

BI2 : Un profil UML pour les Indicateurs Décisionnels

Sandro Bimonte

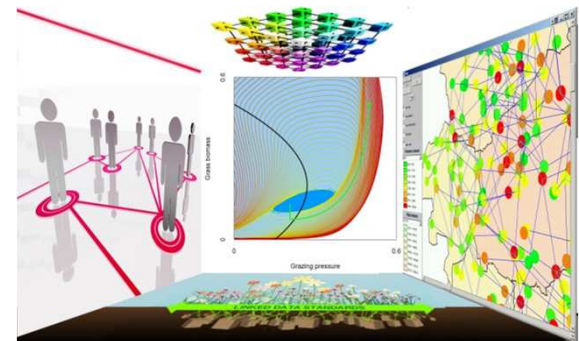
Irstea, TSCF, 9 Av. Blaise Pascal, 63178,
Aubière, France

sandro.bimonte@irstea.fr



www.irstea.fr

- Thème de Recherche MOTIVE





Plan

Motivations

BI2 UML profil

Implémentation

Travaux existants

Conclusion

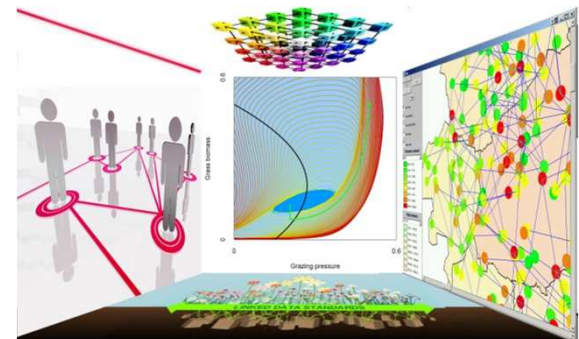


Motivations



www.irstea.fr

- Thème de Recherche MOTIVE





Motivations : Contexte

La mise en place d'un système décisionnel (**Business Intelligence**) implique différents types d'indicateurs

Les décideurs peuvent s'appuyer sur plusieurs méthodes et technologies de BI

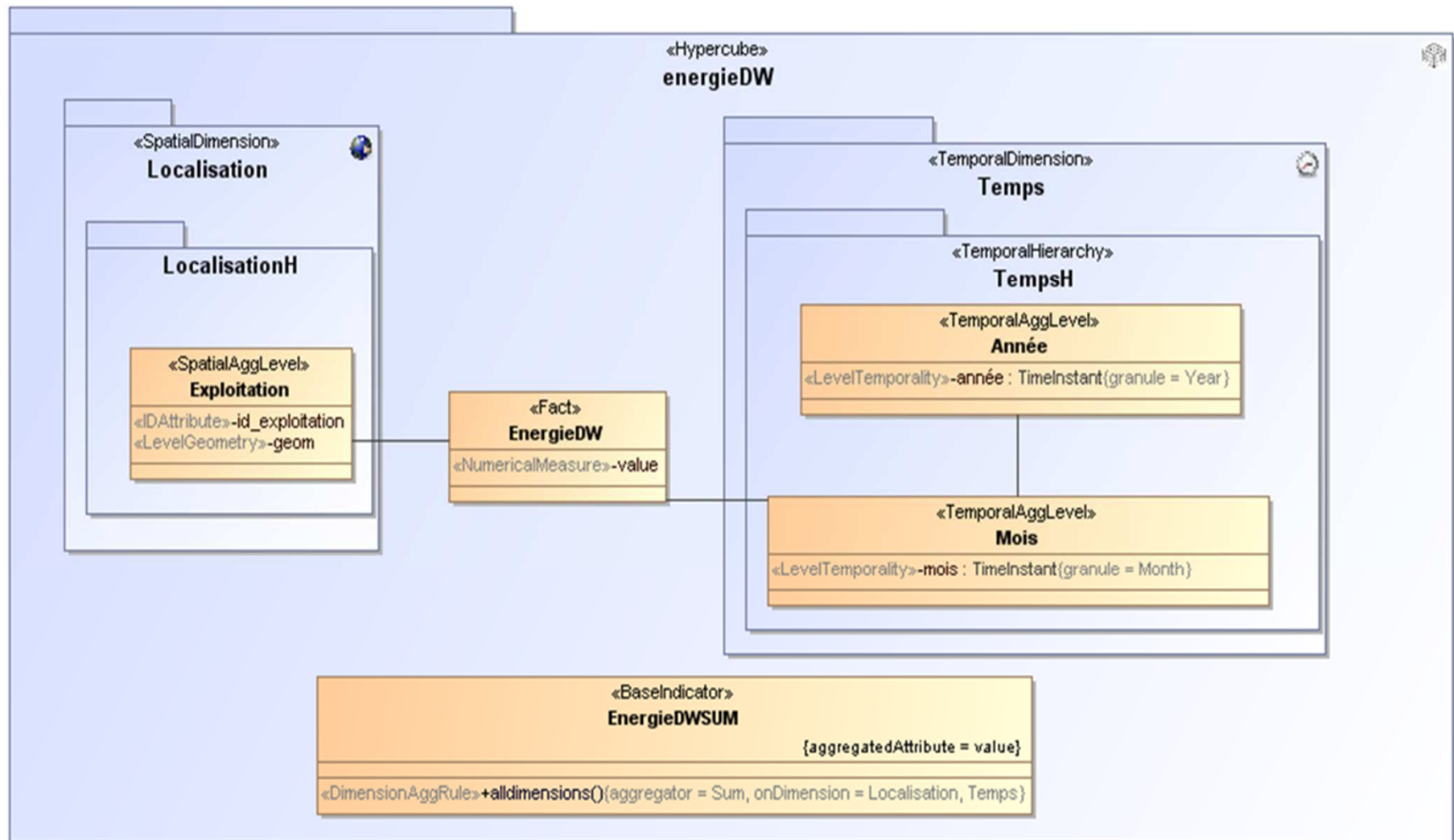
La modélisation conceptuelle des systèmes complexes est indispensable à la réussite des projets de BI (Torlone, 2003)

Permet aux concepteurs de définir des modèles faciles à comprendre par les décideurs

Manque de cadre conceptuel unique pour les indicateurs BI (cf. travaux existants)

Motivations : Cas d'étude (1/3)

Projet EDEN





Motivations : Cas d'étude (2/3)

Les exploitants souhaitent étudier deux scénarios :

Alerte :

Ex: Décideurs sont intéressés à être alertés (Middelfart et Pedersen, 2010), le cas où une machine présente un dysfonctionnement ou un mauvais réglage

Analyse :

Ex : Les exploitants souhaitent analyser les consommations en utilisant le modèle multidimensionnel

Ex : Analyser la consommation énergétique par jour par capteur sans avoir besoin d'accéder à d'autres affichages graphiques à partir de celui par jour et capteur



Motivations : Cas d'étude (3/3)

Les exploitants souhaitent étudier l'Alerte et l'Analyse avec 3 requêtes :

	Requête
Q1	Toutes les 3 secondes : Quelle est la consommation totale, maximale et minimale d'électricité au cours des 3 dernières secondes ?
Q2	Quelle est la consommation totale d'électricité par exploitation, mois et année ?
Q3	Quelle est la consommation totale d'électricité par jour et capteur ?

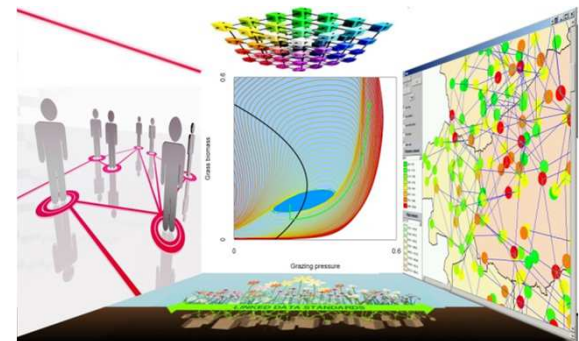


BI2 UML profil



www.irstea.fr

- Thème de Recherche MOTIVE





BI2 UML profil : Cadre conceptuel

Les requêtes présentent différentes caractéristiques :

Agrégation d'une valeur numérique

Données Historiques : données qui sont seulement ajoutées à la base de données sans qu'aucune donnée ne soit effacée

Fréquence continue : le temps d'exécution de la requête est inclus dans la requête même

Navigation hiérarchique : permettre une navigation hiérarchique dans l'outil de restitution



BI2 UML profil : Cadre conceptuel

Classes d'indicateurs :

Données Historiques, Fréquence non continue, Navigation hiérarchique (**Indicateur OLAP**)

Indicateurs de type OLAP

Données Historiques, Fréquence non continue, Navigation non hiérarchique (**Indicateur OLTP**)

Indicateurs définis dans les outils de reporting implémentés à partir d'un système OLTP classique via des requêtes SQL

Données non Historiques, Fréquence continue, Navigation non hiérarchique (**Indicateur Stream**)

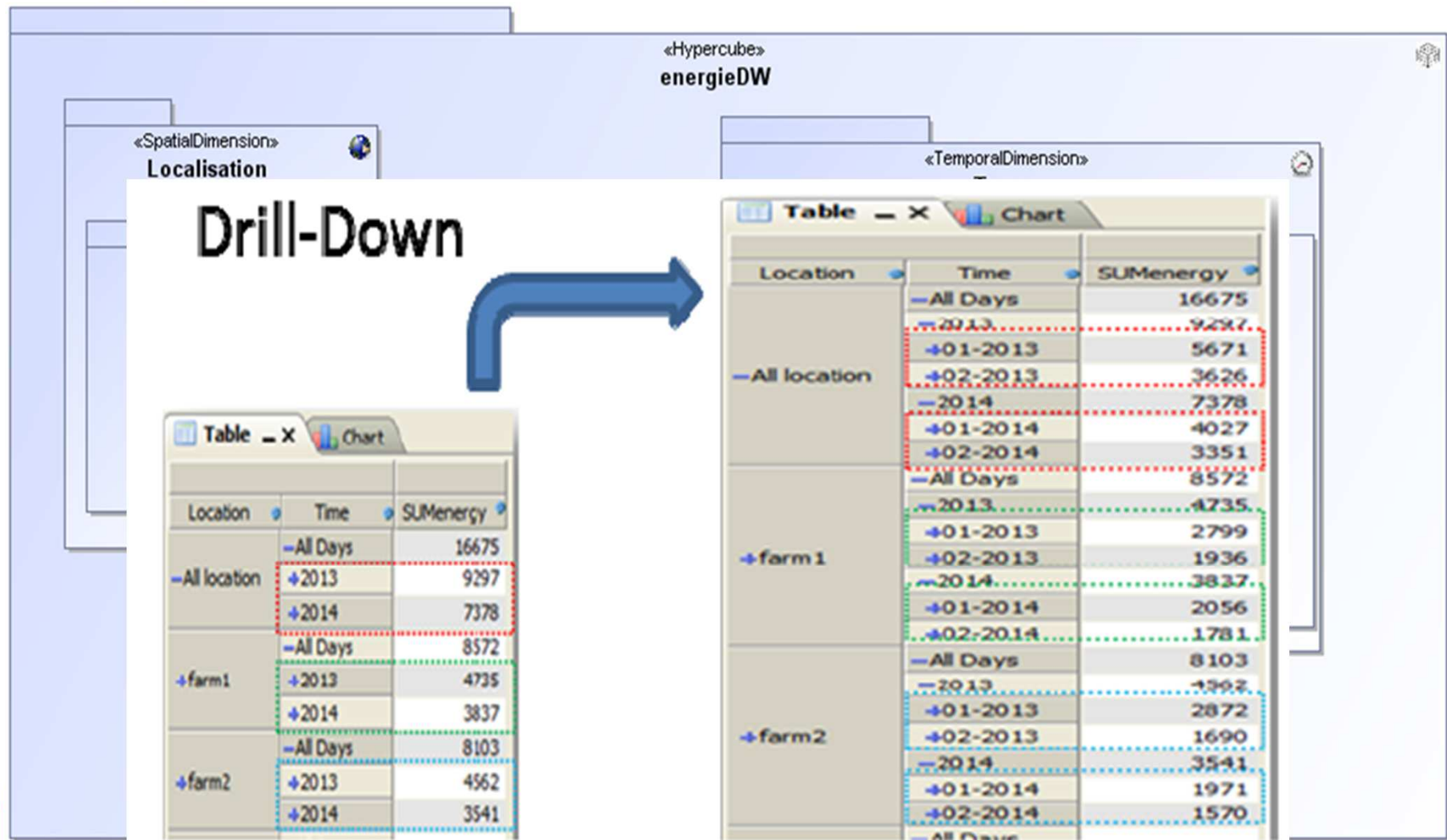
Indicateurs calculés au sein des Data Stream Management Systems

BI2 UML profil : Cadre conceptuel

Vision d'ensemble

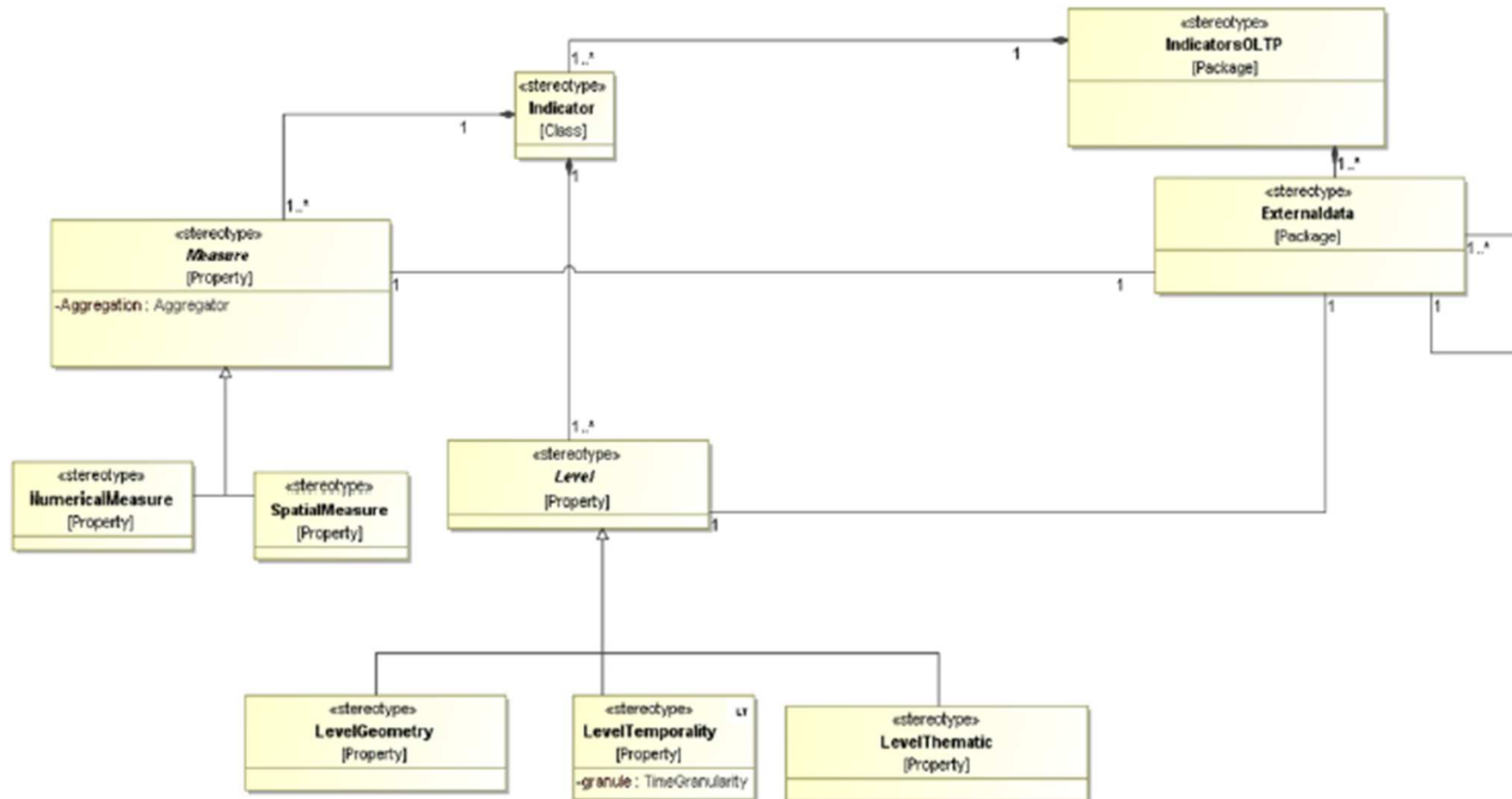
Requête	Données		
	<i>Fréquence</i>	<i>Historiques</i>	<i>Non Historiques</i>
<i>Navigation</i>			
Hiérarchique	Continue		
Hiérarchique	Non Continue	OLAP (Q2)	OLAP
Non Hiérarchique	Continue		DSMS (Q1)
Non Hiérarchique	Non Continue	OLTP (Q3)	OLTP

BI2 UML profil : Cadre conceptuel



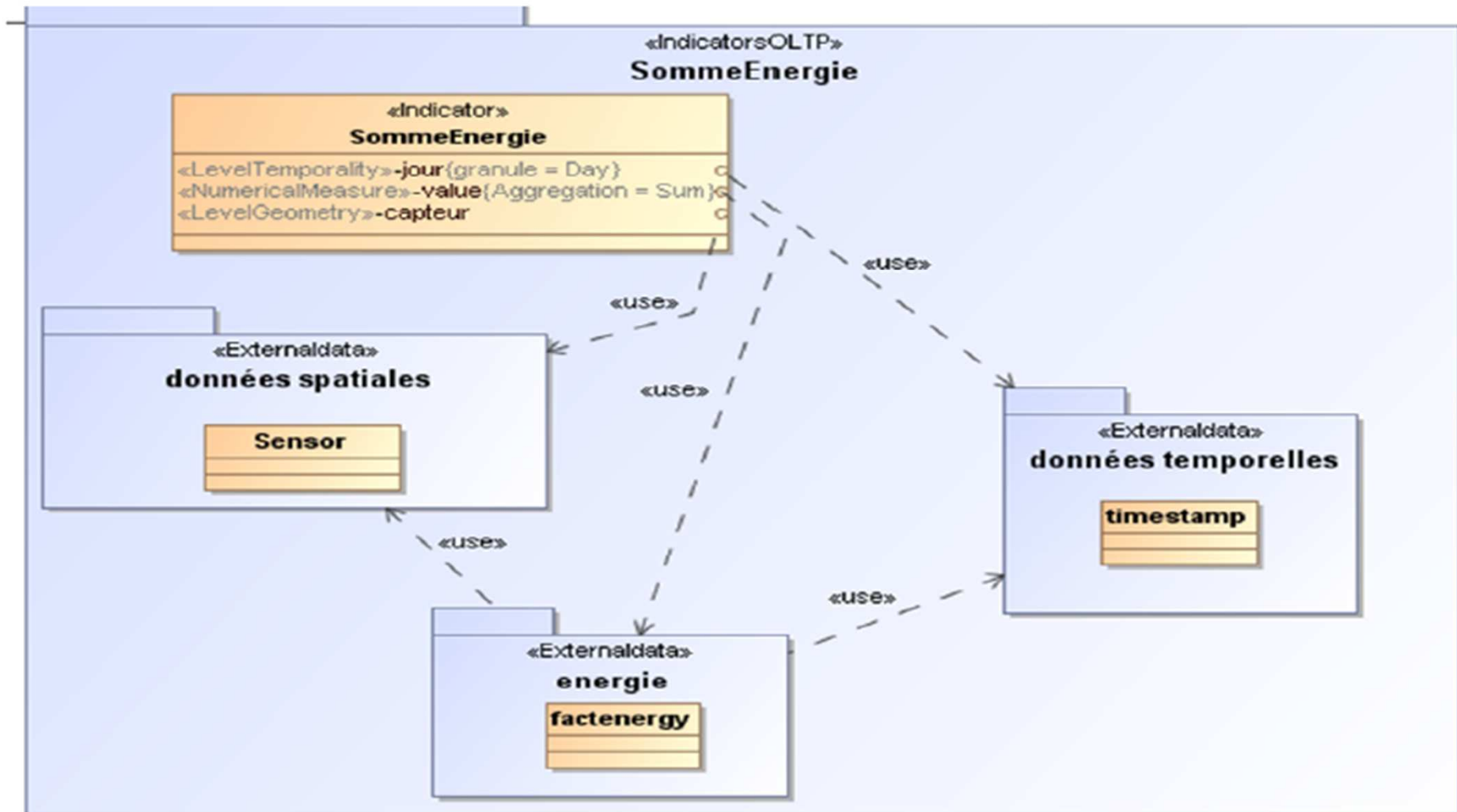
BI2 UML profil : Indicateur OLTP (1/2)

Les indicateurs de type Données Historiques, Fréquence non continue, Navigation non hiérarchique



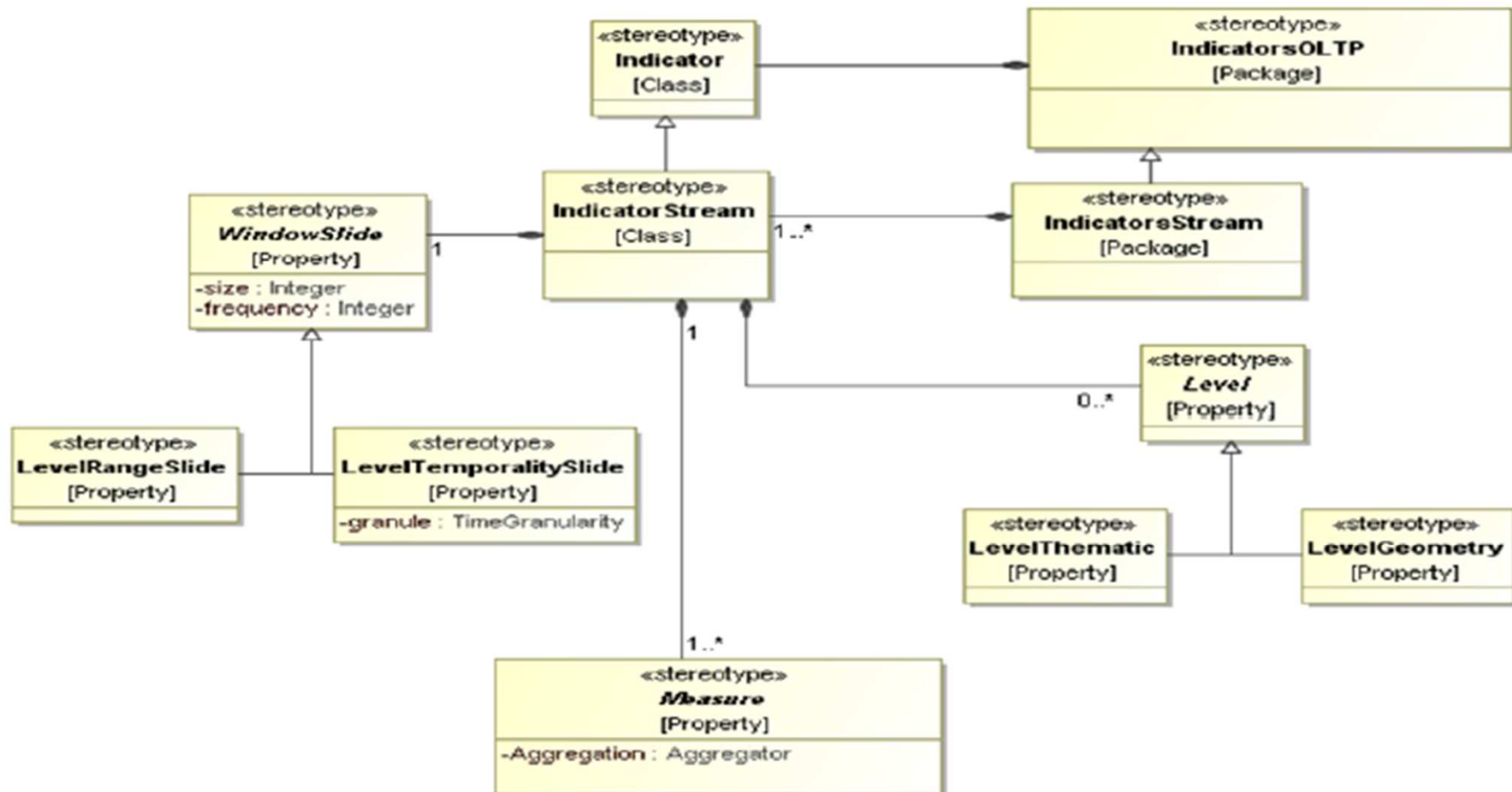
BI2 UML profil : Indicateur OLTP (2/2)

Ex : Quelle est la consommation totale d'électricité par jour et capteur ?



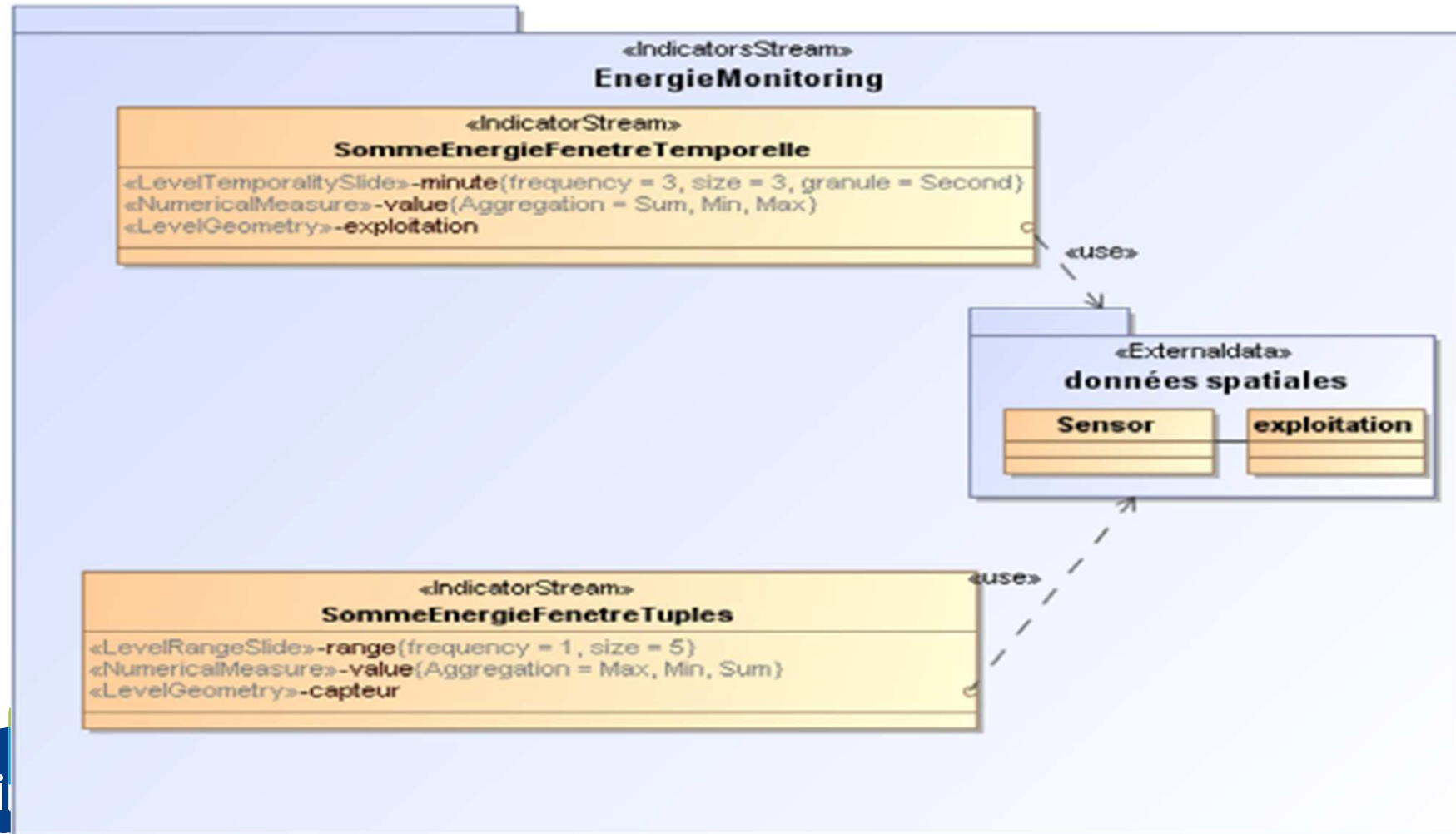
BI2 UML profil : Indicateur Stream (1/2)

Un indicateur de type Données non Historiques, Fréquence continue, Navigation non hiérarchique



BI2 UML profil : Indicateur Stream (2/2)

Toutes les 3 secondes : Quelle est la consommation totale, maximale et minimale d'électricité au cours des 3 dernières secondes ?





BI2 UML profil : AGL MagicDraw (1/2)

Le profil UML est implémenté dans l'AGL **MagicDraw**

MagicDraw permet la définition de profils qui étendent les éléments du méta-modèle d'UML et implémente les **contraintes OCL**

Lors de la conception de modèles, ces contraintes peuvent être ensuite contrôlées afin d'éviter la création de modèle incorrect

BI2 UML profil : AGL : MagicDraw (2/2)

Ex : un package « IndicatorsOLTP » ne doit pas contenir d'indicateur stream («IndicatorStream»)

The screenshot shows the MagicDraw software interface. On the left, a project tree displays a package named 'IndicatorsOLTP'. The main workspace shows a UML diagram for the package 'RetailCompanySDW' with a constraint 'constraints'. A class diagram element is highlighted, showing a class named 'indicatorS_' with a property 'value' of type '<numericalMeasure>'. A red box highlights the class and its property.

At the bottom, the 'Résultats de la validation' (Validation Results) window shows an error message:

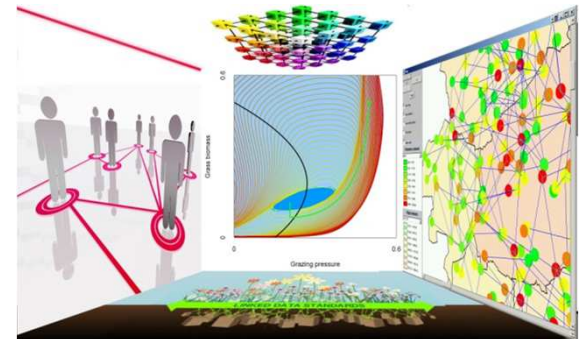
Élément	Sévérité	Abréviation	Message
indicatorOLTP_[RetailCompanySDW]	error	no indicatorS	no stream indicator can be defined

Implémentation



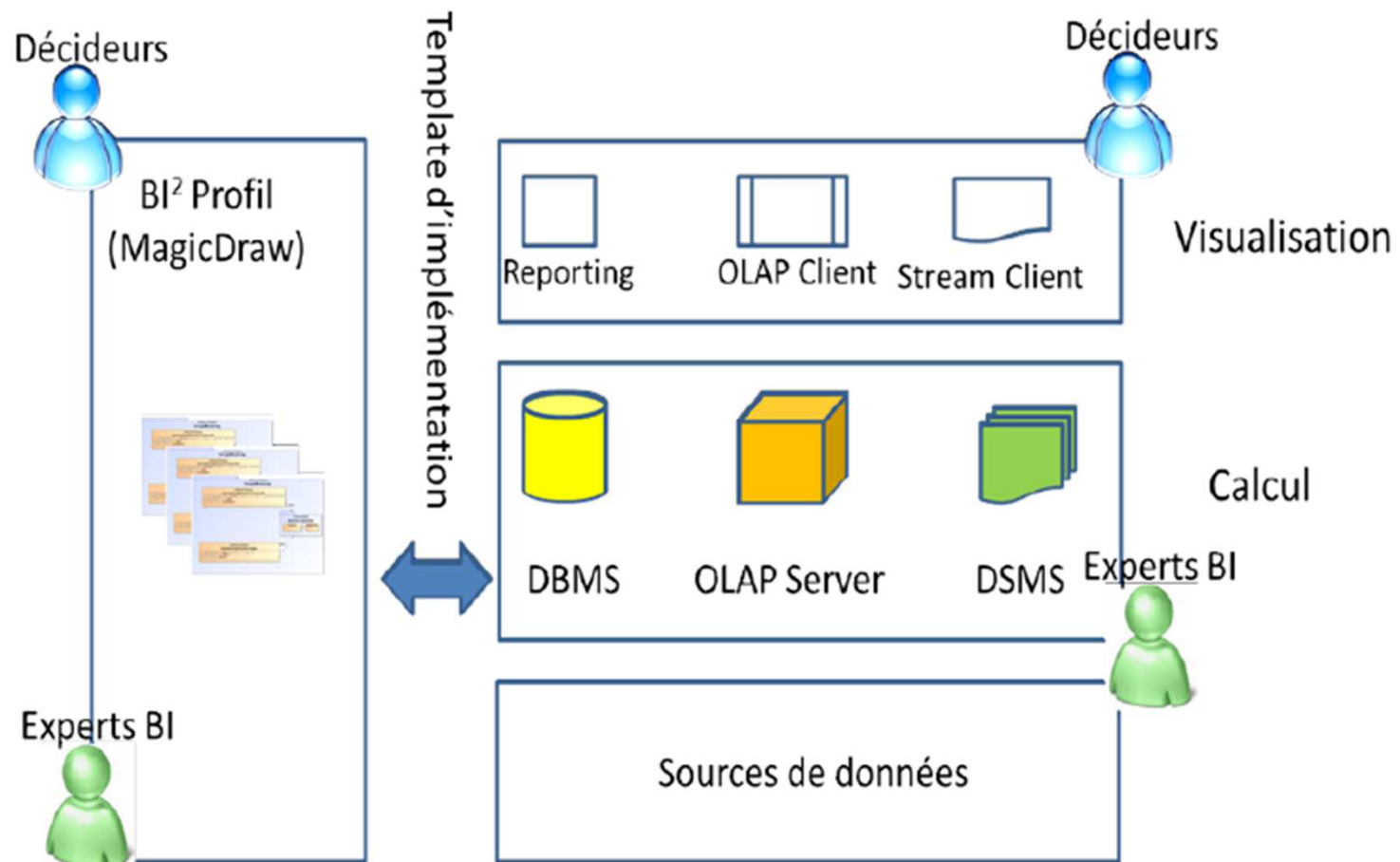
www.irstea.fr

- Thème de Recherche MOTIVE



Implémentation

BI2 UML profil permet la conceptions de différents indicateurs dans le même cadre de conception et implémentation



Implémentation : Indicateur OLTP (1/2)

Implémentation dans un SGDB relationnel

Template SQL : Mapping entre les éléments du profil et une requête SQL (Select From Group By)

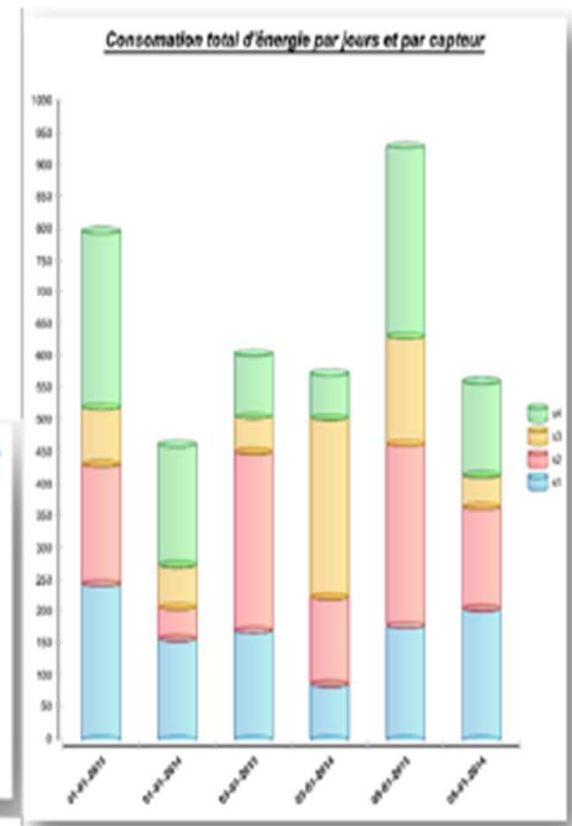
```
SELECT Aggregation(Indicator.Measure1.value),  
...,Aggregation(Indicator.Measuren.value),  
Externaldata2.Level2, ...,Externaldatam.Levelm  
  
FROM Externaldatameasure ...,Externaldatam  
WHERE (Externaldatameasure.attmeasure = Externaldata2.att AND  
...  
AND (Externaldatameasure.attmeasure = Externaldatam.attm)  
  
GROUP BY Externaldata2.Level2, ..., Externaldatam.Levelm
```

Implémentation : Indicateur OLTP (2/2)

Ex : SQL - Quelle est la consommation totale d'électricité par jour et capteur ?

Visualisation : Outil de reporting SpagoBI

```
select time.day as Jours,sensor.id_sensor as Capteur,sum(value) as Valeur
from energy_fact,sensor,time
where energy_fact.id_time=time.id_time
and energy_fact.id_sensor=sensor.id_sensor
group by Jours,Capteur
order by Jours
```



Implémentation : Indicateur Stream (1/2)

Implémentation dans un DSMS

Template Continuous QL : Mapping entre les éléments du profil et une requête CQL (Select From window Group By)

```

SELECT
Aggregation(Indicator.Measure1.value),...,Aggregation(Indicator.Measuren.value),
Externaldata2.Level2, ..., Externaldatam.Levelm

FROM Event.win:LevelTemporalitySlide | LevelRangeSlide as Event
+
"sql:Data2 ['select Externaldata2.Level2 from Externaldata2']as Externaldata2 " +
...
"sql:Datam ['select Externaldatam.Levelm from Externaldatam'] as Externaldatam "
WHERE (Eventmeasure.attmeasure = Externaldata2.att AND
...
AND (Eventmeasure.attmeasure = Externaldatam.attm)

GROUP BY Externaldata2.Level2, ..., Externaldatam.Levelm

```

Implémentation : Indicateur Stream (2/2)

Ex : Esper DSMS - Toutes les 3 secondes : Quelle est la consommation totale, maximale et minimale d'électricité au cours des 3 dernières secondes ?

```
"select labelfarm as labelfarm,sum(consomationStream) as sum_val from SensorEvent.win:time_batch(3 sec) as sensorevent," +
"sql:streamOlap['select labelfarm,farm.id_farm as farm,sensor.id_farm as s_farm,id_sensor from farm,sensor']as sensorbdfarm " +
"where sensorevent.label=sensorbdfarm.id_sensor " +
"and sensorbdfarm.s_farm=sensorbdfarm.farm " +
"group by labelfarm";
```

(a)

```
SensorEvent [Mon Dec 01 16:10:30 CET 2014 : s4 consumed 225]
SensorEvent [Mon Dec 01 16:10:30 CET 2014 : s5 consumed 139]
SensorEvent [Mon Dec 01 16:10:30 CET 2014 : s5 consumed 51]
SensorEvent [Mon Dec 01 16:10:30 CET 2014 : s2 consumed 136]
SensorEvent [Mon Dec 01 16:10:30 CET 2014 : s3 consumed 155]
SensorEvent [Mon Dec 01 16:10:30 CET 2014 : s7 consumed 250]
SensorEvent [Mon Dec 01 16:10:30 CET 2014 : s4 consumed 87]
SensorEvent [Mon Dec 01 16:10:30 CET 2014 : s6 consumed 170]
SensorEvent [Mon Dec 01 16:10:30 CET 2014 : s4 consumed 20]
SensorEvent [Mon Dec 01 16:10:40 CET 2014 : s5 consumed 93]
SensorEvent [Mon Dec 01 16:10:40 CET 2014 : s5 consumed 251]
SensorEvent [Mon Dec 01 16:10:40 CET 2014 : s7 consumed 255]
SensorEvent [Mon Dec 01 16:10:40 CET 2014 : s5 consumed 121]
-----
+ [MONITOR] ValueSUM of | fare3 ---> 170
-----
- [MONITOR] ValueSUM of | fare2 ---> 866
-----
- [MONITOR] ValueSUM of | fare1 ---> 935
-----
SensorEvent [Mon Dec 01 16:10:41 CET 2014 : s3 consumed 230]
SensorEvent [Mon Dec 01 16:10:41 CET 2014 : s2 consumed 200]
SensorEvent [Mon Dec 01 16:10:41 CET 2014 : s3 consumed 51]
SensorEvent [Mon Dec 01 16:10:41 CET 2014 : s5 consumed 177]
SensorEvent [Mon Dec 01 16:10:41 CET 2014 : s7 consumed 30]
SensorEvent [Mon Dec 01 16:10:42 CET 2014 : s2 consumed 60]
SensorEvent [Mon Dec 01 16:10:42 CET 2014 : s2 consumed 167]
SensorEvent [Mon Dec 01 16:10:42 CET 2014 : s3 consumed 294]
SensorEvent [Mon Dec 01 16:10:42 CET 2014 : s2 consumed 226]
SensorEvent [Mon Dec 01 16:10:42 CET 2014 : s2 consumed 101]
SensorEvent [Mon Dec 01 16:10:43 CET 2014 : s1 consumed 275]
SensorEvent [Mon Dec 01 16:10:43 CET 2014 : s4 consumed 248]
SensorEvent [Mon Dec 01 16:10:43 CET 2014 : s6 consumed 38]
-----
- [MONITOR] ValueSUM of | fare3 ---> 108
-----
- [MONITOR] ValueSUM of | fare2 ---> 546
-----
- [MONITOR] ValueSUM of | fare1 ---> 1736
```

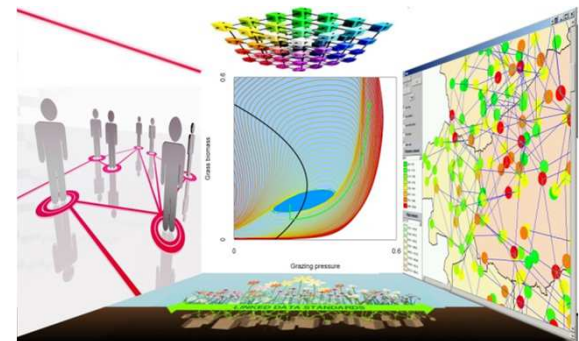



Travaux existants



www.irstea.fr

- Thème de Recherche MOTIVE





Travaux existants (1/2)

La modélisation conceptuelle des streams n'a pas été abordée (Bimonte et al., 2012) :

(Malinowski et Zimányi, 2008) définissent un modèle conceptuel pour les Eds temporelles.

types temporels dans les niveaux, les hiérarchies et les mesures

(Douglass, 1998) définissent un profil UML pour les types de données en temps réel.



Travaux existants (2/2)

Modélisation des indicateurs décisionnels :

Expression des besoins d'analyse (Giorgini et al., 2008)

Structures des bases de données et ED (i.e. OLTP et OLAP)

Processus qui lie les données OLTP et OLAP (ETL) (Akkaoui et al., 2012)

mais ces modèles reposent tous sur une modélisation multidimensionnelle

La représentation de leur interrogation est définie avec :

Des langages textuels (MDX)

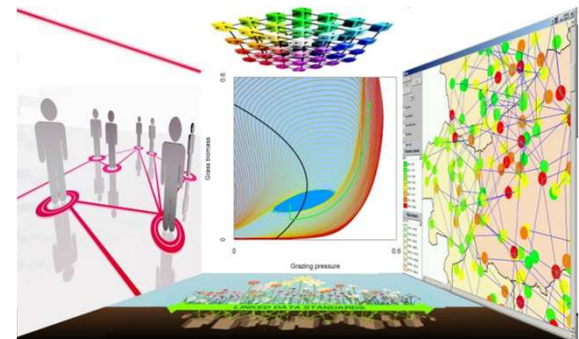
Des langages visuels (Bonhomme et al., 1999)

Conclusions



www.irstea.fr

- Thème de Recherche MOTIVE





Conclusions

De plus en plus de données sont disponibles pour une analyse décisionnelle via des indicateurs décisionnels différents

Nous présentons une première classification de ces indicateurs

Nous présentons un profil UML BI2 unique qui permet de représenter des indicateurs OLAP, OLTP et stream

Nous présentons également leur implémentation dans des outils industriels existants



Travaux futurs

Indicateurs de type *Données non Historiques, Fréquence continue, Navigation hiérarchique*

Génération automatique des requêtes d'implémentation

Définition d'une méthodologie de conception orientée besoins d'analyse, comme dans les EDs, pour la conception de ces différents indicateurs