

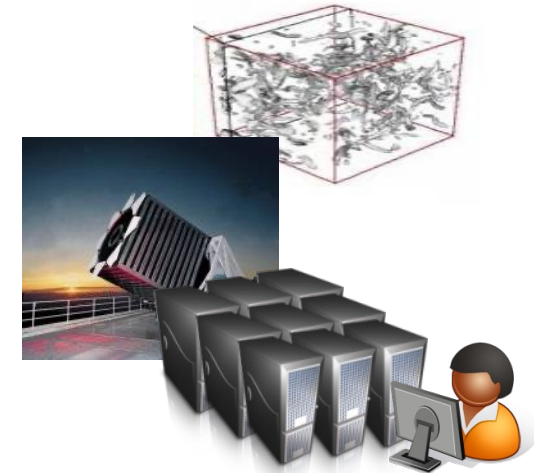
Les paillasses pour les humanités numériques

2000 : vision E-Science : quatrième paradigme sciences physiques et expérimentales

- ▶ Thousand years ago – **Experimental Science**
 - Description of natural phenomena
 - ▶ Last few hundred years – **Theoretical Science**
 - Newton's Laws, Maxwell's Equations...
 - ▶ Last few decades – **Computational Science**
 - Simulation of complex phenomena
 - ▶ Today – **eScience or Data-centric Science**
 - Unify theory, experiment, and simulation
 - Using data exploration and data mining
 - Data captured by instruments
 - Data generated by simulations
 - Data generated by sensor networks
 - Scientists over-whelmed with data
 - Computer Science and IT companies have technologies that will help
- (With thanks to Jim Gray)



$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - K \frac{c^2}{a^2}$$



... Merci à Tony Hey (Microsoft Research)

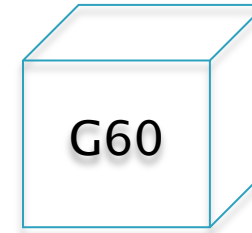
Retour aux années 1960

Le TLF, fiches et groupes binaires

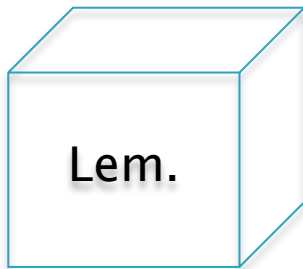
- ▶ Le Gamma 60 produit des concordances
 - parfait pour les mots de faible fréquence.
 - Exemple : *dodécaphonique* (4)
 - Les concordances sont éditées sur « listing »
 - Dans les autres cas : les groupes binaires
 - Co occurrences sur 5 termes sémantiques avant ou après un exemple
 - *Le corbeau fait son nid en haut des arbres.*
 - Le système propose (sur un listing)
 - un ensemble de mots associés avec des listes d'exemples
 - Multiples algorithmes de classement sélection
 - en fonction des fréquences des termes
- ▶ Programmé en assembleur Gamma 60 !!!!

[...] *pour chaque mot*
[...] *pour chaque occurrence*
un contexte de trois lignes,
le mot étudié figurant
obligatoirement
dans la ligne du milieu

Groupes binaires (pour la mémoire)

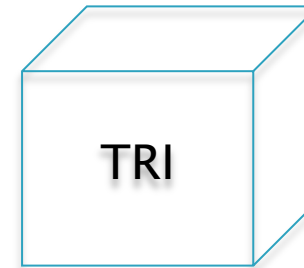


Texte corrigé



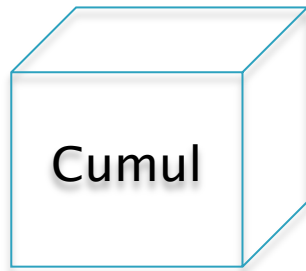
texte

mots



Mots triés

Groupes binaires (suite)

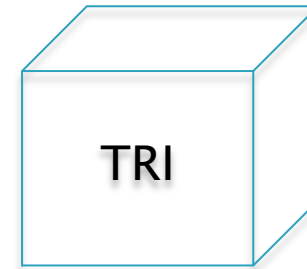


Mots triés

Abaca
Abaca
Abat
....

Couples
Fréquence - mot

2 abaca
5 abat
...



Couples triés

...
5 abat
...
2 abaca

ETC durée 1 mois