



# Conception d'un entrepôt de données géographiques comme outil d'aide à l'analyse criminelle

Jean-Paul Kasprzyk, doctorant  
Université de Liège



# Hypothèse: un entrepôt de données géographiques est un outil adapté à l'analyse criminelle

- Qu'est-ce que l'analyse criminelle?
- Quelles sont les données criminelles?
- Qu'est-ce qu'un entrepôt de données ou comment intégrer les données d'analyse criminelle ?
- En quoi l'entrepôt de données est-il adapté à l'analyse criminelle ?



# Qu'est-ce que l'analyse criminelle?



# Analyse criminelle

*« Utilisation de techniques informatiques pour schématiser et analyser les données d'enquêtes complexes afin d'identifier des relations entre événements »*





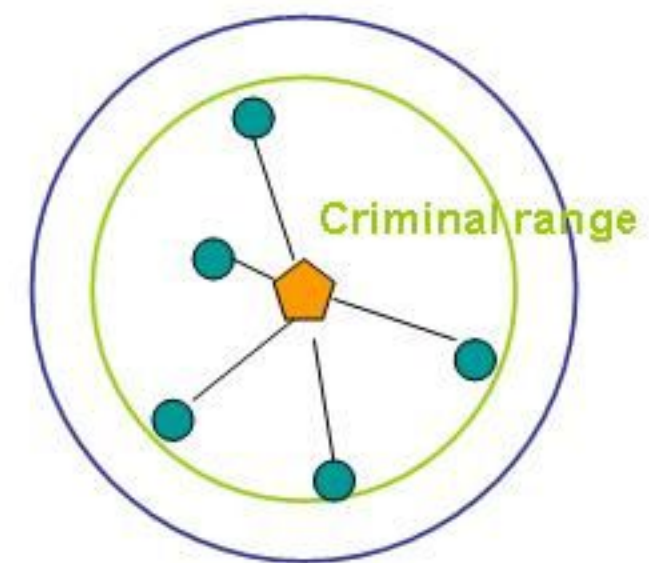
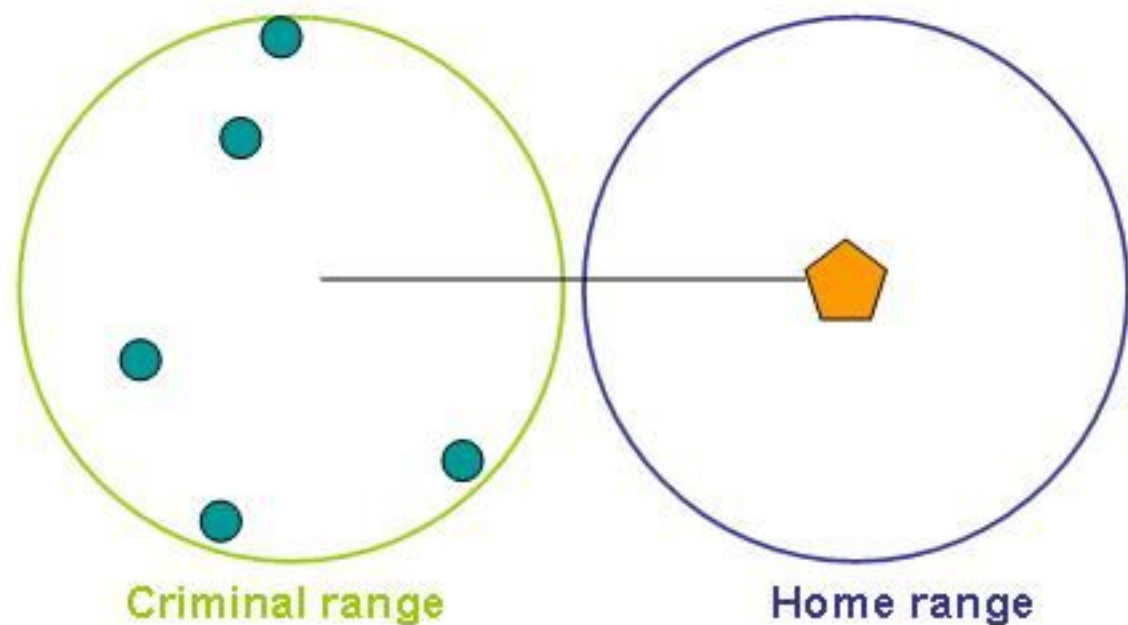
# Analyse criminelle et relations spatiales

- **Profilage géographique**: méthodologie utilisant la localisation d'une série de crimes attribués au même auteur pour déterminer une zone de recherche de son point d'ancrage.
- Distinction **navetteur** et **maraudeur**
- Techniques de profilage dans le cas du maraudeur
  - **Dispersion spatiale**
    - Ellipses de dispersion
    - Surface de densité de Kernel
  - **Interaction spatiale**
    - « Journey to crime »: fonction de distance décroissante

# Analyse criminelle et relations spatiales

Commuter

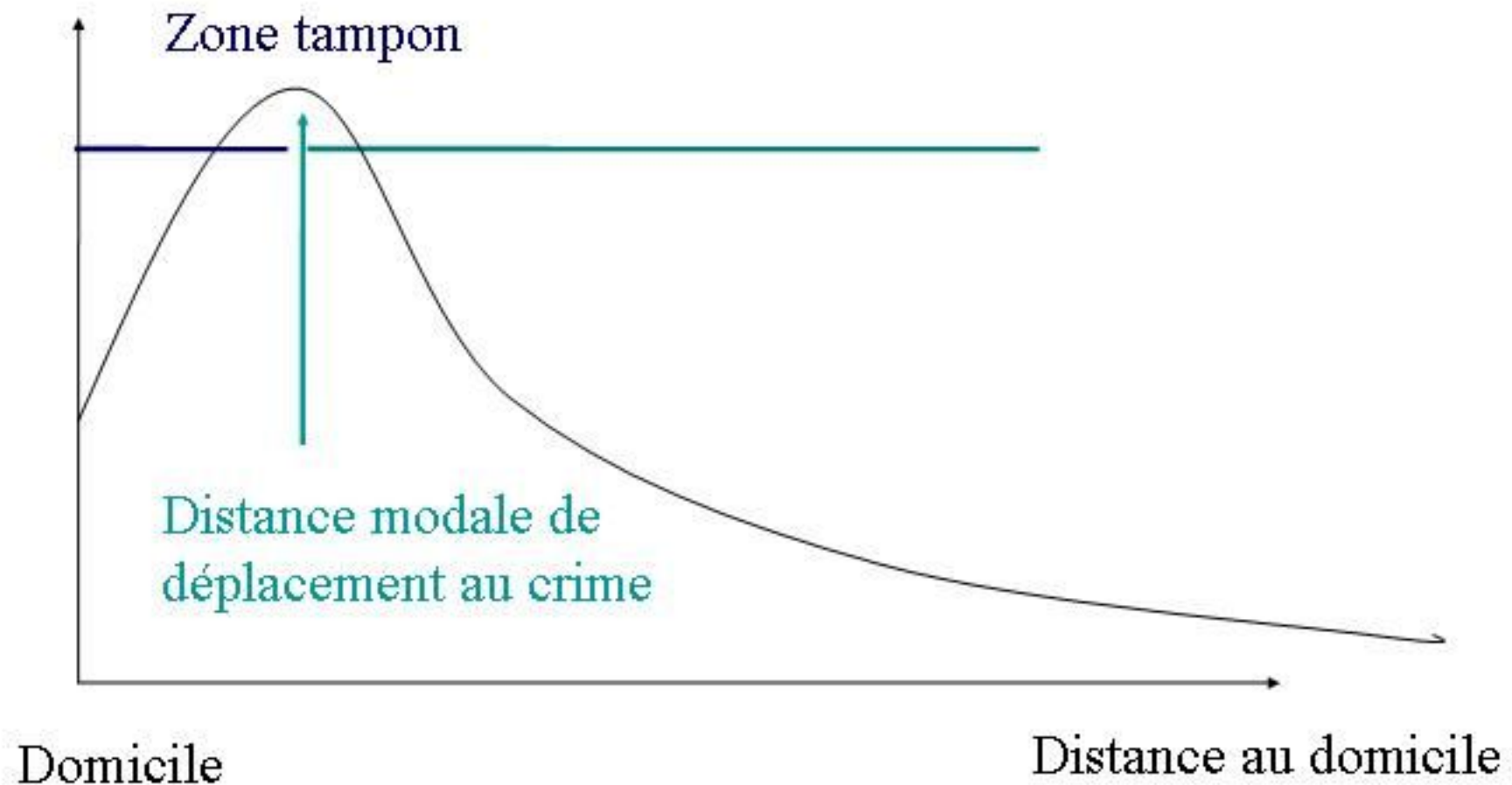
Marauder



*Distinction entre commuter et marauder*  
*Source: Trotta, 2010*

# Analyse criminelle et relations spatiales

Probabilité de crime



*Illustration du principe de distance décroissante*  
*Source: Trotta, 2010*





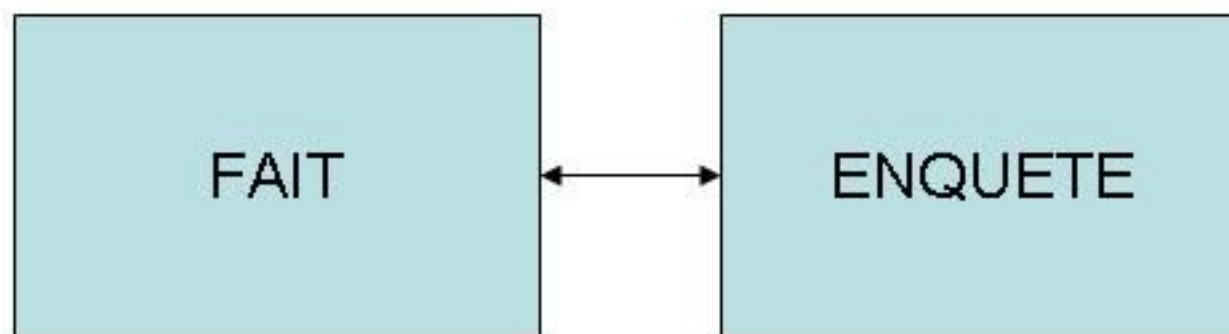
# Quelles sont les données exploitées en analyse criminelle ?





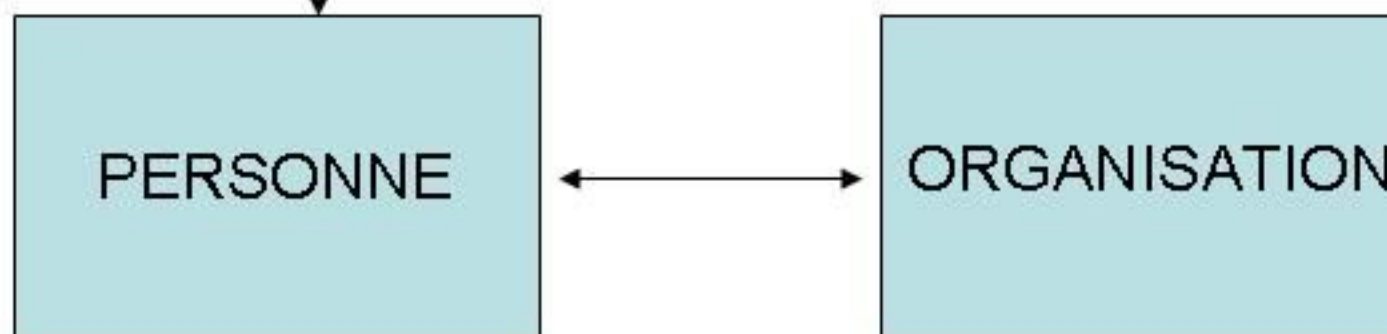
# Données criminelles (BNG)

Niveau 1

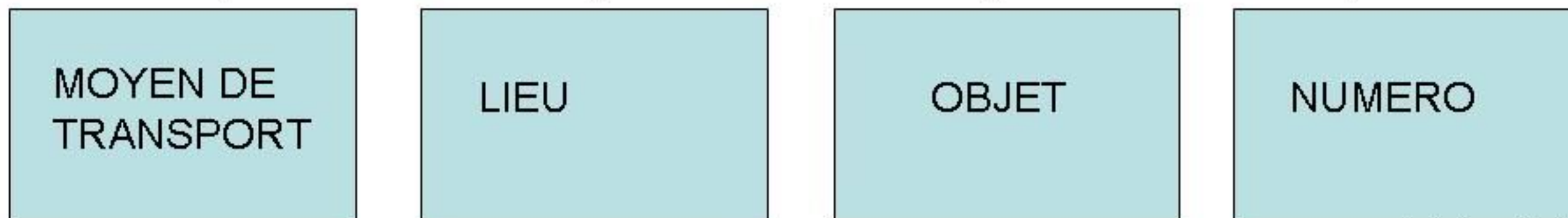


*Structure de la BNG*

Niveau 2



Niveau 3





# Autres données

- Télécommunication, données bancaires, e-mails, données démographiques, données géographiques
  - Caractère **hétérogène** des données utilisées en analyse criminelle:
    - contenu
    - source
    - forme
  - Croisement des données à travers des dimensions thématiques, temporelles et spatiales
- ➔ Pertinence d'un outil informatique permettant l'intégration, le stockage et l'exploitation des données spatio-temporelles: **entrepôt de données**



# Qu'est-ce qu'un entrepôt de données ou comment intégrer les données d'analyse criminelle ?





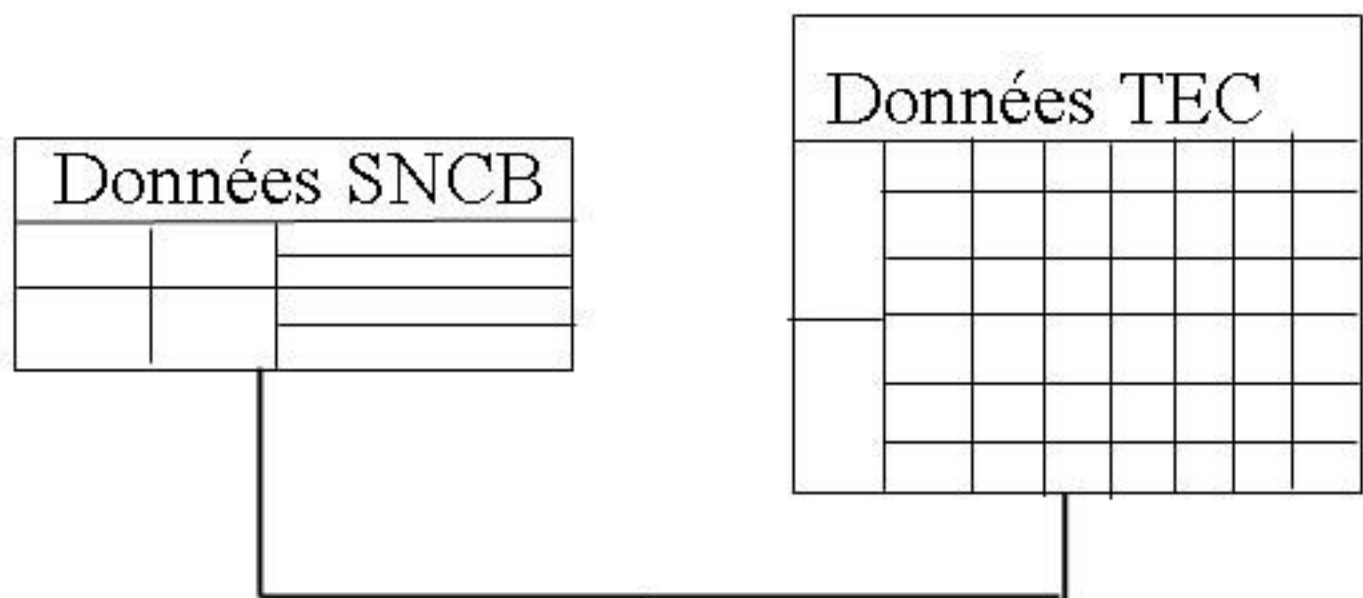
# Entrepôt de données: généralités

- Entrepôt de données = base de données **intégrant** des données de sources diverses afin d'en obtenir une vue compréhensible
- Outil d'**aide à la décision**
- Structure **multidimensionnelle**
- Différents niveaux d'**agrégation**
- Outil **d'archivage**





# Entrepôt de données: intégration



*Exemple d'intégration de données dans un entrepôt*

ID	Arrêt	Transport	X	Y
<b>1</b>	<b>Liège-Guillemins</b>	<b>Train</b>	<b>234660</b>	<b>146560</b>
<b>2</b>	<b>Liège-Guillemins</b>	<b>Bus</b>	<b>234660</b>	<b>146560</b>

Départ	Ligne
<b>14:02</b>	<b>235</b>
<b>14:05</b>	<b>245</b>
<b>14:08</b>	<b>160</b>
<b>14:10</b>	<b>82</b>



# Entrepôt de données: structure

- **Dimensions**

- Chaque donnée contient une dimension spatiale, une dimension temporelle et un certain nombre de dimensions thématiques

- **Membres**

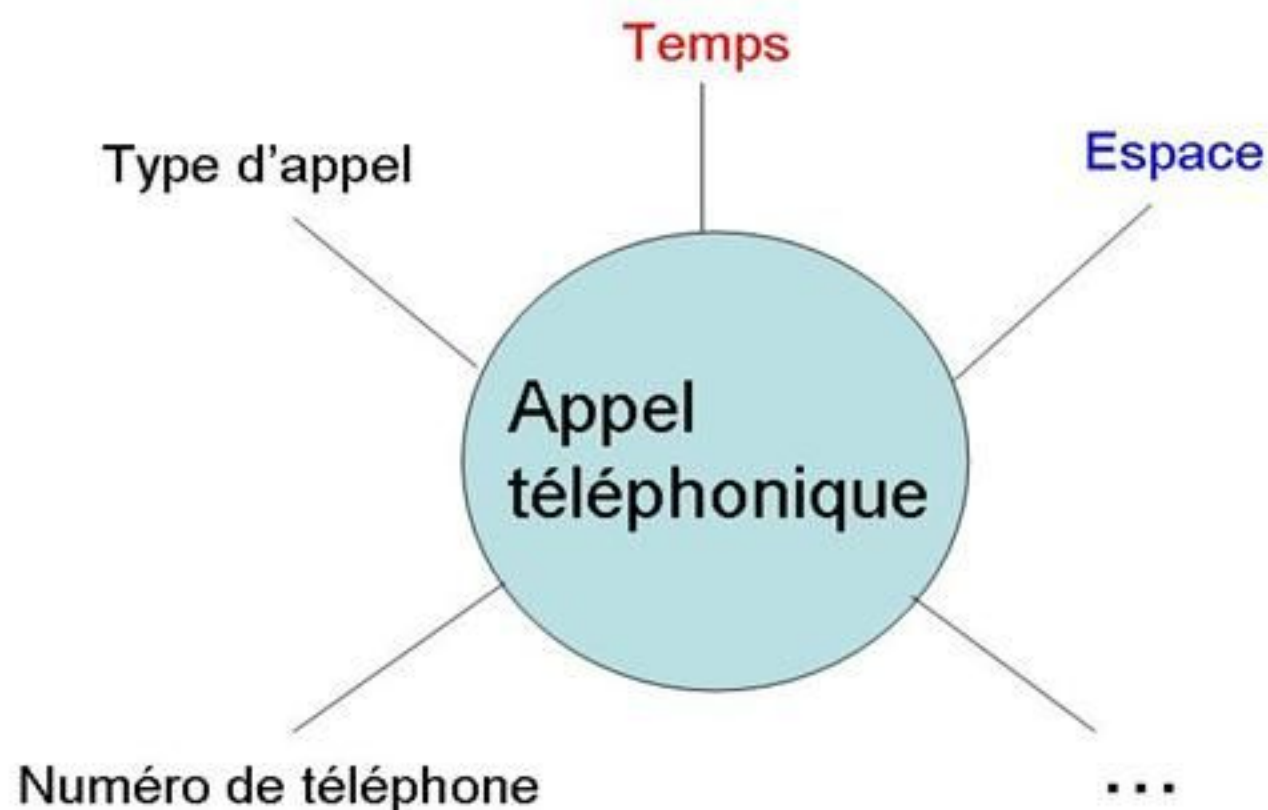
- Une dimension contient des membres :
  - « Septembre 2009 » est un membre de la dimension « temps »
- Les membres d'un niveau peuvent être agrégés au niveau supérieur :
  - « Janvier 2009 » + « Février 2009 » + ... + « Décembre 2009 » = Ensemble des données pour « Année 2009 »
  - Passage du niveau « mois » au niveau « année »

- **Mesures**

- Une mesure est un attribut numérique se retrouvant à l'intersection de plusieurs dimensions



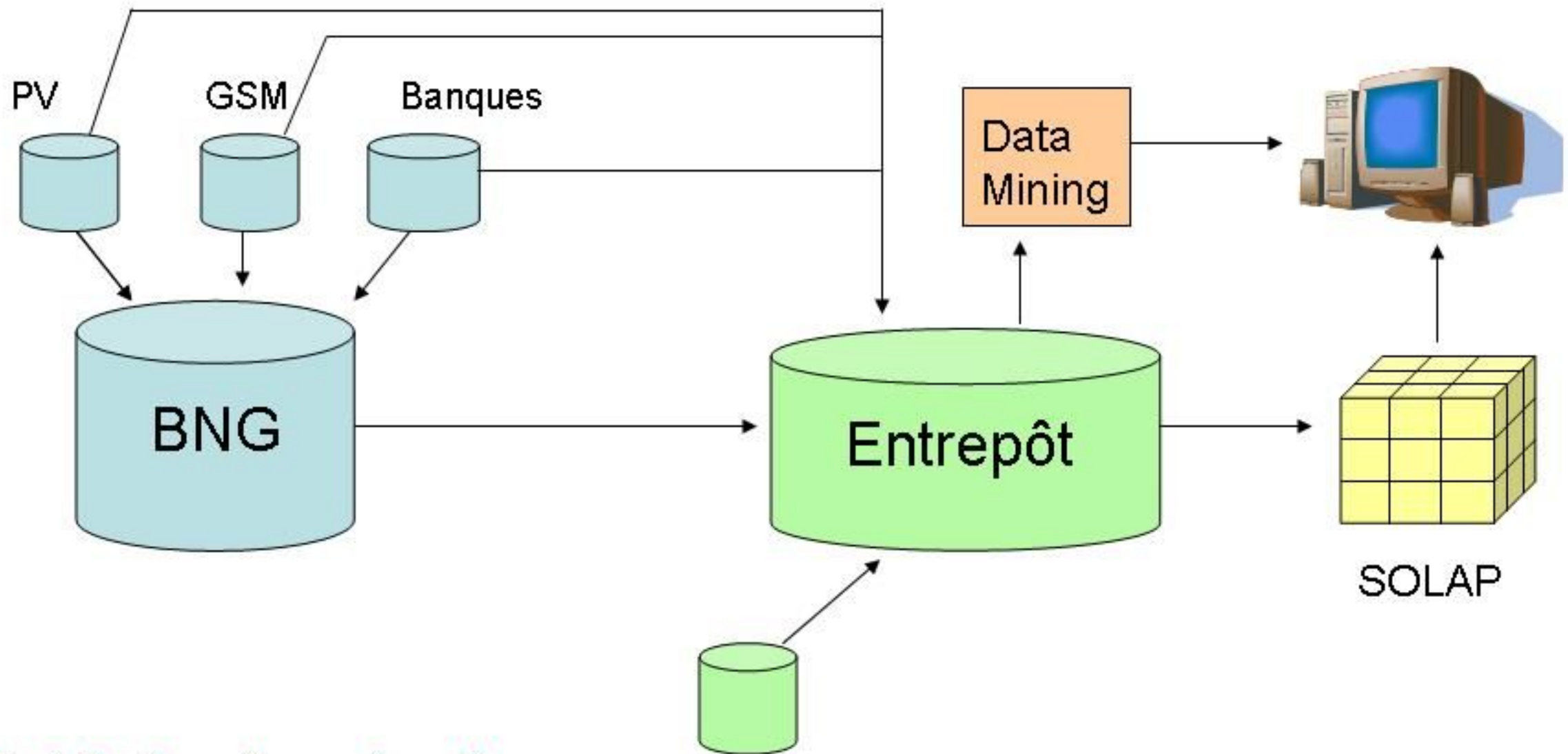
# Entrepôt de données: structure



Différentes agrégations pré-calculées pour améliorer la vitesse des requêtes



# Entrepôt de données: architecture



*Architecture d'un entrepôt de données*





**En quoi l'entrepôt de données  
est-il adapté à l'analyse  
criminelle ?**

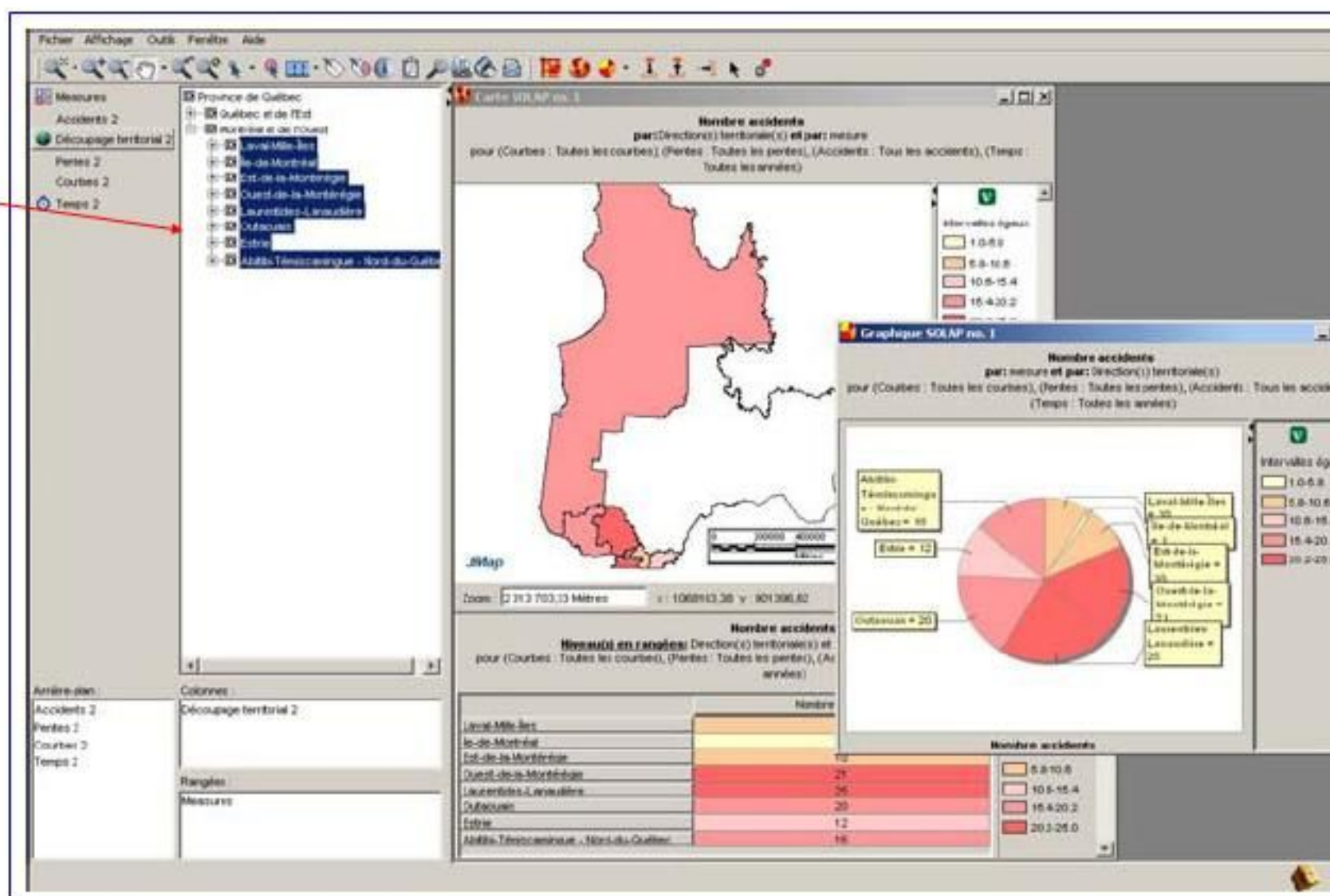


# Serveur SOLAP et interface client

- **Serveur SOLAP et interface client**
  - Le serveur SOLAP permet à l'utilisateur de **voyager dans l'hypercube de données** de manière conviviale
    - « *point & click* »
      - pas de connaissance nécessaire en matière de base de données
  - Actions possibles via l'interface SOLAP
    - Figurer une ou plusieurs dimensions (ex : espace et temps)
    - Passer d'une dimension à l'autre (thématique ou spatio-temporelle)
    - Voyager dans le niveau de détails (= niveau d'agrégation)
    - Ajouter des dimensions (avec respect de la sémiologie cartographique)
    - ...
  - Affichage des données sous plusieurs formes
    - Cartes
    - Diagrammes
    - ...

# Serveur SOLAP et interface client

Arbre des dimensions



*Interface de JMAP Spatial OLAP ©  
(ici une application québécoise sur les accidents de la route)*

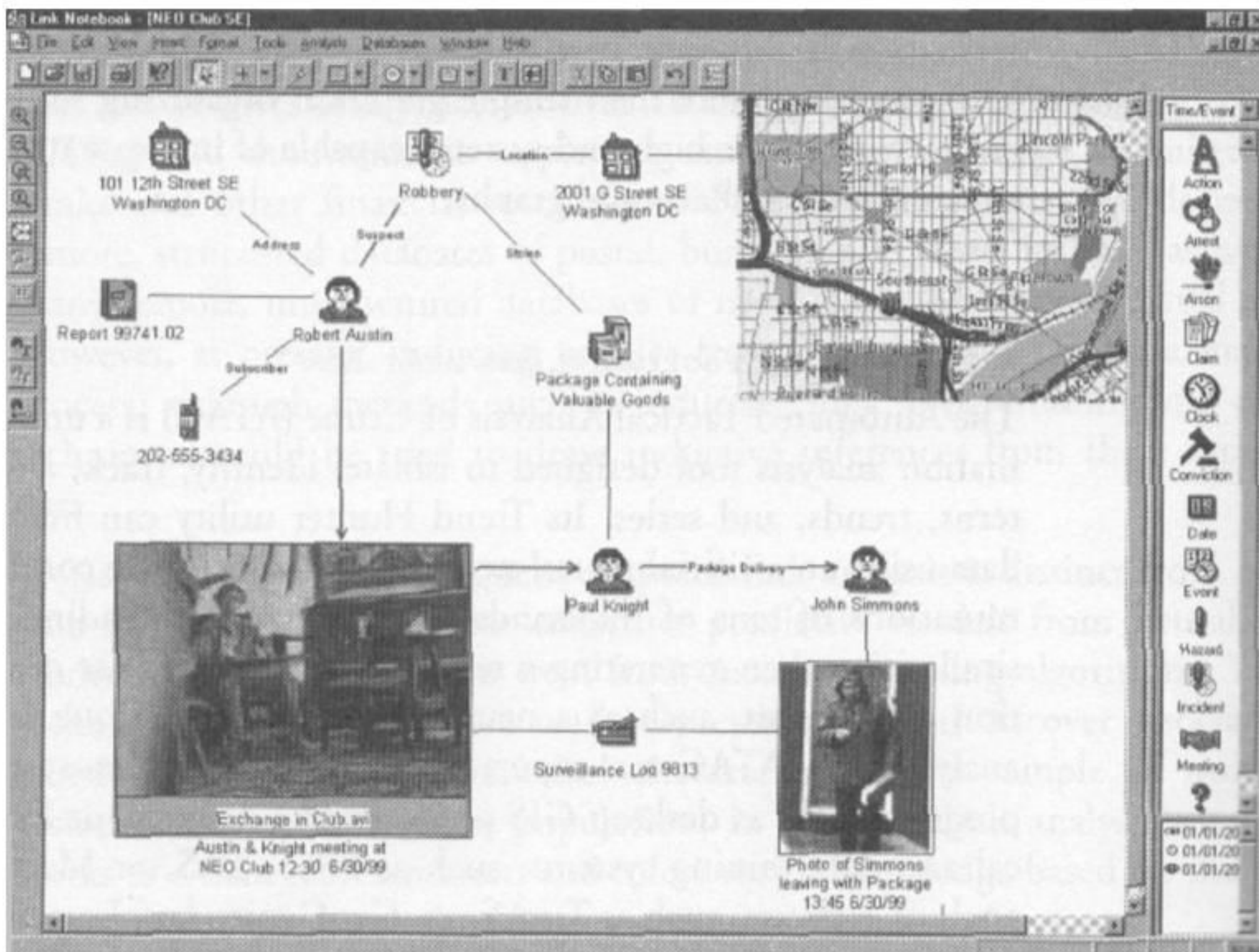


# Fouille de données: généralités

- Fouille de données = ensemble des techniques de visualisation, organisation, classement, groupement et segmentation des données
- Utilisé par les entreprises pour **prévoir** la bonne offre, à la bonne personne, au bon moment
- Analyse criminelle:
  - Quel type de crime, à quel moment et à quel endroit?
  - Associer des crimes à des séries
- Analyse basée sur de la modélisation statistique
- Plusieurs types de data mining en analyse criminelle
  - Fouille de texte
  - Arbres de décision
  - Réseaux neuronaux
  - Cartes auto-organisées
  - Analyse de liens
  - ...



# Fouilles de données: analyse de liens



*Interface de Link Notebook*

©  
*Source: Mena, 2003*



# Conclusions

- Entrepôt de données adapté à l'analyse criminelle:
  - Intégrations des données de source et formats divers dans l'**entrepôt**
  - Exploration et croisement des données à différents niveaux d'agrégation via **SOLAP**
  - Visualisation et groupement des données via **data mining**



# Merci de votre attention!



## Questions?