

**COST Action C21 - 3rd
Workshop - Zaragoza**



Managing Heterogeneity in Urban Ontologies

Chantal BERDIER- EVS/EDU

Samuel GESCHE - Liris

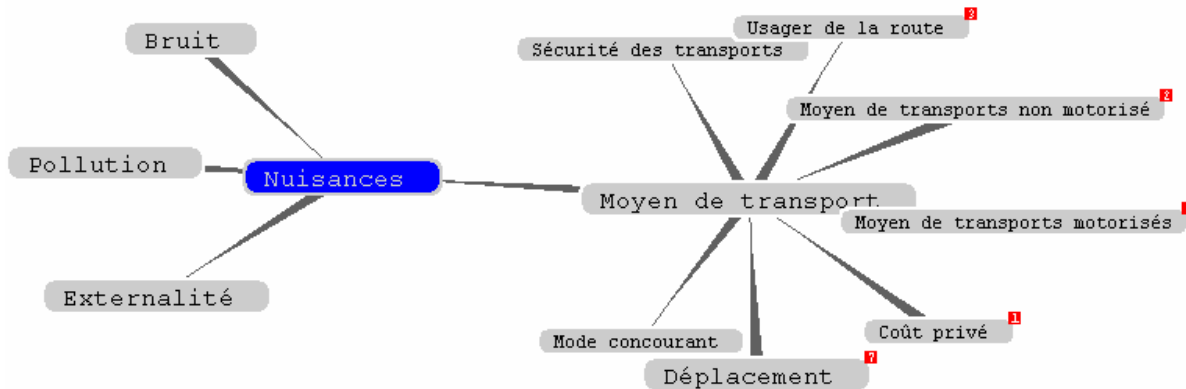
Sylvie CALABRETTO - Liris

Guy CAPLAT

October 20th 2008

Context

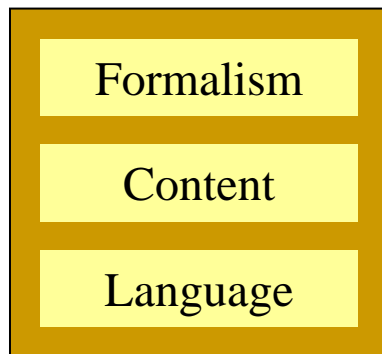
- **Towntology**
 - Urban ontologies
 - Pre-consensual ontologies



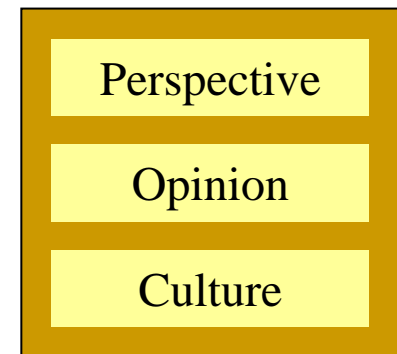
- **Multiple sources of heterogeneity: language, culture, profession...**

About heterogeneity

- Heterogeneity has many forms.



“Superficial” forms

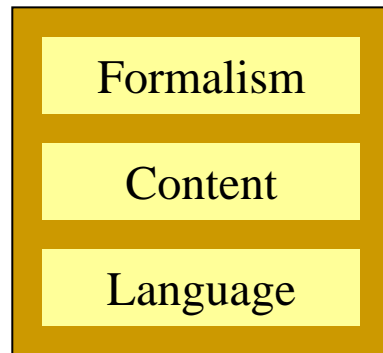


“Deep” forms

- Three ways of managing heterogeneity:
 - ❑ Reducing it (in the best cases, completely getting rid of it);
 - ❑ Working around it;
 - ❑ Taking advantage of it.

About heterogeneity

- (More or less) easily managed forms.
- “Superficial” forms of heterogeneity.



About heterogeneity

- Heterogeneity among Formalisms
 - There are many formalisms that can be used to build ontologies.
 - The W3C standard is OWL, which is based on RDF.
 - But before OWL has really become a standard, many languages have been designed (among which is the Towntology format).
 - There are many structures that claim the name of ontology.
 - What is an ontology again?
 - The towntology XML format: “lightweight ontologies”.

- Reduction: import/export systems

About heterogeneity

- Heterogeneity among Content
 - Empirical rule among ontologies:
 - “When working on the same subject, two teams are not likely to build the same ontology”.
 - True for students, true for experts.
- Reduction: ontology matching

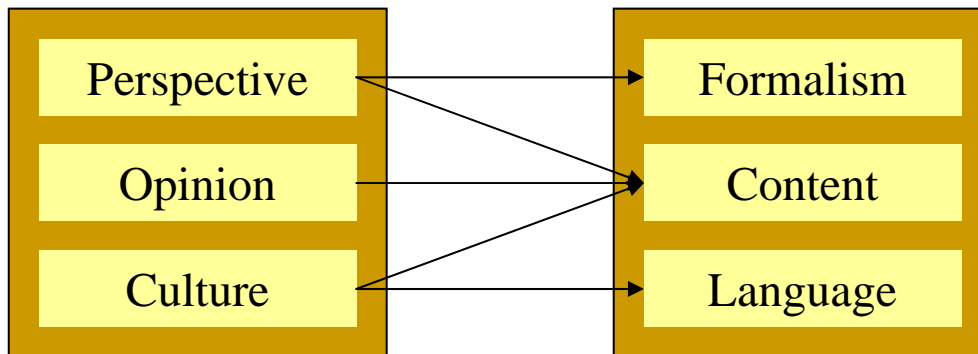
About heterogeneity

- Heterogeneity among Languages
 - If only everybody wrote in English... well yes, if only...
 - It is easier to give an accurate description in one's mother tongue.
 - Not all concepts exist in another language;
 - If they do, their meaning can be quite different;
 - Not all experts can write another language fluently.

- Reduction: translation

About heterogeneity

- Not easily managed forms.
- The “deep” forms of heterogeneity that result in the previous “superficial” forms.



About heterogeneity

- Heterogeneity among Perspectives
 - Urbanism (and therefore Towntology) includes many professions.
 - Each of these can have its own meaning for the same term.
- Workaround: multi-viewpoint approach
 - One viewpoint for each profession.
 - Most useful when building views of a single integrated system.

About heterogeneity

- Heterogeneity among Opinions
 - Several experts of the same field can have different positions on a given matter.
- Reduction: multi-expertise
 - A cognition specialist works with the experts to bring their knowledge to a consensus.

About heterogeneity

- Heterogeneity among Culture
 - Culture highly interferes in the expertise.
 - This is why towns are not the same in different countries.
- Workaround: is there any?
 - There are few international initiatives that take into account the cultural aspect.
 - Often the most worldwide recognized culture prevails (the American culture for many situations).

About heterogeneity

- ... and let us not forget heterogeneity among publication policies.
 - Standard or Specific ?
 - Protection of Information
 - Confidential or Shared ?
 - Hints or Full data ?

Taking advantage of heterogeneity

- Before reducing or even working around heterogeneity, it is necessary to know how it works.
 - Is there any way to map concepts from French to Italian ? Which ones ?
 - What if well-known experts disagree ? And how to detect how exactly they disagree ?
 - Is it realistic to integrate knowledge from several professions ? If not, what to do ?
- It often depends on the studied case.

Taking advantage of heterogeneity

- An expert can improve his knowledge and proficiency if allowed to learn from others.
 - ❑ By knowing how his work is carried out in other cultures.
 - ❑ By learning how to discuss efficiently with experts from other professions.
- Ontologies are designed to be easily learnt.
 - ❑ Confronting viewpoints expressed as ontologies can be especially enriching.

Taking advantage of heterogeneity

- ArcEnCiel: a system dedicated to viewpoint confrontation.
 - ❑ Using comparison of heterogeneous structures to mine new knowledge.
 - ❑ Written in Java
 - ❑ 10000 code lines
 - ❑ Used for several projects. Development of version 3 ('Hippodamos') was focused on the study of Townontology ontologies.

ArcEnCiel

- ArcEnCiel approach: Individual confrontation.
 - With the viewpoints expressed as ontologies, one expert can either:
 - compare his viewpoint with that of other experts;
 - or compare the viewpoints of several experts.
- It is easier than having several experts together.
 - Although complementary.

ArcEnCiel

- ArcEnCiel approach: Local confrontation.
 - Find locations that “show potential for connection”.
 - Visualize differences in neighborhood.
 - Wondering :
 - Whether the connection is relevant;
 - Whether there are another neighbor location that could be connected.
 - Aim: find differences of opinion by exploring similarity.
 - Serendipity versus Exhaustiveness

ArcEnCiel

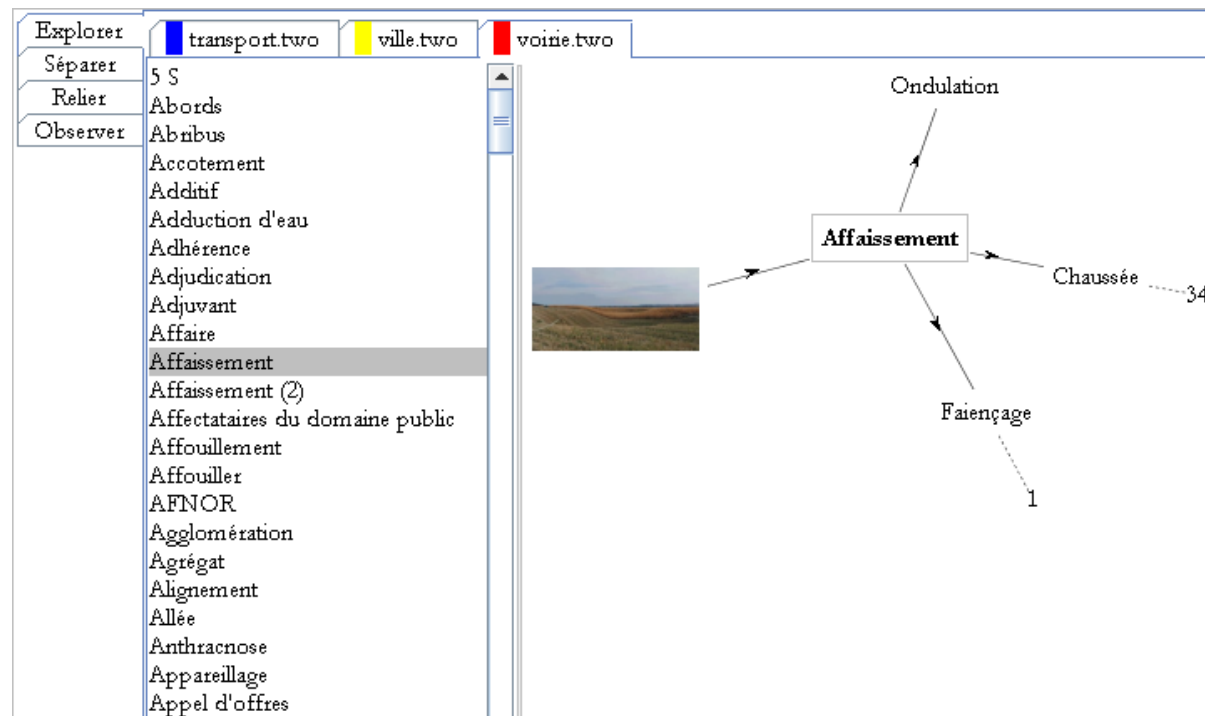
- ArcEnCiel approach: Multi-formalism confrontation
 - Each studied case had its own formalism.
- About multilingual confrontation:
 - We already do not rely on a static (or at least consensual) terminology.
 - What is the difference (apart from the scale) between different terminologies and different languages ?
 - The expert must know the languages, however he can discover the differences.

ArcEnCiel

- Confrontation can be made in different ways,
 - Influenced by the practices,
 - Or by the purpose.
- We cannot impose a sequence of actions.
- Computer-assisted Confrontation:
 - A set of tools,
 - All usable at any time,
 - No predefined order.

ArcEnCiel - Actions

- Explore
 - Explore viewpoints independantly



ArcEnCiel - Actions

■ Connect

□ Connect items

Bruit Bruit

Piste cyclable Piste cyclable

Plan de déplacements urbains Plan de déplacements urbains

Stationnement Stationnement

Voie Voie

Why connect the following terms:
accessibilité;accessibilité;accessibility

These terms are synonyms

Another reason (if relevant):

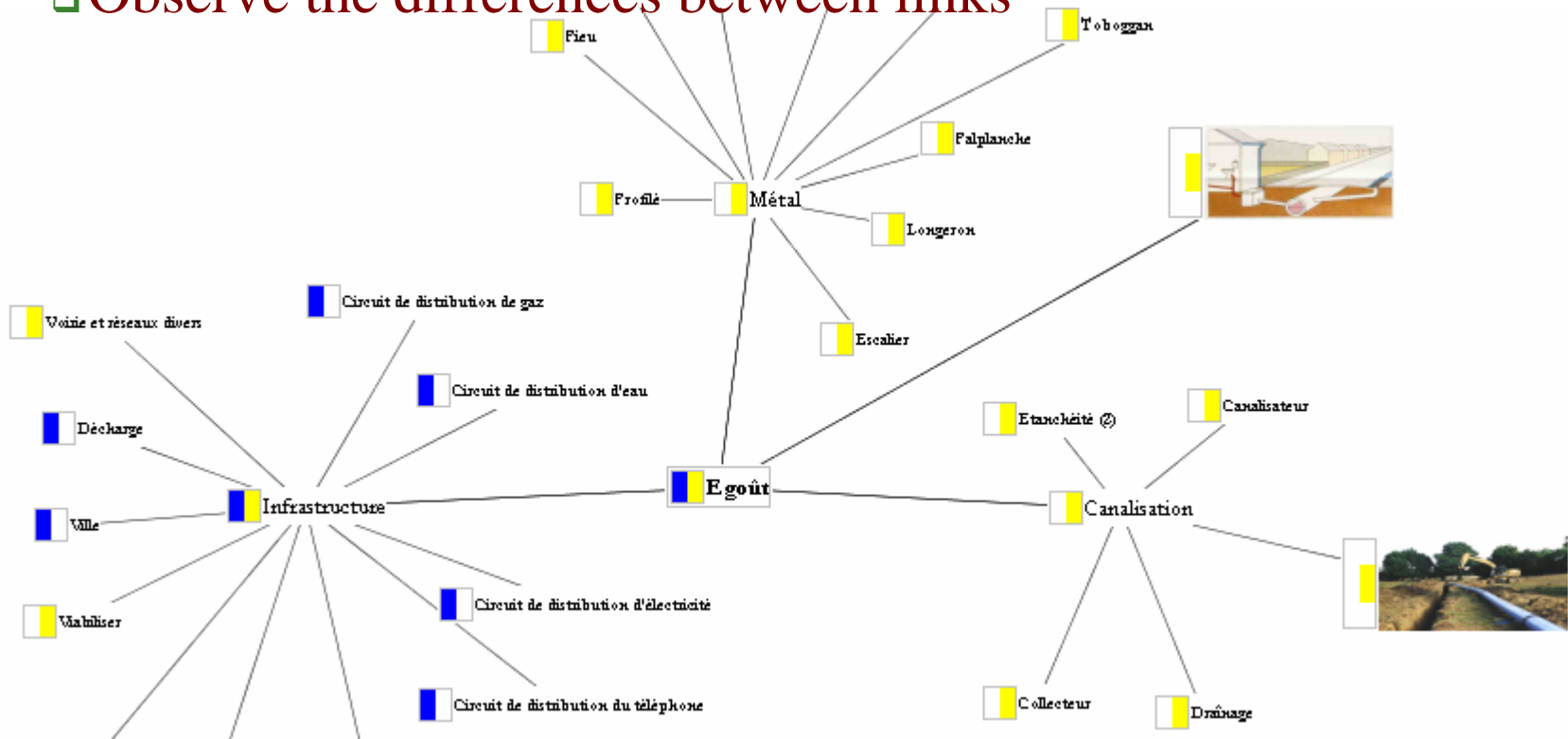
Comments

OK Never mind...

ArcEnCiel - Actions

■ Observe

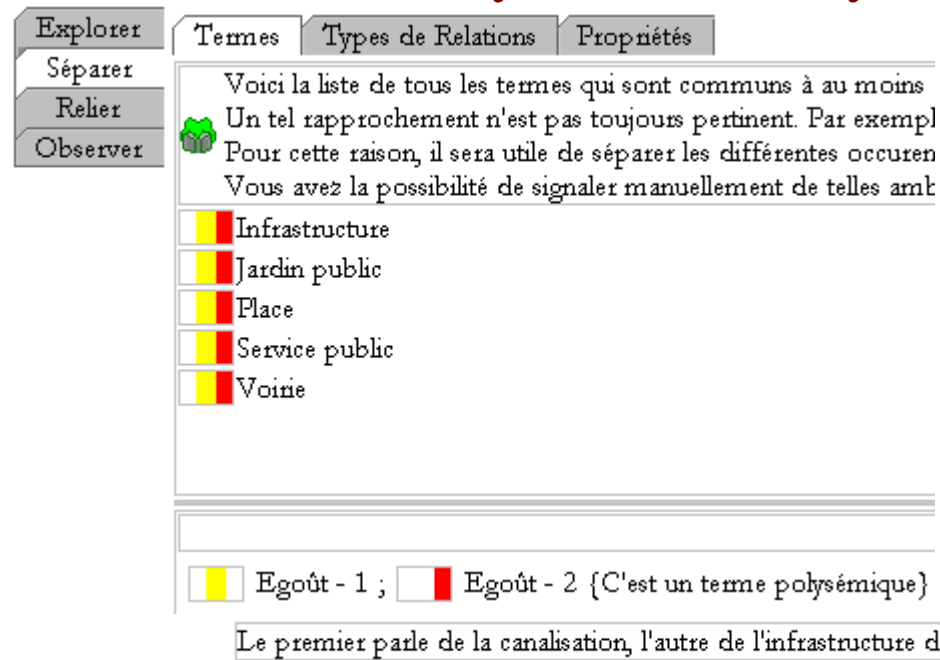
▣ Observe the differences between links



ArcEnCiel – Actions

■ Divide

- ❑ Cancel previously made connections, or split items which would have been automatically connected by error.





The screenshot shows the ArcEnCiel software interface. On the left, a vertical menu contains the following options: Explorer, Séparer, Relier, and Observer. The 'Séparer' option is highlighted. The main window has three tabs: Termes, Types de Relations, and Propriétés. The 'Termes' tab is active, displaying a text box with the following content:

Voici la liste de tous les termes qui sont communs à au moins
Un tel rapprochement n'est pas toujours pertinent. Par exemple
Pour cette raison, il sera utile de séparer les différentes occurrences
Vous avez la possibilité de signaler manuellement de telles anomalies

Below the text box, there is a list of terms, each preceded by a small icon consisting of a yellow square and a red square:

- Infrastructure
- Jardin public
- Place
- Service public
- Voie

At the bottom of the window, there is a legend for the terms:

 Egoût - 1 ;  Egoût - 2 {C'est un terme polysémique}

Below the legend, there is a text box containing the following text:

Le premier parle de la canalisation, l'autre de l'infrastructure d'évacuation entière.

Conclusion

- Computer-assisted Viewpoint Confrontation
- Managing heterogeneity:
 - Different formalisms
 - Different languages
- In order to focus on the difference of content.
 - Learning from heterogeneity before attempting to reduce it.
 - Taking into account “deep” aspects of heterogeneity.

Questions