>→ Date</i>	<i>→ Date</i>	<i>→ Date</i>	<i>Date</i>	
Provider	<i>Name</i>	<i>→ Name</i>	<i>→ Name</i>	<i>→ Name</i>	<i>Name</i>	
Deficit rate	<i>Percent</i>	<i>Percent</i>	<i>Percent</i>	<i>Percent</i>		
Precision		<i>cm</i>	<i>cm</i>	<i>cm</i>		
Accuracy		<i>cm</i>	<i>cm</i>	<i>cm</i>		
Consistency						<i>Percent</i>



# User Query

**REVIGIS Project : User Interface**

Applications Edit View Windows Help

**Spreadsheet model for quality-to-feature**

Elements queried by user with expected quality

Rows: (fixed list of) quality parameters      Columns: (variable list of) elements describing the [FLOODING-Problem] Ontology

	parcel	Height	Topo	Geometry	parcel-flow
<b>Spatial coverage</b>	Hérault				
<b>Validity date</b>	_/11/1994				
<b>Provider</b>					
<b>Deficit rate</b>		<= 25%			
<b>Precision</b>		<= 10 cm			
<b>Accuracy</b>					
<b>Consistency</b>					<= 1%

☒ **parcel**

- ☒ Id
- ☒ Height
- ☒ Topo
- ☒ Error
- ☒ Geometry

☐ **flow**

OK





# Logical Consistency

REVIGIS Project : User Interface

Applications Edit View Windows Help

Spreadsheet model for quality-to-feature

Elements queried by user with expected quality

Rows: (fixed list of) quality parameters Columns: (variable list of) elements describing the [FLOODING-Problem] Ontology

	parcel	Height	Topo	Geometry	parcel-flow
Spatial coverage	Hérault				
Validity date	23/11/1994				
Provider	Cémagref				
Deficit rate		13.5%			
Precision		10 cm			
Accuracy					
Consistency					?

Quality mediator daemon

WrOp  
 launching Catalog with C:\Drodo\WrapperOperator\Done  
 launching with C:\Drodo\Todo\Catalog.txt  
 Open file C:\Drodo\WrapperOperator\Catalogue.txt  
 with HERAULT is Vrai  
 with < 10 YEARS is Faux  
 with FLOODING is Vrai  
 with HERAULT is Faux  
 with < 10 YEARS is Faux  
 with FLOODING is Faux  
 with HERAULT is Faux  
 with < 10 YEARS is Faux  
 with FLOODING is Faux  
 with HERAULT is Faux  
 with < 10 YEARS is Faux  
 with FLOODING is Faux  
 with HERAULT is Faux  
 with < 10 YEARS is Faux  
 with FLOODING is Faux

Operations found

The Mediator system has identified these programs as convenient operation for you

LCC

Logical Consistency Control operation for water levels and water flows.  
 Version 1.0, Provider: REVIGIS, Author: M. Khelfallah, Speed: High,

Apply Exit

MapFlooding

Appuyez sur F1 pour obtenir l'aide en ligne.



# Logical Revision

**REVIGIS Project : User Interface**

Applications Edit View Windows Help

**Spreadsheet model for quality-to-feature**

Elements queried by user with expected quality  
Rows: (fixed list of) quality parameters Columns: (variable list of) elements describing the [FLOODING-Problem] Ontology

	parcel	Height	Topo	Geometry	parcel-flow
<b>Spatial coverage</b>	Hérault				
<b>Validity date</b>	23/11/1994				
<b>Provider</b>	Cémagref				
<b>Deficit rate</b>		13.5%			
<b>Precision</b>		10 cm			
<b>Accuracy</b>					
<b>Consistency</b>					2.6%

**Quality mediator daemon**

WrOp  
launching Catalog with C:\Drodo\WrapperOperator\Done  
launching with C:\Drodo\Todo\Catalog.txt  
Open file C:\Drodo\WrapperOperator\Catalogue.txt  
with HERAULT is Vrai  
with < 10 YEARS is Faux  
with FLOODING is Vrai  
with HERAULT is Faux  
with < 10 YEARS is Faux  
with FLOODING is Faux  
with HERAULT is Faux  
with < 10 YEARS is Faux  
with FLOODING is Faux  
with HERAULT is Faux  
with < 10 YEARS is Faux  
with FLOODING is Faux  
with HERAULT is Faux  
with < 10 YEARS is Faux  
with FLOODING is Faux  
with HERAULT is Faux

**Operations found**

The Mediator system has identified these programs as convenient operation for you

LCR

Logical Consistency Revision operation for water levels and flows  
Version 2.0, Provider: REVIGIS, Author: M. Khelfallah, Speed: Low,

Apply Exit

**GeoMedia Professional - Wrap...**

Fichier Edition Affichage Insertion Outils Analyse  
Warehouse Légende Metagate Fenêtre Aide  
Terrain Grid

**MapFlooding**

Appuyez sur F1 pour obtenir l'aide en ligne.



# Correcting Data

**REVIGIS Project : User Interface**

Applications Edit View Windows Help

**Spreadsheet model for quality-to-feature**

Elements queried by user with expected quality

Rows: (fixed list of) quality parameters      Columns: (variable list of) elements describing the [FLOODING-Problem] Ontology

	parcel	Height	Topo	Geometry	parcel-flow
<b>Spatial coverage</b>	Hérault				
<b>Validity date</b>	23/11/1994	23/11/1994	23/11/1994	23/11/1994	
<b>Provider</b>	?	REVIGIS	Cémagref	Cémagref	
<b>Deficit rate</b>	?	13.5%	13.5%	0%	
<b>Precision</b>		10 cm	10 cm	?	
<b>Accuracy</b>		10 cm	10 cm	100 cm	
<b>Consistency</b>					0%

Launching[L:\Drodo\Done\partition.txt]

**Quality mediator daemon**

```

launching CSP with C:\Drodo\WrapperOperator\Done\C:
launching with C:\Drodo\Todo\CSP.txt
Launch operator CSP
launching PARTITION with C:\Drodo\WrapperOperator\
launching with C:\Drodo\Todo\PARTITION.txt
Launch operator PARTITION
launching CSP with C:\Drodo\WrapperOperator\Done\C:
launching with C:\Drodo\Todo\CSP.txt
Launch operator CSP
launching CSP with C:\Drodo\WrapperOperator\Done\C:
launching with C:\Drodo\Todo\CSP.txt
Launch operator CSP
launching LLC with C:\Drodo\WrapperOperator\Done\LL
launching with C:\Drodo\Todo\LLC.txt
Launch operator LLC
launching REVISION with C:\Drodo\WrapperOperator\
launching with C:\Drodo\Todo\REVISION.txt

```

**GIS-Wrapper**

```

GeoMedia: C:\Drodo\Flooding\Wrapper.gws - C:\Drodo\GIS\Wra
GIS[C:\Drodo\GIS\Wrapper\Done\LLCresult.txt]: SELECT CENTROI
GeoMedia: C:\Drodo\Flooding\Wrapper.gws - C:\Drodo\GIS\Wra
GIS[C:\Drodo\GIS\Wrapper\Done\RevisionResult.txt]: SELECT CEN
GeoMedia: C:\Drodo\Flooding\Wrapper.gws - C:\Drodo\GIS\Wra

```

**GeoMedia Professional - Wra...**

Fichier Edition Affichage Insertion Outils

Analyse Warehouse Légende Metagate Fenêtre

Aide Terrain Grid

**MapWindow1**

Appuyez sur F1 pour obtenir l'aide en ligne.



# Outline

1. Motivation
2. Starting point: the VirGIS project
3. Reasoning with (Data) Quality
4. Quality-Aware Query Processing
5. Service-Based Quality Approach
6. Conclusion



# Solution

- XQuality: an XQuery with quality features
- Basic Quality Services
- Composite Service



# XQuality

`doc("./Multisources/elevationcanada/001k.xml")//citation paramquality aligvertical="20" alighorizontal="100"`

Requête XQuery      Clé de qualité      Critères

- Query

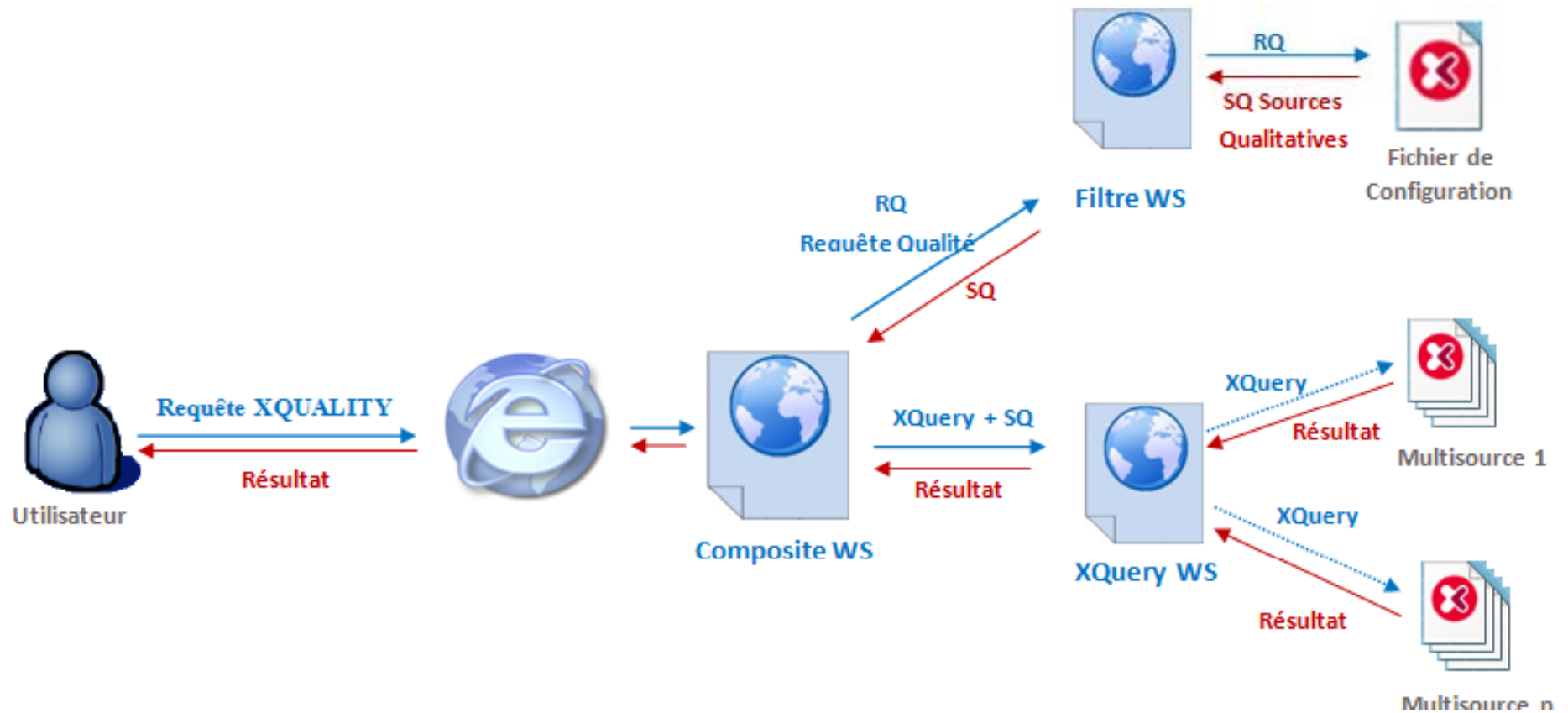
```
doc("./src/main/config-sources.xml")//source
[facteurQualite/facteur="aligvertical" and facteurQualite/valeur="20"
and facteurQualite/facteur="aligvertical" and facteurQualite/valeur="100"] /url
```

- Result

```
<url>./Multisources/elevationcanada/001k12fgdc.xml</url>
<url>./Multisources/elevationcanada/001k13fgdc.xml</url>
<url>./Multisources/elevationcanada/001k14fgdc.xml</url>
```

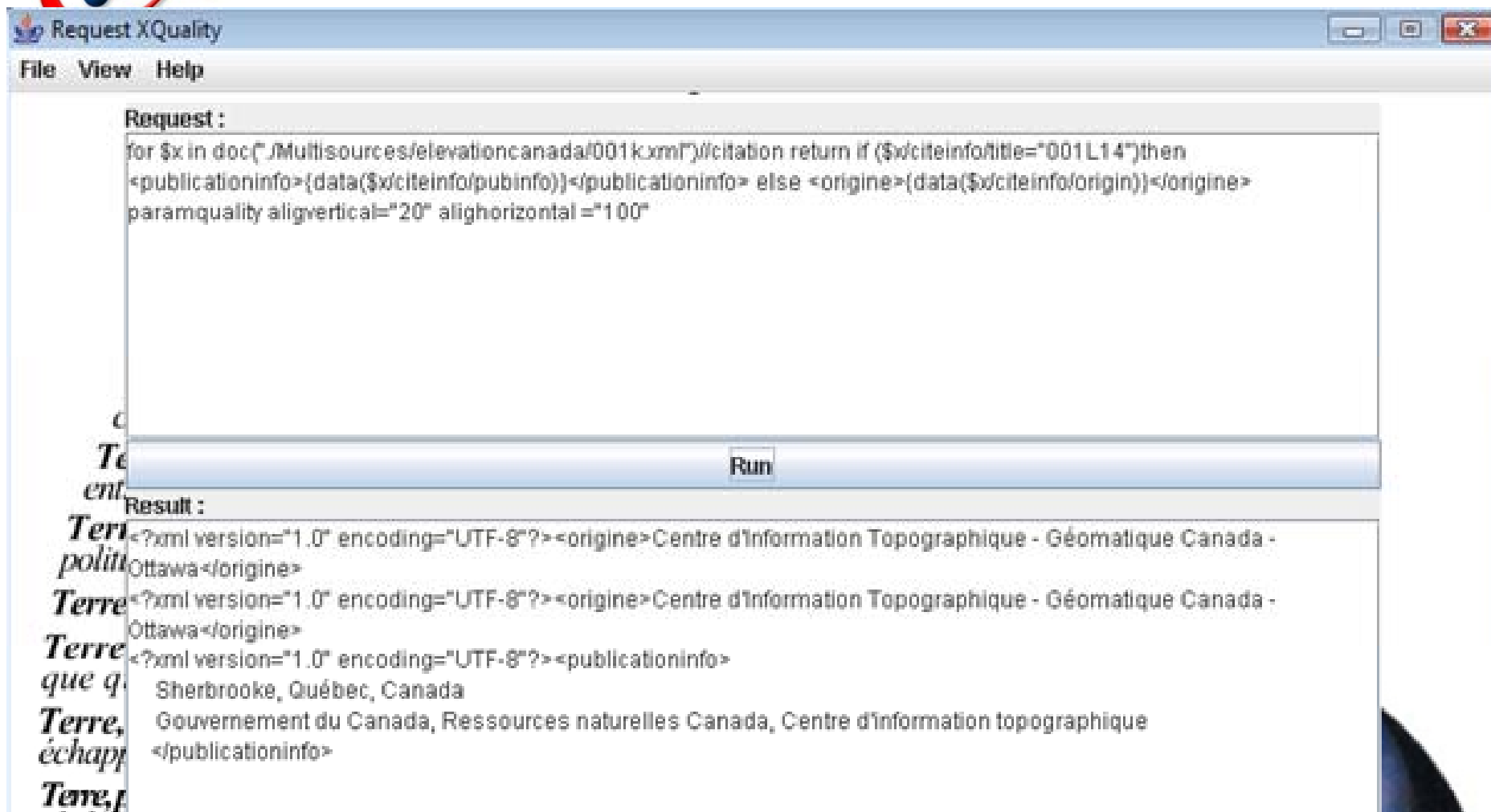


# Service Architecture





# Query Interface







# Wrap-up

- Spatial Data Mediation System
- Different Quality Experimentations
- Management of ISO 19115 Metadata Quality



# Future Work

- Towards Geospatial Web Reasoning
  - OGC's Geospatial Semantic Web Interoperability Experiment
  - Development and encoding of formal geospatial ontologies
  - Geospatial service interfaces which can provide service information formulated in OWL-S
  - WFS that allow requests for ontologically expressed service and content descriptions
  - Semantic query language processing interfaces for WFS