

— université
— LUMIÈRE
— LYON 2

Data face !



**Tout savoir sur
les données**

Qu'est-ce qu'une donnée ?



Définitions (dictionnaire Larousse)

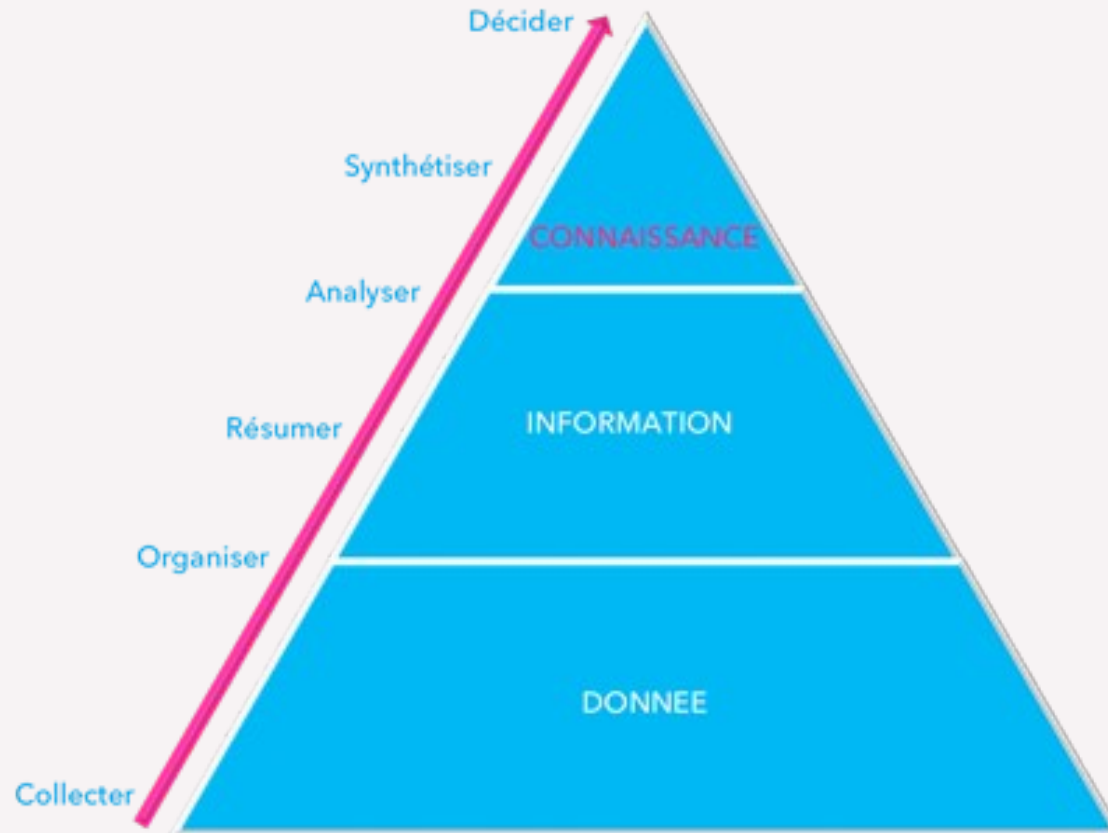
Ce qui est connu ou admis comme tel, sur lequel on peut fonder un raisonnement (...)

Dans un problème de mathématiques, hypothèse figurant dans l'énoncé.

Représentation conventionnelle d'une information en vue de son traitement informatique.

Résultats d'observations ou d'expériences faites délibérément ou à l'occasion d'autres tâches et soumis aux méthodes statistiques.

Donnée, information, connaissance



Information comprise

Donnée interprétée

Élément brut

www.semdee.com

Sondage express

A) Jusqu'ici, tout va bien

B) Je suis déjà perdu·e



Répondre sur <https://toreply.univ-lille.fr>

Question n° 427

Sommaire



Différentes facettes des données



Brève histoire de la gestion de données



Traitement des données, IA et compagnie

Principaux types de données



Type	Description
boolean, bool	Booléen (vrai/faux)
smallint, int2	Entier de -32768 à +32767
integer, int, int4	Entier de -2147483648 à +2147483647
bigint, int8	Entier de -9223372036854775808 à +9223372036854775807
real, float4	Réel simple précision (6 décimales)
double precision, float8	Réel double précision (15 décimales)
numeric/decimal, numeric/decimal(P , S)	Nombre jusqu'à 131072 chiffres avant la virgule (P), 16383 après (S)
character(N), char(N)	Chaîne de caractères de longueur N fixe
varchar, varchar(N)	Chaîne de caractères de longueur variable (maximum N si précisé)
text	Texte long de longueur variable (nécessite des opérateurs spécifiques)
date	Date (ex. 'jj/mm/aaaa')
time	Heure (ex. 'hh:mm:ss')
timestamp	Date et heure (ex. 'aaaa-mm-jj hh:mm:ss')

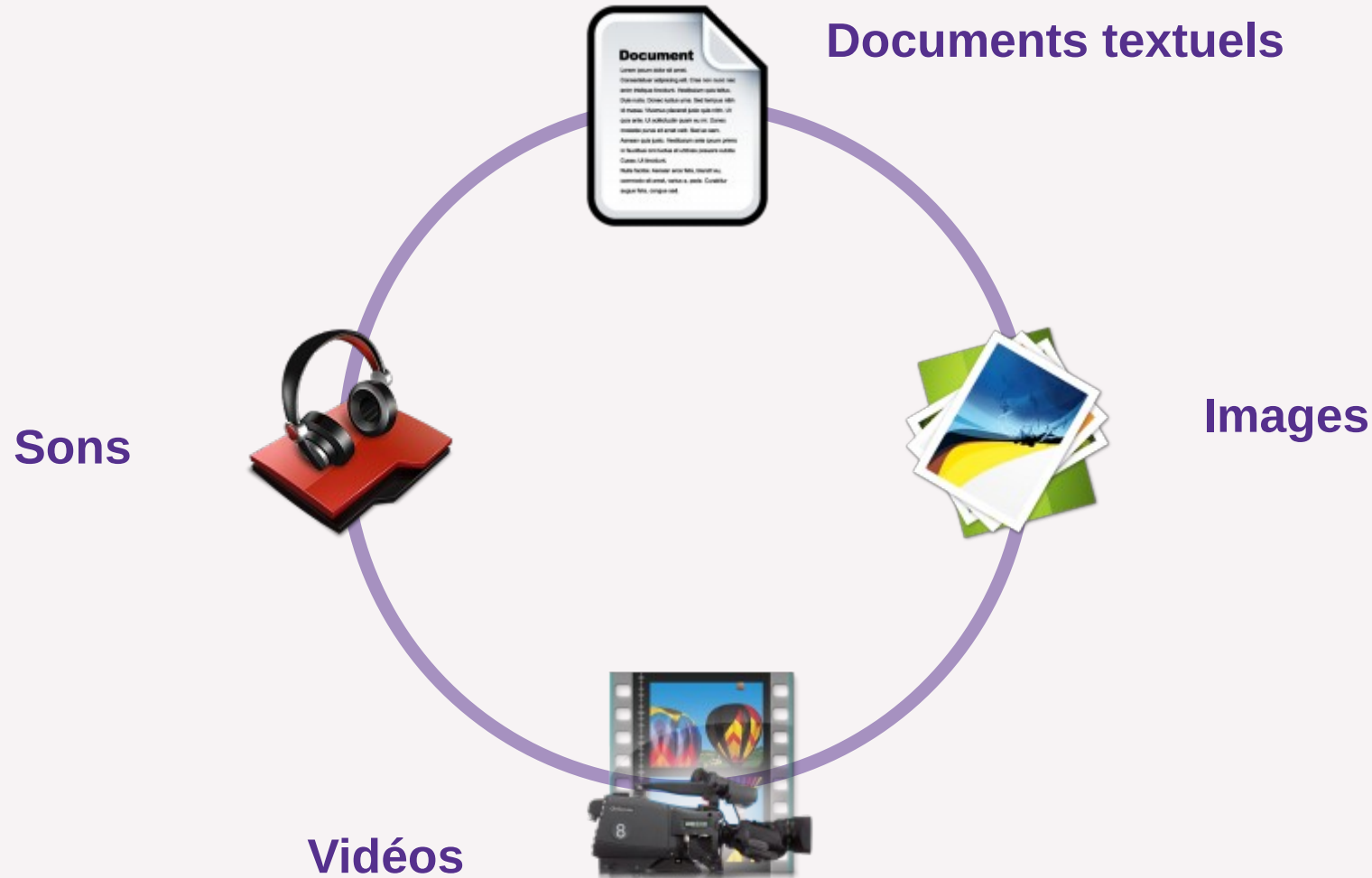


Données structurées

Nom	Prénom	Courriel	Affiliation
Darmont	Jérôme	jerome.darmont@univ-lyon2.fr	Université Lyon 2
Bentayeb	Fadila	fadila.bentayeb@univ-lyon2.fr	Université Lyon 2
Gruenwald	Le	ggruenwald@ou.edu	Oklahoma University



Données non structurées





Données semi-structurées

Bases de données

Documents



Années 1990

N° employé	Date embauche	Nom	Prénom	Adresse	Ville	Province	Code postal	Date naissance
0090	12-jan-89	Charest	Eric	1075 rue Duth	Montreal	QC	H4T 2L9	05-sep-49
0101	20-mar-89	Roux	Kennedy	888 ch. Lacombe	Donat	QC	H9P 3T2	05-déc-53
0105	05-mars-89	Langlois	Claude	308 rue Ste Pierre	Montreal	QC	H2P 1A0	03-mars-58
0105	14-oct-89	Belard	Julie	440 rue Taylor	Laval	QC	H7P 3G8	14-juil-65
0160	07-mars-89	Furte	Jean-Paul	4380 rue Racicot	Montreal-Nord	QC	H2V 2P5	25-juil-59
0330	09-mars-89	Lalonde	Monique	402 rue Victoria	Lingwood	QC	H3P 2T2	02-juil-37
0334	12-avr-89	Therrien	Jacques	320 rue Verchères	Greenfield Park	QC	J4V 2V8	02-juil-71
0367	21-avr-89	Pruss	Marcel	6985 rue St-Laurent	Montreal	QC	H3A 2V3	09-mars-56
0380	12-mai-89	Lalonde	Andrée	89 rue du Collège	Ville St-Laurent	QC	H4P 2M9	05-juil-64
0434	04-juin-89	Connel	Marl	1211 rue de Gault	St-Hubert	QC	J4T 3L9	29-juil-69
0500	10-juin-89	Royette	Alain	888 rue Guertin	Osherson	QC	H3V 2V2	25-juil-67
0550	15-juin-89	Simard	Martin	2575 rue Victoria	St-Lambert	QC	J4L 6P7	03-avr-74
0728	15-juil-89	Belard	Sébastien	3520 rue Rowles	St-Théodore	QC	J7P 3V5	14-juin-53
0839	14-oct-89	Lescur	Josée	684 de Montanille	St-Benoit	QC	J4L 3Y4	12-sep-71
1162	17-oct-89	Foucher	Math	45 rue Harcel	Cardoux	QC	J2T 2B7	31-août-54
1161	14-oct-89	Eric	Julia	621 rue Camille	Laval	QC	H7P 3G7	02-juil-48
1241	07-jan-90	Eric	Julia	621 rue Camille	Laval	QC	J3V 3P5	25-mai-55
1280	01-fév-90	M. in				QC	H3V 2V3	05-mai-58
1341	04-mars-90	Charest	Eric	1075 rue Duth	Montreal	QC	J4P 1A0	13-mai-65
1480	05-mars-90	Charest	Eric	1075 rue Duth	Montreal	QC	H3V 2P2	28-oct-65
1473	07-mars-90	Yassin	Jean-Claude	2058 rue Fleury	Montreal	QC	H4A 1C5	12-oct-62
1481	01-avr-90	Wady	Marwa	301 de Lefebvre	Boucherville	QC	J4C 1T7	07-sep-68
1543	09-juin-91	Francoeur	Louise	8378 rue Sherbrooke est	Montréal	QC	H2T 2V6	30-mai-66
1560	09-juin-91	Charest	Eric	1075 rue Duth	Montreal	QC	H7P 3P8	03-juin-65
1571	14-juin-91	Charest	Eric	1075 rue Duth	Montreal	QC	G01 2X0	17-oct-62
2000	04-juin-91	Charest	Eric	1075 rue Duth	Montreal	QC	J4P 1C3	28-oct-63
2312	04-juin-91	Charest	Eric	1075 rue Duth	Montreal	QC	H3L 5A3	03-oct-70
2335	04-juin-91	Charest	Eric	1075 rue Duth	Montreal	QC	H4T 2Z2	29-oct-69
2329	04-juin-91	Hamel	Naïkate	365 ave. Hudson	Osherson	QC	H2V 2T4	15-mars-61
2321	01-juin-91	Mart	Bernard	6987 rue Viger	Montreal	QC	H3V 2B5	10-mars-73
2322	01-juin-91	Axel	Gilbert	436 rue Dawson	Montreal	QC	H2L 6L9	10-oct-67
2340	01-déc-91	Ouellet	Patricia	7228 boul. St-Joseph	Montreal	QC	H2S 3P5	13-avr-61
2341	05-jan-92	Archambault	Marc	673 rue Bruneau	Montreal	QC	H4S 2T1	31-déc-61
3389	04-fév-92	Brousseau	Guy	2575 Tasseand	Brossard	QC	J4P 2V3	29-jan-60
3480	04-mars-92	Lapard	Monica	802 St-Joseph	Montreal	QC	H2T 6S2	24-mars-73
3500	04-mars-92	Piquette	Pierre	3887 Christophe Colomb	Montreal	QC	H3V 5M1	09-sep-72
3501	15-mars-92	Tremblay	Alexandra	1500 rue Villeneuve	Lingwood	QC	J3P 6V3	11-mars-72
3550	15-mars-92	Tremblay	Alexandra	524 ave. Dumont	Boucherville	QC	J4C 8A2	14-avr-69
3551	15-mars-92	Tremblay	Francine	2502 rue Penson	Montreal	QC	H4T 3A9	30-mai-66
3560	21-mars-92	Wester	Hans	52 rue Ste Marie	Verchères	QC	J5G 4J2	20-sep-72
3561	21-mars-92	Coppen	Eric	2228 de Beaulieu	Montreal	QC	H2P 2J5	08-oct-49
3580	09-juin-92	Cerniers	Marie-Claire	1275 rue Van Home	Osherson	QC	H2V 6P5	05-nov-58
3600	08-oct-92	Robitaille	Celine	7129 Edouard-Montpetit	Montreal	QC	H3V 6L3	29-jan-74



L'ESTA* la nouvelle procédure d'autorisation de voyage pour aller ou transiter aux États-Unis ¶

De nouvelles formalités d'entrée et de transit sur le territoire américain vont entrer en vigueur dès le 12 janvier 2009. Cela concerne tous les voyageurs qui sont exemptés de visa. ¶

En Europe, ce sont les ressortissants de 22 pays bénéficiaires du Programme d'Exemption de Visa, dont la France, qui sont concernés par cette réforme auxqels s'ajoutent 5 pays asiatiques ¶

Désormais, il faudra avant d'embarquer pour un voyage à destination des États-Unis ou avec un transit aux États-Unis, obtenir une autorisation de voyage "ESTA" ¶

Qui doit soumettre une demande d'autorisation de voyage ESTA ? ¶

A partir du 12 janvier prochain, avant de monter dans un avion ou d'embarquer à bord d'un paquebot à destination des États-Unis, tout voyageur devra remplir via internet le nouveau formulaire d'inscription en ligne, l'ESTA ¶

Attention* Cette formalité est obligatoire pour ¶

TOUT voyageur, qu'il se rende sur le territoire américain ou qu'il y transite ¶

que ce soit pour un séjour de tourisme ou affaires de moins de 90 jours ¶

qu'il soit majeur ou mineur (donc même pour votre bébé de 8 mois et 1 par enfant), accompagné ou non ¶

quel que soit le mode de transport, à bord d'une compagnie maritime ou aérienne. ¶

L'autorisation ESTA n'est pas un visa. Elle nécessite un contrôle supplémentaire préalable afin de faciliter la tâche des officiers d'immigration pour l'identification des voyageurs à risque. ¶

L'autorisation de voyage ESTA assure-t-elle l'admission sur le territoire américain ? ¶

Ce n'est pas automatique. En effet, tout ressortissant ESTI prim d'un pays éligible de la commission sur l'immigration doit se soumettre à un contrôle d'admission à l'entrée aux États-Unis. ¶

Attention* ¶

Dans tous les cas, l'officier d'immigration au poste frontière se prononce sur l'admission comme c'est déjà le cas aujourd'hui ¶

Quelle est la procédure d'obtention d'une autorisation ESTA ? ¶

Le système "ESTA" utilise Internet uniquement ¶

Pour soumettre votre demande d'autorisation ESTA depuis le 1er août 2008, vous devez vous rendre sur le site suivant: <https://esta.cbp.dhs.gov> et suivre les instructions pour répondre aux questions posées ¶

L'accès au site est gratuit et disponible en français (sélectionnez pour cela la langue française dans le menu déroulant en haut à droite) ¶

Attention* ¶

Si vous n'avez pas accès à Internet, vous devez recourir à une tierce personne de votre entourage ou à un agent de voyage. Vous restez légalement responsable des réponses fournies ¶

La plupart des voyageurs obtiendront la confirmation de l'autorisation de façon quasi immédiate ce qui permet les voyages de dernière minute ¶

Langages de requête

Moteurs de recherche

UNIVERSITÉ
LUMIÈRE
LYON 2



Données semi-structurées

Nom	Jérôme Darmont	
Courriel	jerome.darmont@univ-lyon2.fr	
	jerome.darmont@msh-lse.fr	
Affiliation	Université Lyon 2	
Courriel		fadila.bentayeb@univ-lyon2.fr
Nom	Prénom	Fadila
	Nom de famille	Bentayeb
Affiliation		Université Lyon 2
Nom	Gruenwald	
Prénom	Le	
Affiliation	Oklahoma University	

Qui pèse le plus lourd ?

Données
non structurées



Données
structurées et
semi-structurées





Big (mais pas que) data

Volume

Taille des données

Véracité

Qualité des données
Sécurité

Variété

Données (semi-)structurées
et non structurées



#76651093

www.lebigdata.fr

Vélocité

Capteurs, réseaux sociaux

Données opérationnelles vs. données décisionnelles



Données opérationnelles en évolution

Image de la base en Mai 2005

Répertoire

Nom	Ville
Dupont	Paris
Durand	Lyon

Image de la base en Juillet 2006

Répertoire

Nom	Ville
Dupont	Marseille
Durand	Lyon

Données décisionnelles historisées

Schémas de données

Lydie Soler, AgroParisTech/INRA

Calendrier

Code	Année	Mois
1	2005	Mai
2	2006	Juillet

Répertoire

Code	Année	Mois
1	Dupont	Paris
1	Durand	Lyon
2	Dupont	Marseille

Quizz

Les *big data* peuvent être ?

- A) Structurées
- B) Semi-structurées
- C) Non structurées
- D) Opérationnelles
- E) Décisionnelles

Répondre sur <https://toreply.univ-lille.fr>

Question n° 278

<http://eric.univ-lyon2.fr/jdarmont/>

N° de la question : 278



Sommaire



Différentes facettes des données



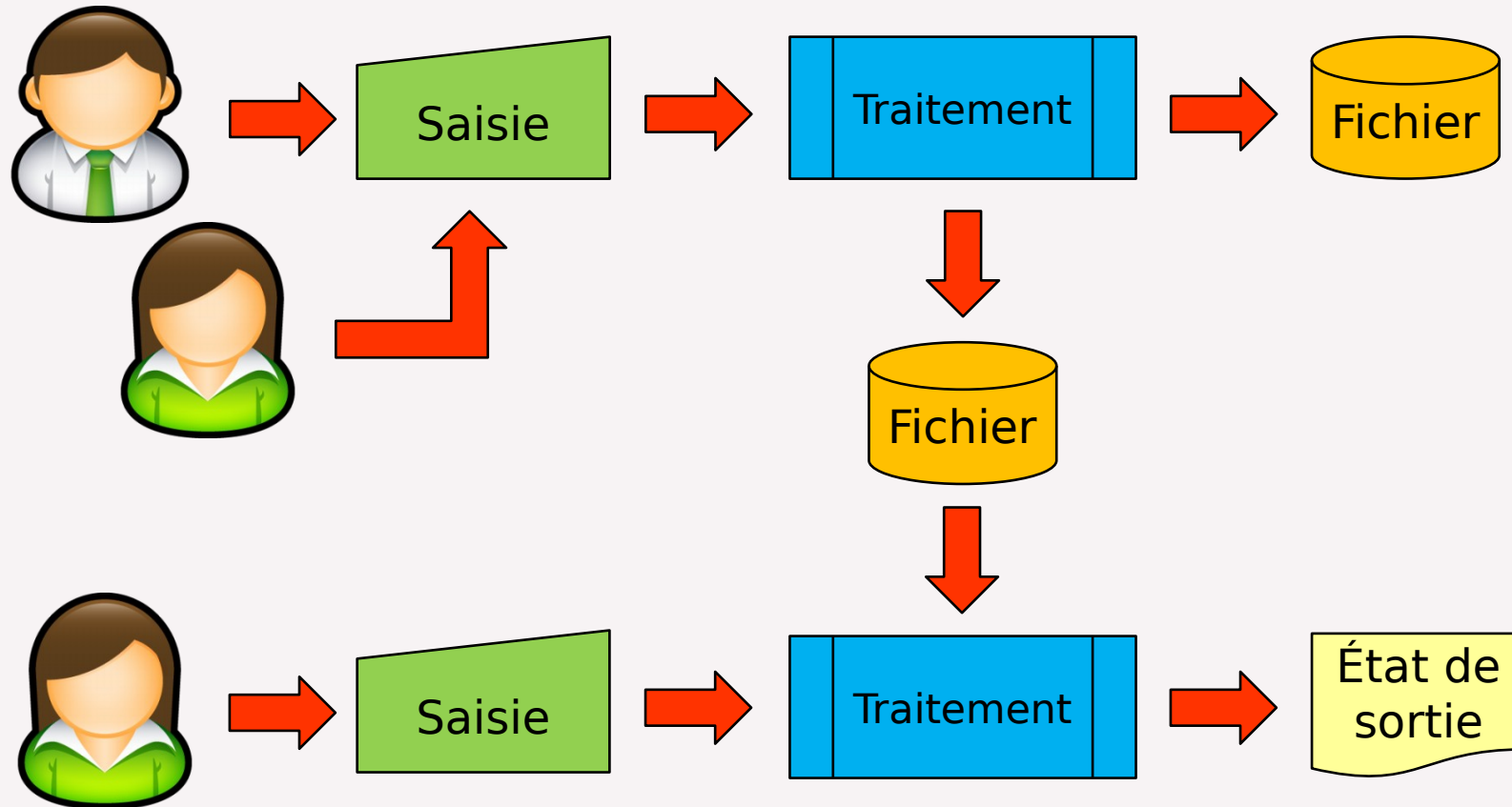
Brève histoire de la gestion de données



Traitement des données, IA et compagnie

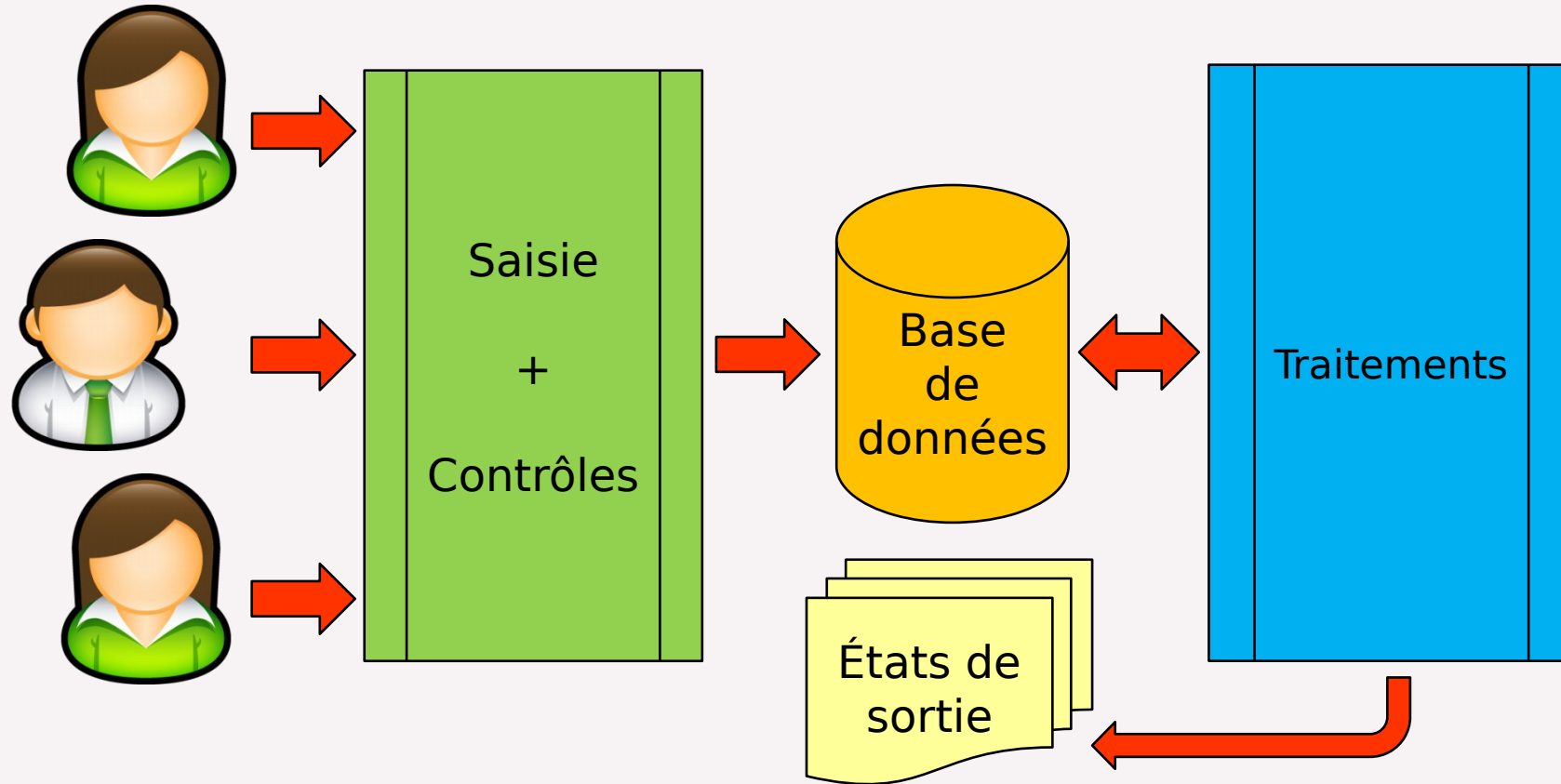


1950-1960 Fichiers





1960-1970 Bases de données



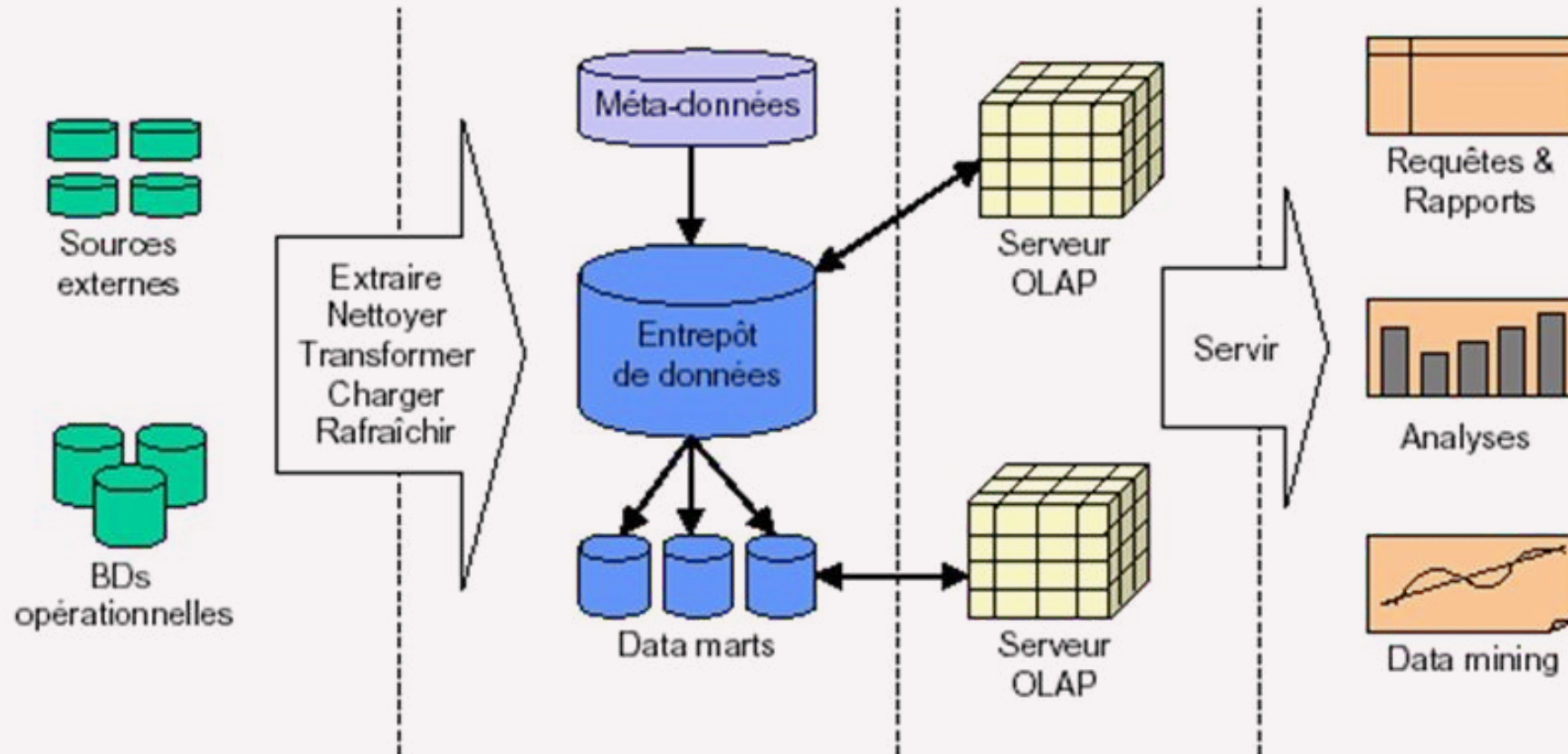
Bases de données relationnelles



Marie Duflot-Kremer
INRIA

L'HISTOIRE DES BASES DE DONNÉES
ou presque ...

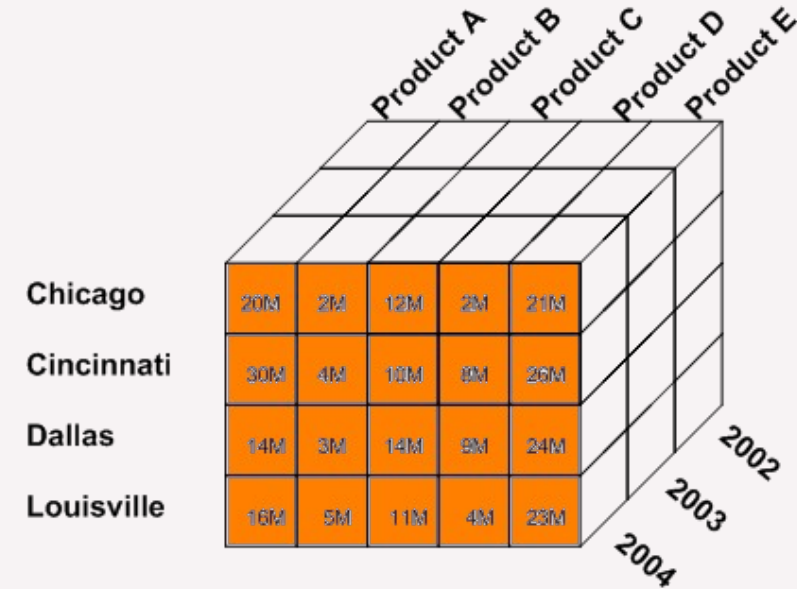
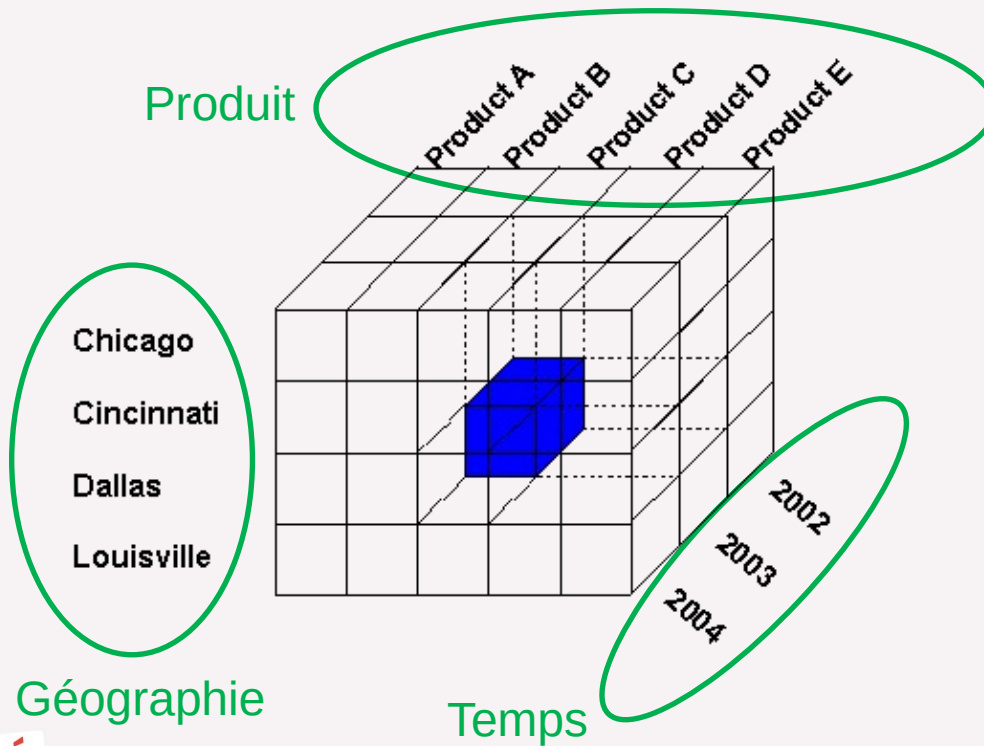
1980-1990 Entrepôts de données





(Hyper)cube de données

- **Fait** : sujet d'analyse
 - Ensemble de **mesures**
- **Dimensions** : axes d'analyse



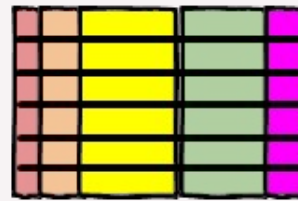
Schémas : docs.oracle.com



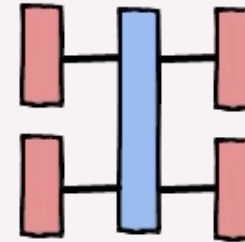
1990-2000 Bases de données NoSQL

Base de données SQL

Relationnelle

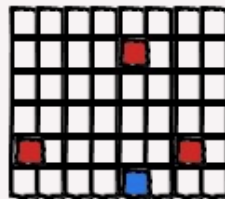


OLAP

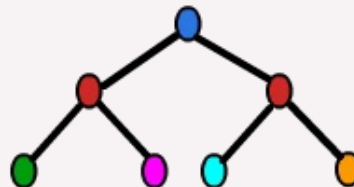


Base de données NoSQL

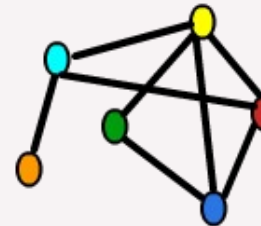
Orienté colonne



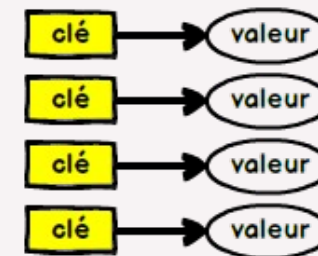
Document



Graphe



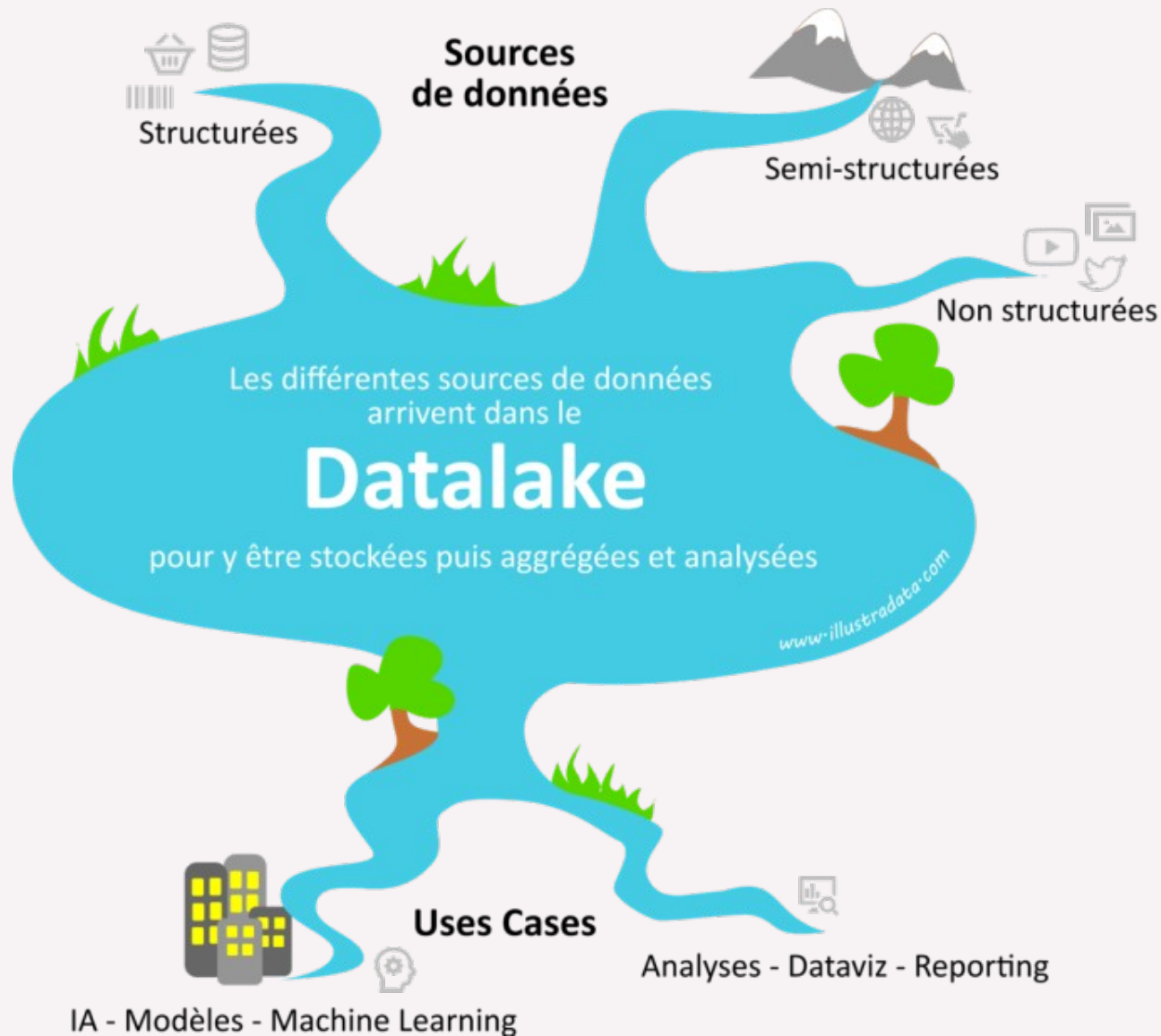
Clé-Valeur



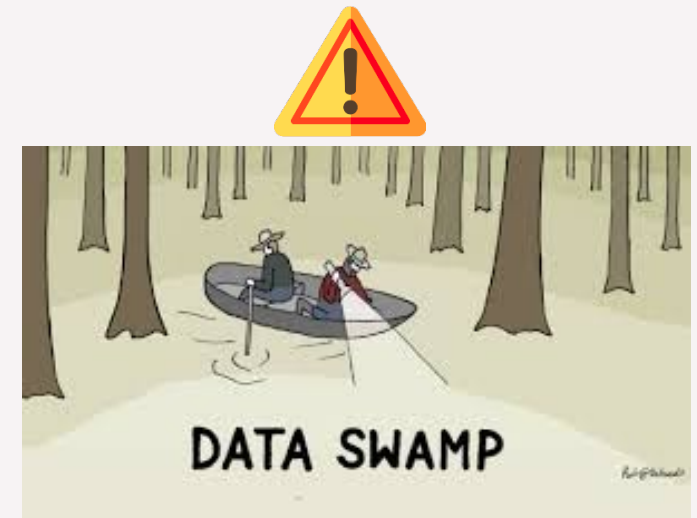
waytolearnx.com

<http://eric.univ-lyon2.fr/jdarmont/>

2010-2020 Lacs de données



illustradata.com



dataedo.com

Quizz

Quels sont les outils de gestion de données les plus courants aujourd'hui ?

- A) Fichiers
- B) Bases de données relationnelles
- C) Entrepôts de données
- D) Bases de données NoSQL
- E) Lacs de données

N° de la question : 925



Répondre sur <https://toreply.univ-lille.fr>

Question n° 925

<http://eric.univ-lyon2.fr/jdarmont/>

Sommaire



Différentes facettes des données



Brève histoire de la gestion de données



Traitement des données, IA et compagnie



Langages de requêtes



Données structurées

```
SELECT nom, prenom  
FROM listeEleves  
WHERE ville = 'Lyon'
```



Données semi-structurées

```
for $e in /listeEleves/eleve  
where ville = 'Lyon'  
return <eleve>  
    $e/nom  
    $e/prenom  
</eleve>
```



Résultat des requêtes



nom	prenom
Lovelace	Ada
Hopper	Grace
Hamilton	Margaret



```
<eleve>
  <nom>Lovelace</nom>
  <prenom>Ada</prenom>
</eleve>

<eleve>
  <nom>Hopper</nom>
  <prenom>Grace</prenom>
</eleve>

<eleve>
  <nom>Hamilton</nom>
  <prenom>Margaret</prenom>
</eleve>
```

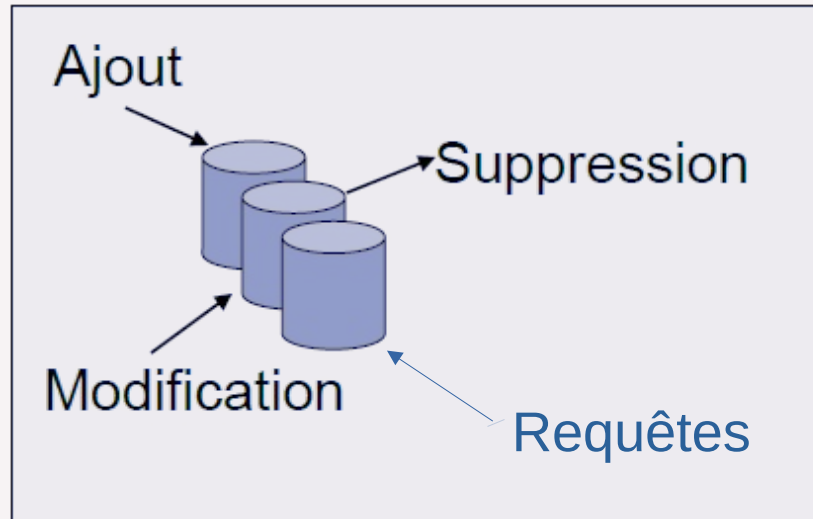


Photos : blog.iakaa.com

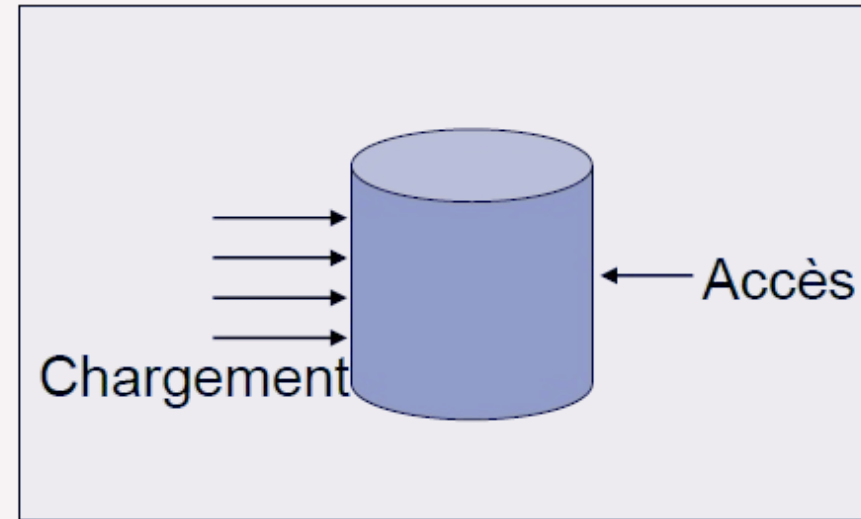


OLTP vs. OLAP

Base de données opérationnelle



Entrepôt de données (décisionnel)



Schémas : Lydie Soler, AgroParisTech/INRA

On **L**ine **T**ransaction **P**rocessing

On **L**ine **A**nalytical **P**rocessing

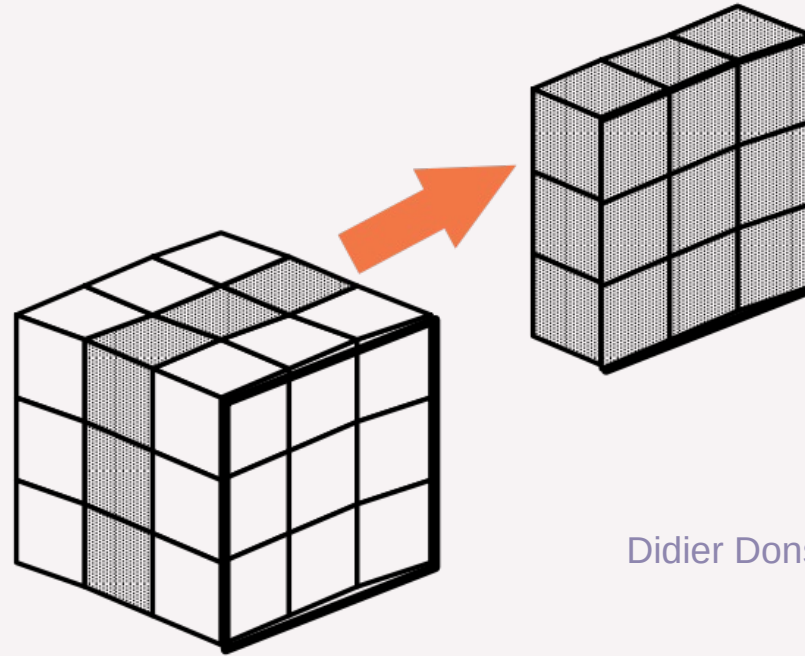
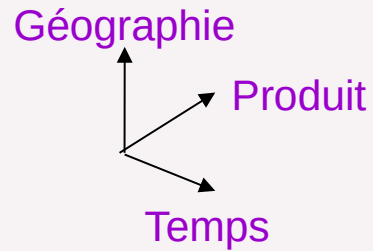


Requêtes OLAP Slice

		1995	1996	1997
Frais	IdF	220	265	284
	Province	225	245	240
Liquide	IdF	163	152	145
	Province	187	174	184



		1996
Frais	IdF	265
	Province	245
Liquide	IdF	152
	Province	174



Didier Donsez, Université Grenoble 1

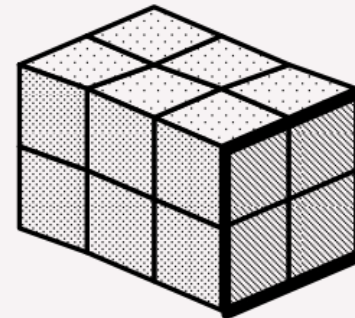
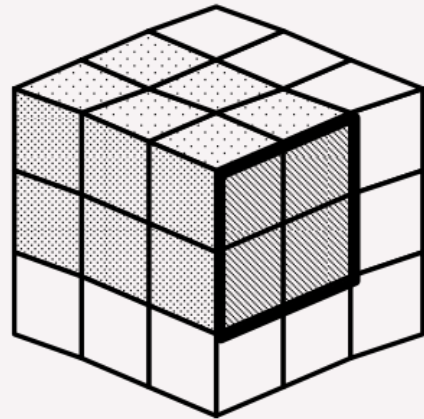


Requêtes OLAP Dice

		1995	1996	1997
Frais	IdF	220	265	284
	Province	225	245	240
Liquide	IdF	163	152	145
	Province	187	174	184



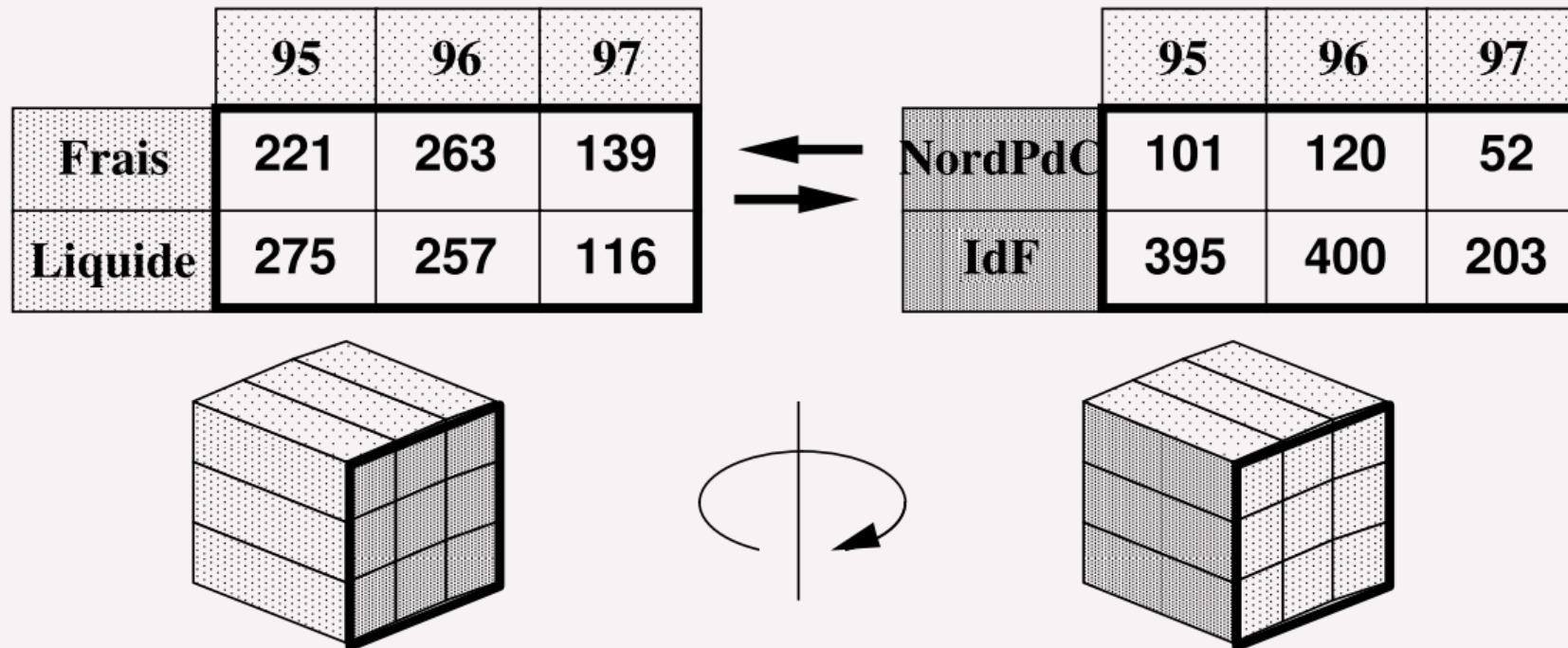
		1995	1996
Frais	IdF	220	265
	Province	225	245



Didier Donsez, Université Grenoble 1



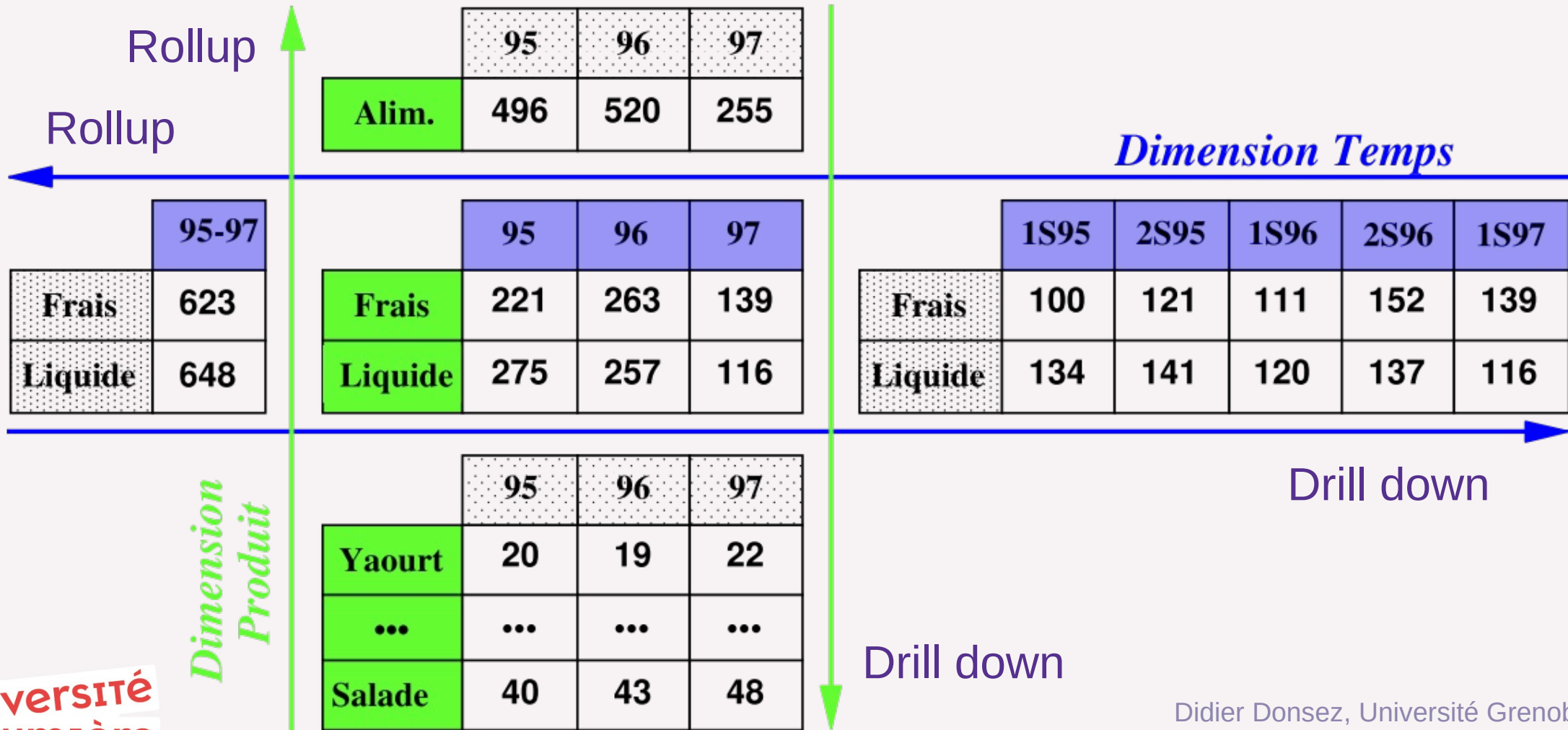
Requêtes OLAP Rotate



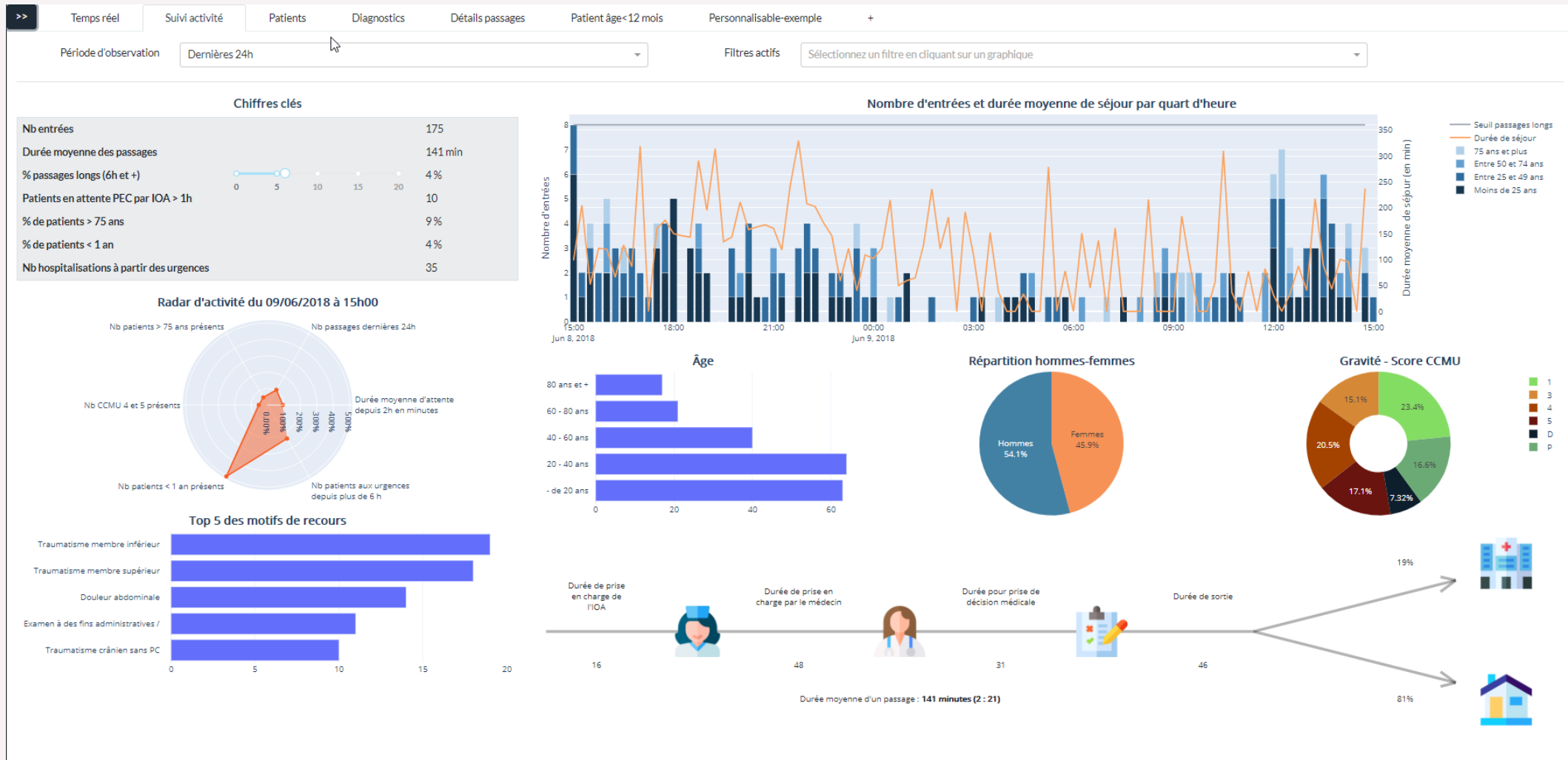
Didier Donsez, Université Grenoble 1



Requêtes OLAP Rollup/Drilldown

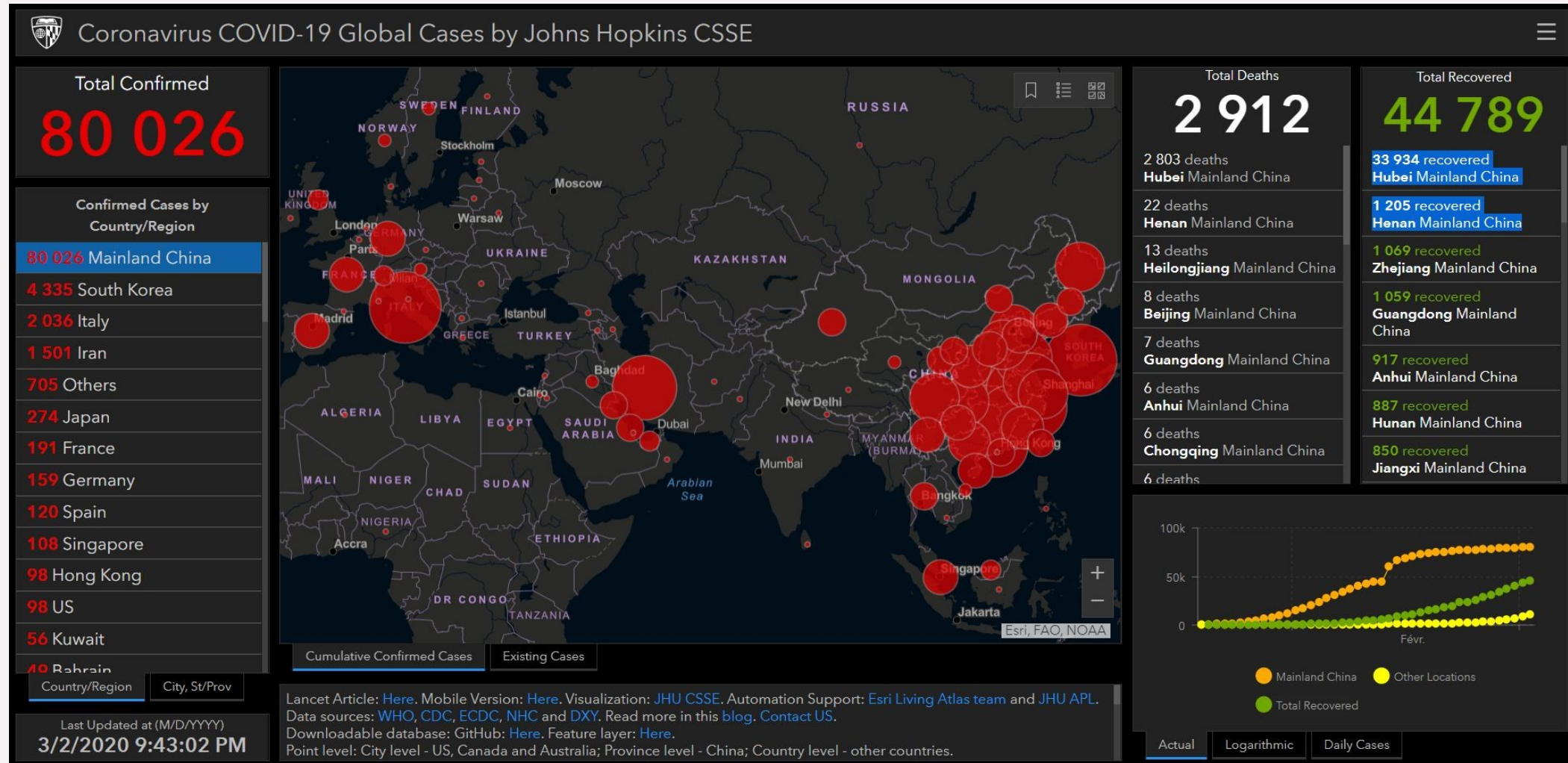


Exemple de tableau de bord





Exemple de tableau de bord



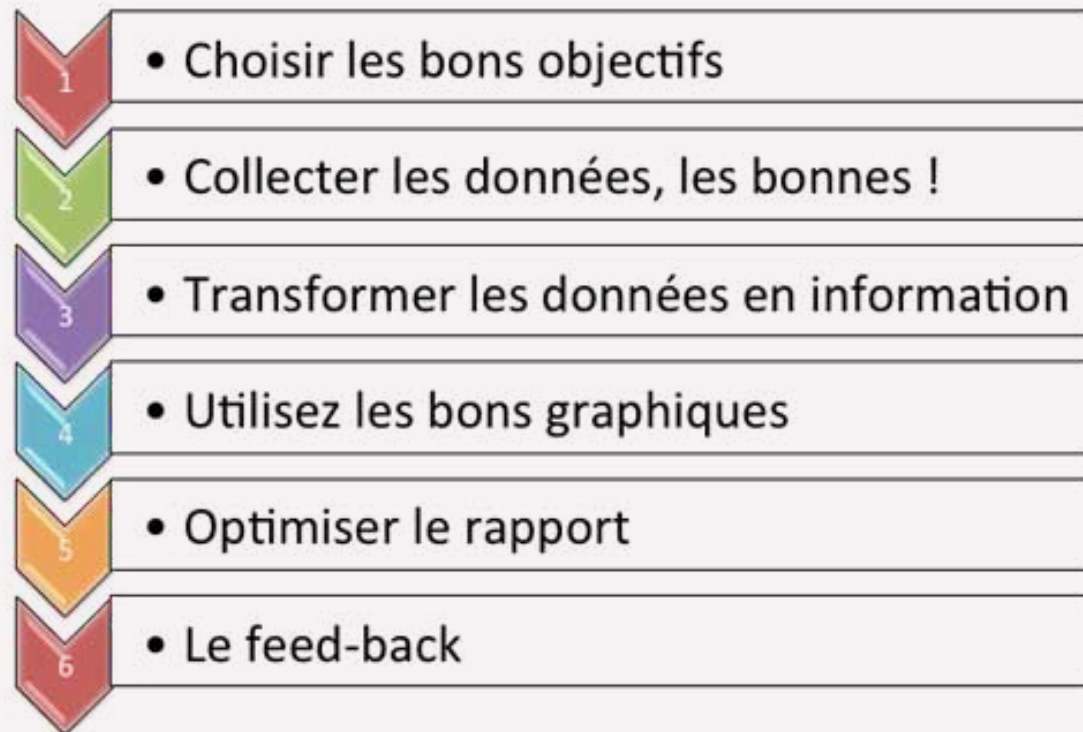
delladata.fr

<http://eric.univ-lyon2.fr/jdarmont/>



Conception d'un tableau de bord

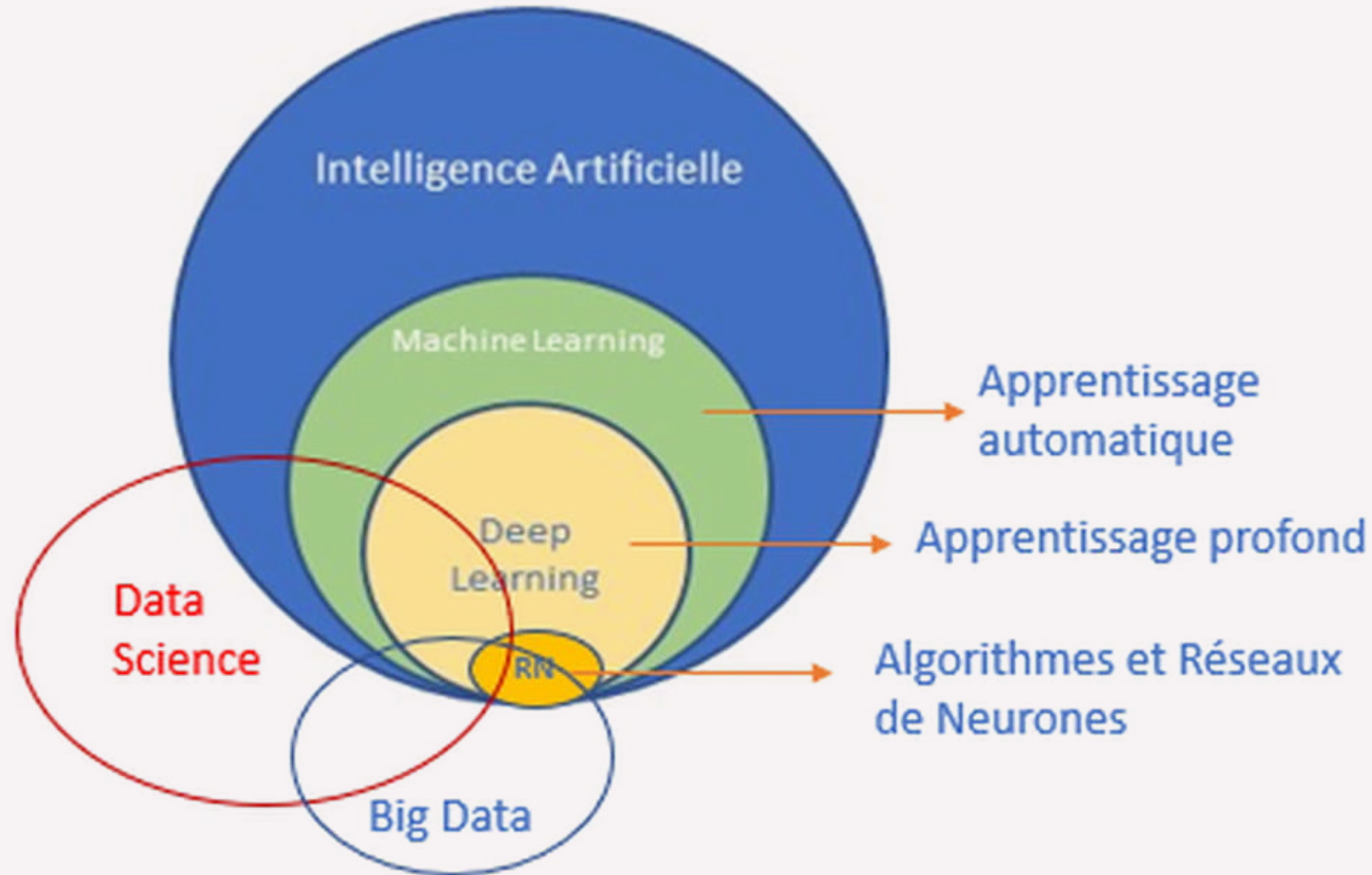
Un reporting réussi



www.piloter.org



Un peu d'intelligence dans ce cours ?



www.journaldunet.com

Apprentissage automatique



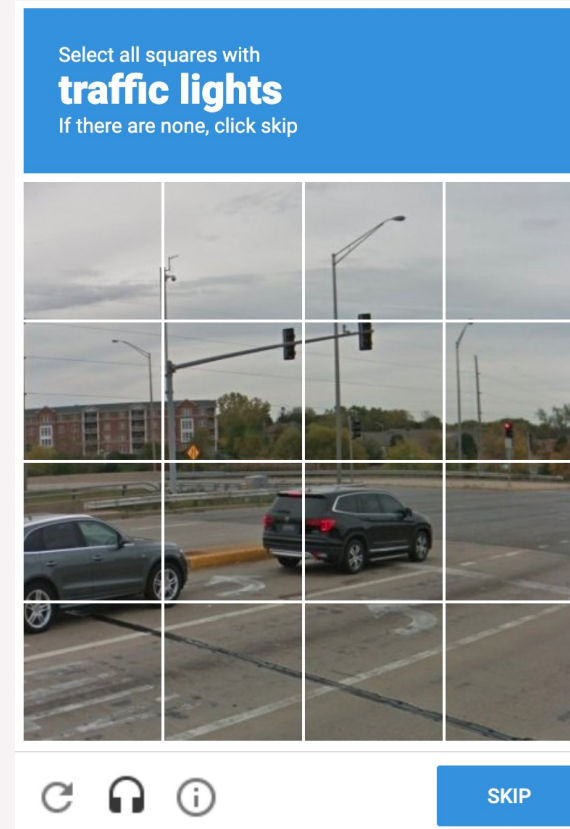
neodigital

Quizz

Quand vous répondez à un captcha, vous participez à de l'apprentissage ?

A) Supervisé

B) Non-supervisé



N° de la question : 374



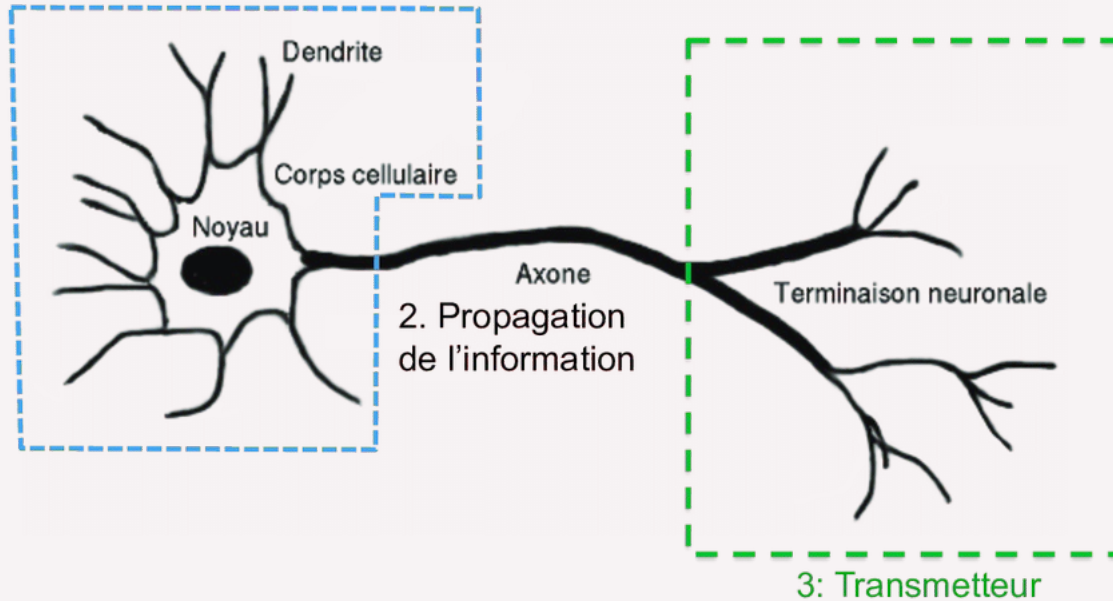
Répondre sur <https://toreply.univ-lille.fr>

Question n° 374



Réseaux de neurones... naturels !

1: Récepteur



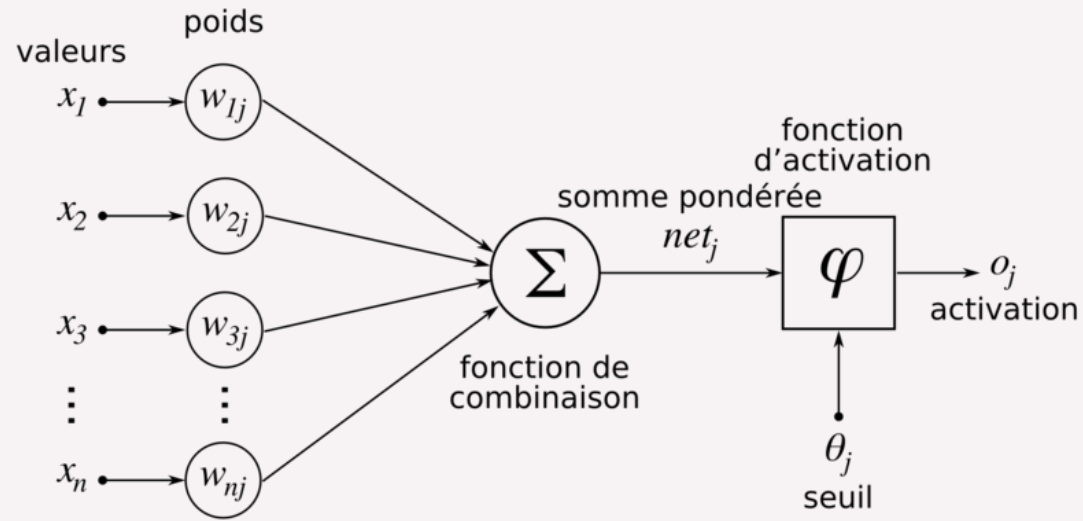
Clément Hébert – www.researchgate.net

inrs.ca

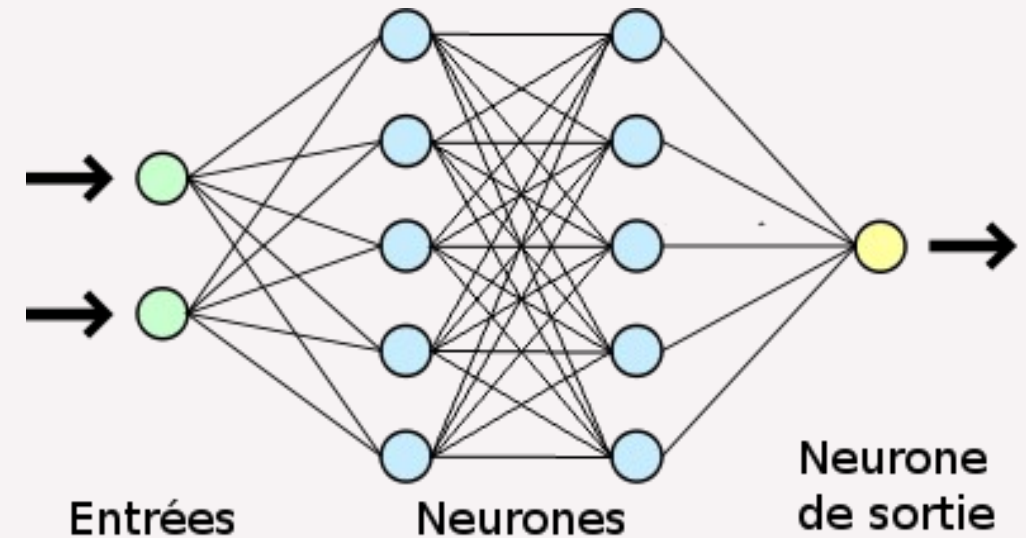




Réseaux de neurones artificiels



www.juripredis.com



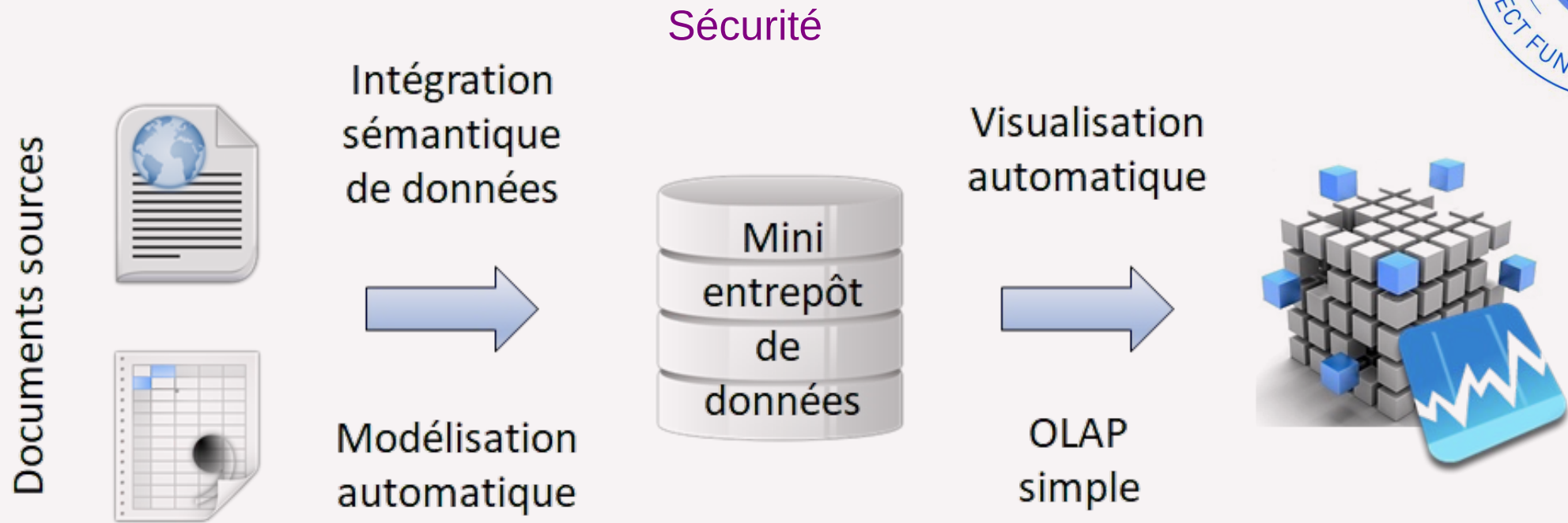
Sommaire

	Différents facteurs des données
	Bibliothèque et la gestion de données
	Travail des données, IA et compagnie

BONUS

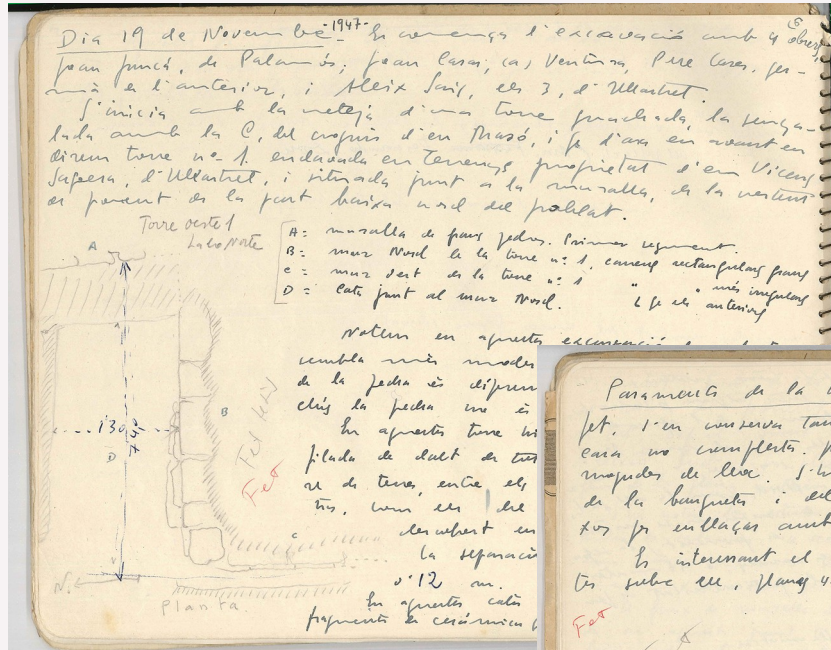
Exemples de projets de recherche

BI4people

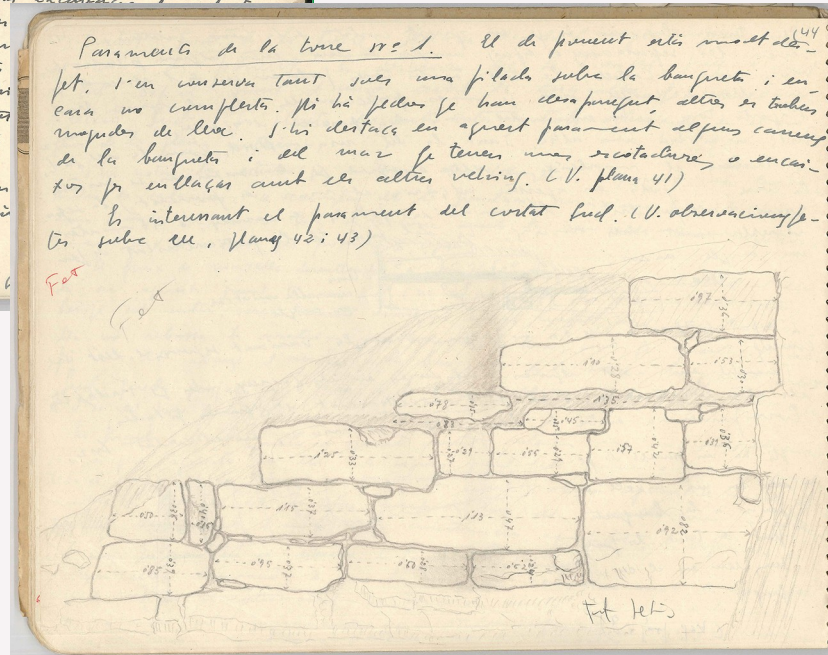


Analyse collaborative

IA et lac de données



Carnets de fouille de Miquel Oliva, 1947

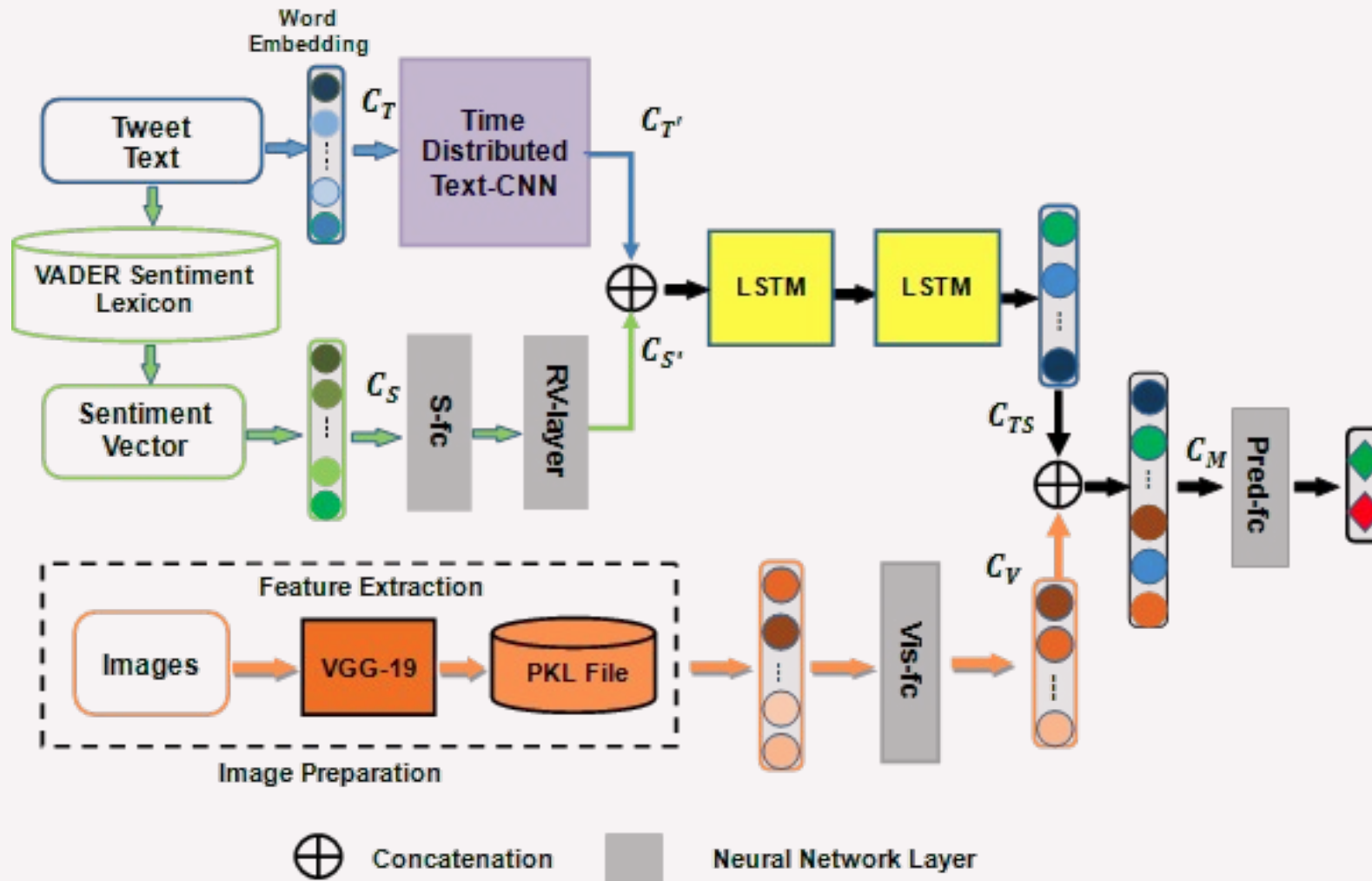


Village ibérique du puy de Sant Andreu d'Ullastret

deepMONITOR



Thèse
Abderrazek
Azri





Sondage express

Votre opinion sur ce cours ?

- A) Trop difficile
- B) Bof...
- C) Intéressant
- D) Génial !



Répondre sur <https://toreply.univ-lille.fr>

Question n° 451