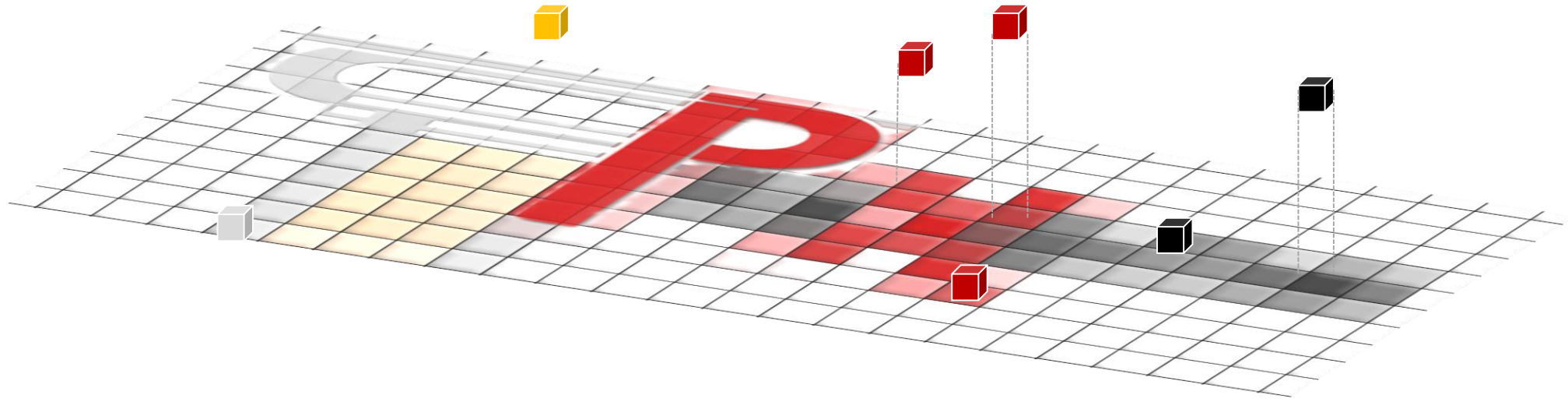




Big data, X-data... le phénomène Data a-t-il besoin d'une méthode ?

Symposium 2014, le 17 décembre 2014

Joël BIZINGRE, directeur-associé de 

Contenu de la présentation



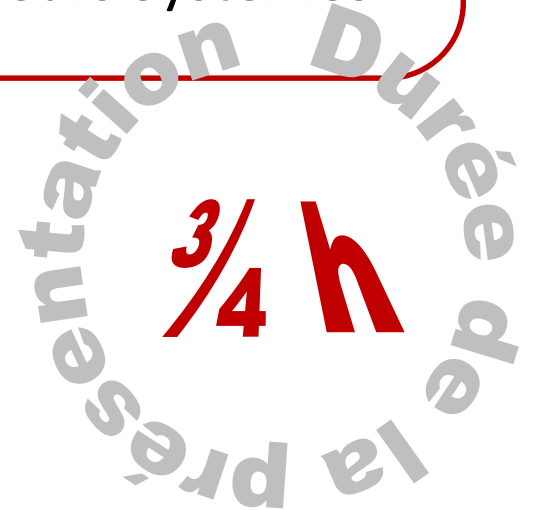
1. Le buzz Data

2. Data World

- Signaux et points de bascule autour des données
- Nouveaux métiers, nouvelles solutions, nouvelles guerres
- Nouveau contexte pour les entreprises ... et pour leurs Systèmes d'information

3. PxData

- Idée et ambition de la contribution
- Démarche
- Les 1^{er} travaux
 - Réflexions sur le Data World – Analyse des termes
 - La méthode Praxeme est nativement orientée données
 - Quels enrichissements de la méthode pour répondre aux attentes du Data World ?



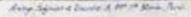
4. Echange avec la salle : Data world et méthode

Le buzz Data

- Data sourcing
- Open Data
- Data management
- Big Data
- Data as a service
- Data centric
- Data analysis
- Data journalism
- Data quality
- Data security
- Data exchange
- Data vizualisation
- Data storytelling
- Data policy
- Smart Data
- Data virtualization
- Master Data
- Data scraping
- DataLab

- Data governance
- Data engineering
- Data Lake
- Data provider
- Data Hub
- Data grid
- Unstructured Data
- Data API
- Data crawler
- Data broker
- Data marketplace
- Data cleansing
- Data science
- Data mining
- Data warehouse
- Metadata
- Linked Data
- Data portal
- Enterprise Data Architecture
- Dark data
- All Data ...





Le buzz Data



Carte interactive de France de la corruption



Data world : points de bascule

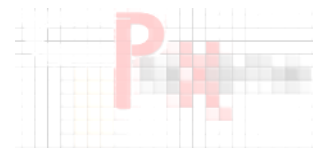


La data révolution

Neuer en libero contre l'Algérie, Bale médiocre défenseur reconverti en attaquant fulgurant. Comment les data transforment le foot.



Data world : points de bascule



**TREATO**
The voice of the patient

Log in | Sign up

See what **millions** of patients are saying

Search

Featured searches: [Lipitor](#), [Bupropion vs. Wellbutrin](#), [Prozac](#)

**Déc 2014**

 268,307,094 patients

 1,887,847,313 posts

 26,884 medications

 14,217 conditions

Plus de limite de portée

Data world : points de bascule

- 2000 : < 10 capteurs pour une voiture standard
- 2013 : 76 capteurs dernière Ford Hybrid
- Moins de 1\$ le cout d'une puce GPS
- 1 Milliards de smartphones vendus en 2013
- > 10 capteurs embarqués par smartphone
- 50 à 100 milliards d'objets connectés à l'horizon 2020



Plus de limite de capture

Data world : points de bascule

- 2000 : 25% de l'information est numérique, 75% analogique
- 2013 : 98% de l'information est numérique, 2% analogique
- Qu'est ce qui pourrait échapper à la numérisation ?

The smart bracelet that can read your emotions

7 July 2014 Last updated at 16:03 BST



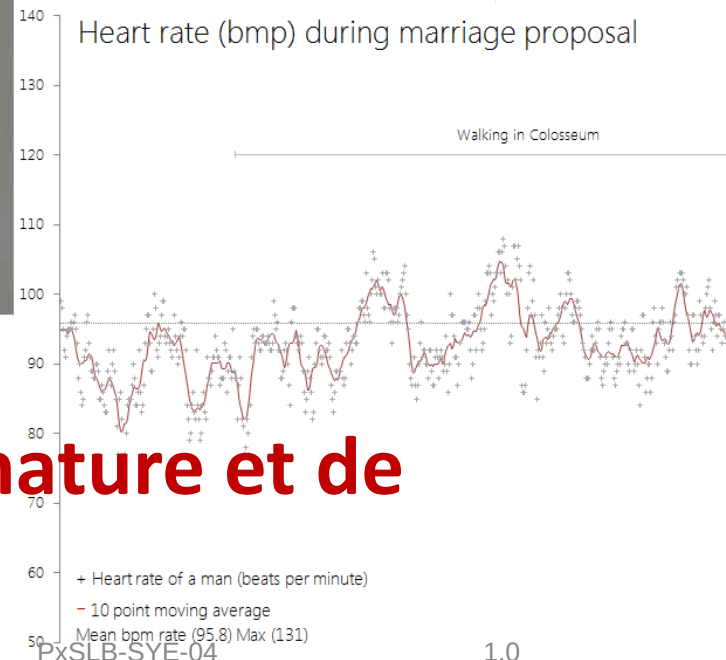
Plus de limite de nature et de type données

IDEES & DEBATS // FOCUS ESSAIS

La « datafication » du monde

Les Echos n° 21622 du 07 Février 2014 • page 10

L'avènement du « Big Data » va changer nos vies et transformer nos villes. Deux essais en anglais font le point sur les promesses et les menaces.



enterprise infrastructure
technology operations
information objectives
scorecards
analysis text mining
metrics applications
connections technical
solution stakeholder

Data world : points de bascule



Google Maps ajoute une machine à remonter le temps



LifeLogging sensor cameras, neck-worn by way of a lanyard to capture one's entire life



Mann, 1998

Microsoft, 2004

Mann, Fung, Lo, 2006

Memoto, 2013



Plus de limite de temps

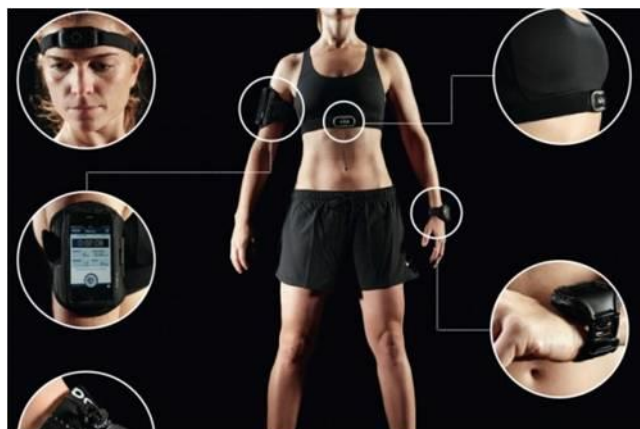
Lifelogging

(source wikipedia)



Data world : points de bascule

- Chaque objet et activité du monde physique peut avoir une ou plusieurs ombres numériques, c'est-à-dire des données et méta-données captées à un moment donné
- Aucune dimension (sociale, économique, artistique, politique, juridique, personnelle...) n'échappe à cela



<http://www.prescouter.com/2013/08/can-the-quantified-self-be-your-competitive-advantage/>



<http://www.numerama.com/magazine/28489-la-medecine-personnalisee-ou-la-sante-confiee-a-google-et-apple.html>

Un journaliste a utilisé un outil appelé **Collusion** pour déterminer qui le traçait : **105 entreprises** ont tracé son usage internet sur une seule période de 36 heures !

<http://www.framablog.org/index.php/post/2013/03/18/internet-etat-policier-surveillance>

et mandats électifs	fonction électifs	et mandats	ou gratifications perçues
Député	2007/2014		(5424,38 €/Mois) 609 € de 2008 à Mai 2010 1819,97 € (500 € jusqu'à Mars 2013 et 144 € depuis)
Maire	2008/2014		
1er Vice Président	2008/2014		
Conseiller d'Agglomération			
9° Les noms des collaborateurs parlementaires ainsi que les activités déclarées par eux :			
Noms des collaborateurs parlementaires	Identification de l'employeur ou de la structure sociale d'emploi	Description d'exercice de l'activité professionnelle	



http://www.lemonde.fr/les-decodeurs/article/2014/07/24/ce-que-devoient-les-declarations-de-patrimoine-des-parlementaires_4462427_4355770.html



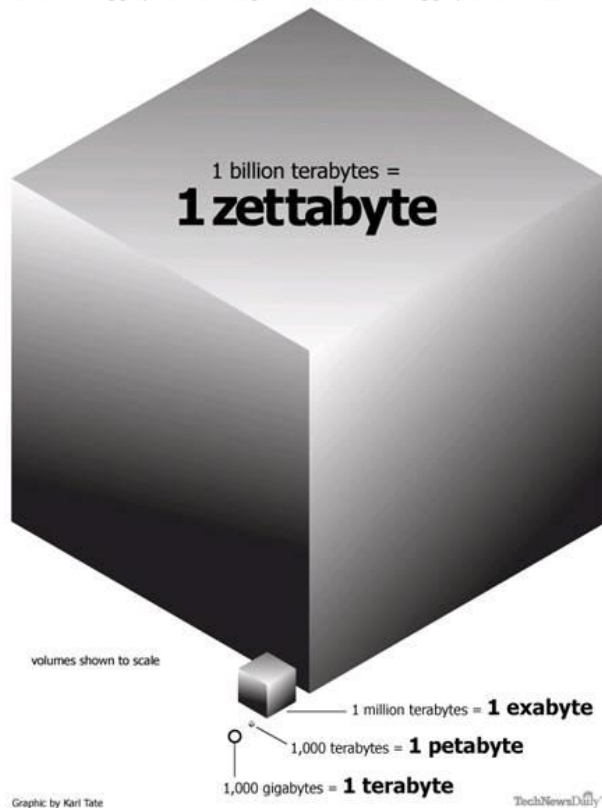
Plus de limites aux ombres numériques

Data world : points de bascule



Humanity Passes 1 Zettabyte Mark in 2010

A zettabyte is 1,000,000,000,000,000,000,000 bytes (that's 21 zeroes for those counting), or one trillion gigabytes. That's enough data to fill 75 billion 16-gigabyte-sized iPads.



Les échanges de données via Internet devraient dépasser le Zettabyte en 2016

**X 2 tous les 2 ans
pendant que le prix est / 2**

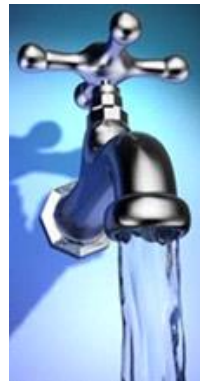


;-) RAM 4K, DD 64K
pour la mission Apollo

**Limites de stockage et de traitement (vélocité)
sans cesse repoussées**

Data world : points de bascule

- Il y a quelques années



Données



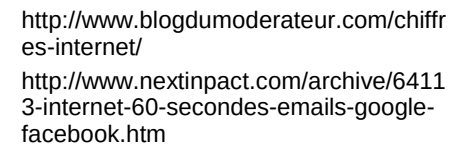
- **Aujourd'hui**

- **90% des données disponibles aujourd'hui ont été créées dans les 2 dernières années.**



Flot de données

4 millions de recherches Google sont effectuées
2,46 millions de contenus sont partagés sur Facebook
347 222 photos sont partagées sur WhatsApp
277 000 tweets sont envoyés
216 000 photos sont partagées sur Instagram
8 333 vidéos Vine sont partagées
3 472 images sont épinglées sur Pinterest
72 heures de vidéo sont téléchargées sur YouTube
204 millions de mails sont envoyés



Sans arrêt ... liquid Data



Recherche Google : internet 60 secondes

Data world : Nouveaux métiers

- Data manager
- Chief data officer
- Data analyst
- Data scientist
- Data miner
- Data strategist
- Data sourcer
- Data broker
- Data architect
- Data steward
- Data protection officer
- Data journalist
- Data coach
- Data artisan ;-)
- Data artist
- Et organisation : Datalab



Le big data créera 4,4 millions d'emplois IT dans le monde d'ici 2015



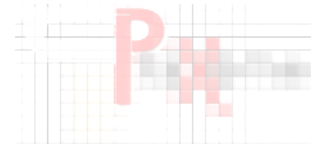
Gartner estime que plus de 4 millions de jobs IT liés au big data seront créés dans le monde d'ici trois ans, mais que les compétences dans ce domaine ne seront pas suffisantes pour répondre à la demande, faute d'un système éducatif inadapté.

Les dépenses mondiales en IT devraient dépasser 3,7 milliards de dollars en 2013, soit une hausse de 3,8% par rapport aux 3,6 milliards attendus d'ici la fin de l'année, estime le cabinet Gartner. Mais ce sont les prévisions concernant le big data qui suscitent beaucoup

d'enthousiasme. « D'ici 2015, 4,4 millions d'emplois informatiques seront créés dans le monde pour soutenir le big data, dont 1,9 millions aux Etats-Unis », a déclaré Peter Sondergaard, senior vice-président et responsable mondial de la recherche chez Gartner lors du Symposium/ITExpo qui s'est déroulé du 21 au 25 octobre à Orlando. « En outre,

LE PREMIER SALON DE RENCONTRE POUR LES METIERS DE LA DATA

Data world : Nouvelles solutions



- **Stockage**
 - Technologies Big Data : Hadoop, bases NOSQL, ...
 - Denedo technologies, composite, data gravity
- **Ateliers d'ingénierie Data (Data preparation)**
 - RelateIQ, OpenDataSoft, enigma.io, Trifacta, Paxata, ClearStory, Tamr, Reportive, Data brewery, Palantir
- **Data analytics**
 - Dataku, Tableau, Qlikview, semantria, visokio, spotfire, datameer, yellowfin, mirador, Google Bigquery, datawatch
- **Data integration**
 - Attivio, Bedrock Data, Google Dataflow, Amazon kinesis, Chartio, dat, vibe – informatica,
- **Data governance**
 - Colibra, adaptive, trillium
- **DataHub**
 - Antidot, OpenDataSoft, enigma.io, freebase, Data publica, Datasift, Dataprovider, zanran, ckan, SAMI, datafox, socrata, factual
- **Capture**
 - Import.io, Quandl, Dataminr, Kazoup, UIMA, Google fit, Apple Healthkit, sigsense, opensensors.io

Data world : Nouvelles guerres

- **Comment allez-vous ?**

Je vais



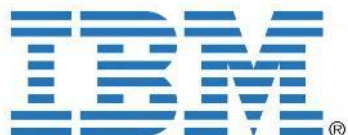
Data world : Nouvelles guerres

- Mais qui est le médecin ?



PHILIPS

SONY



CHU
Hôpitaux de
Bordeaux



Data world : Nouveau contexte pour les entreprises ... et pour leurs Systèmes d'information

AG du Cigref : la data, carburant de la révolution numérique



Edition du 21/10/2014 - par Bertrand Lemaire

Le club informatique des grandes entreprises françaises (Cigref) a tenu son Assemblée Générale le 20 octobre 2014. Ce fut l'occasion de rappeler le rôle essentiel du DSI dans la révolution numérique des entreprises.

« Depuis 2010, le Cigref affirme que le numérique est facteur de croissance pour les entreprises mais ce qui est aujourd'hui banal ne l'était pas à l'époque » a rappelé Pascal Buffard, président réélu du Cigref [en photo], lors de l'Assemblée Générale de cette association, le 20 octobre 2014. Jean-François Pépin, délégué général, s'est réjoui pour sa part que « le Cigref a été le premier à parler d'entreprise numérique ». A l'époque, un travail de fond avait commencé sur la vision de l'entreprise en 2015. Comme ce temps es

<http://www.cio-online.com/actualites/lire-ag-du-cigref%C2%A0-la-data-carburant-de-la-revolution-numerique-7156.html>

- Représentation numérique
- Dématérialisation
- Enrichissement par les données, l'analytique
- Création de nouveaux services
- Interaction et asservissement



Data et transformation numérique sont étroitement liés

Data world : Nouveau contexte pour les entreprises ... et pour leurs Systèmes d'information



Enrichissement



Analytique



Paiement à l'usage



Produits en services



Asservissement



Data et transformation numérique sont étroitement liés

Data world : Nouveau contexte pour les entreprises ... et pour leurs Systèmes d'information



La data révolution

Neuer en libero contre l'Algérie, Bale médiocre défenseur reconverti en attaquant fulgurant... Les stats l'avaient prévu. Comment les data transforment le foot.

Tous footballeurs, cyclistes connectés ?



Sociometric badges

Data et transformation numérique sont étroitement liés

Contenu de la présentation



1. Le buzz Data

2. Data World

- Signaux et points de bascule autour des données
- Nouveaux métiers, nouvelles solutions, nouvelles guerres
- Nouveau contexte pour les entreprises ... et pour leurs Systèmes d'information

3. PxData

- Idée et ambition de la contribution
- Démarche
- Les 1^{er} travaux
 - Réflexions sur le Data World – Analyse des termes
 - La méthode Praxeme est nativement orientée données
 - Quels enrichissements de la méthode pour répondre aux attentes du Data World ?

4. Echange avec la salle : Data world et méthode

PxData : idée et ambition de la contribution



Création 1997



2012 : 19,5 M€
DE CHIFFRE D'AFFAIRE



Membre du Praxeme Institute
depuis sa création



180 Collaborateurs

Banque & Finance	50%
Energie & Industrie	24%
Télécom	13%
Services	13%

4 Business Unit



Risque
& Réglementaire

Intégration des SI



Transformation
& innovation

Sécurité

Consulting :

- Architectes SI, IT, Cloud
- Urbanistes
- Architectes de données
- Experts : innovation, transformation

PxData : idée et ambition de la contribution



• Autour de la donnée



Stratégie Numérique

- Veille et études personnalisées
- Services Innovants
- Nouveaux usages



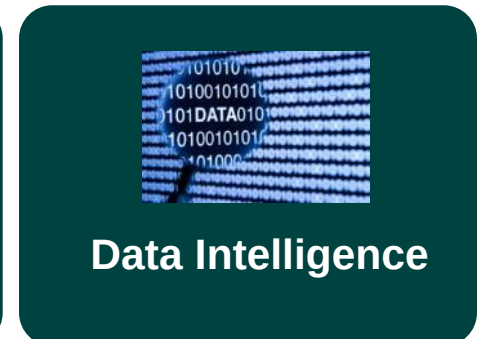
Nouveaux Métiers

- Sensibilisation et éclairage
Data World (Big data, Open data, data analyst, CDO,...)
- Accompagnement des Métiers et DSI



Gouvernance des Données

- Dispositifs de Gouvernance
- Référentiels
- Valorisation des Données
- Protection des Données



Data Intelligence

- Mise en œuvre des nouvelles architectures B.I. / Data Lake
- Etudes, architectures, AMOA
- Cadrage de projets innovants

• Evolutions des Systèmes d'Information

- De la muraille de chine à l'ouverture
- Ingénierie Data transverse : référentiels, échanges, gouvernance...
- La capture des données fait l'objet d'une stratégie explicite
- Du traitement centric au data centric

P

Exploiter



Impuissance ou Influence ?

Que peut apporter une méthode ... et Praxeme ?

 **Assister**

Le big data créera 4,4 millions d'emplois IT dans le monde d'ici 2015

Gartner estime que plus de 4 millions de jobs IT liés au big data seront créés dans le monde d'ici trois ans, mais que les



LE PREMIER SALON DE RENCONTRE POUR LES METIERS DE LA DATA

Innovover



Transformer

Entreprise Système d'Information

PxData : idée et ambition de la contribution



- La contribution à la méthode est centrée sur :
 - « **Se transformer pour tirer profit des opportunités apportées par les X-Data** »

- Objectif du chantier PxData
 - Mise en lumière des forces de Praxeme pour leur exploitation par les projets X-Data
 - Développement des procédés qui permettent aux entreprises de tirer profit des opportunités apportées par les X-Data

- Historique
 - Lancement en juin 2014

PxData : idée et ambition de la contribution



■ Les forces de Praxeme

- 1) La dimension Transformation d'entreprise
 - Data et transformation numérique (d'entreprise) sont étroitement liés



■ Les forces de Praxeme

■ 2) Le retour en grâce de la modélisation ... !

- Communication : est-ce de la même chose dont-on parle ?
- Ouverture : êtes-vous sûr que les données que vous utilisez correspondent au modèle (conclusions) que vous publiez ?
- Transparence : comment mes données sont exploitées – c'est quoi le modèle ?
- Interprétation : d'abord quel est le bon modèle pour exploiter correctement toutes ces données ... parole de Data Scientist
- Intégration : et si on s'échangeait nos données mais suivant quel modèle ? Les données que l'on vient de récupérer à quoi correspondent-elles ?
- Décision et action : sur quel modèle est bâti cet indicateur ? Comment peut-on rapprocher ces données ?
- Innovation : quel modèle inventer à partir de ces données ?

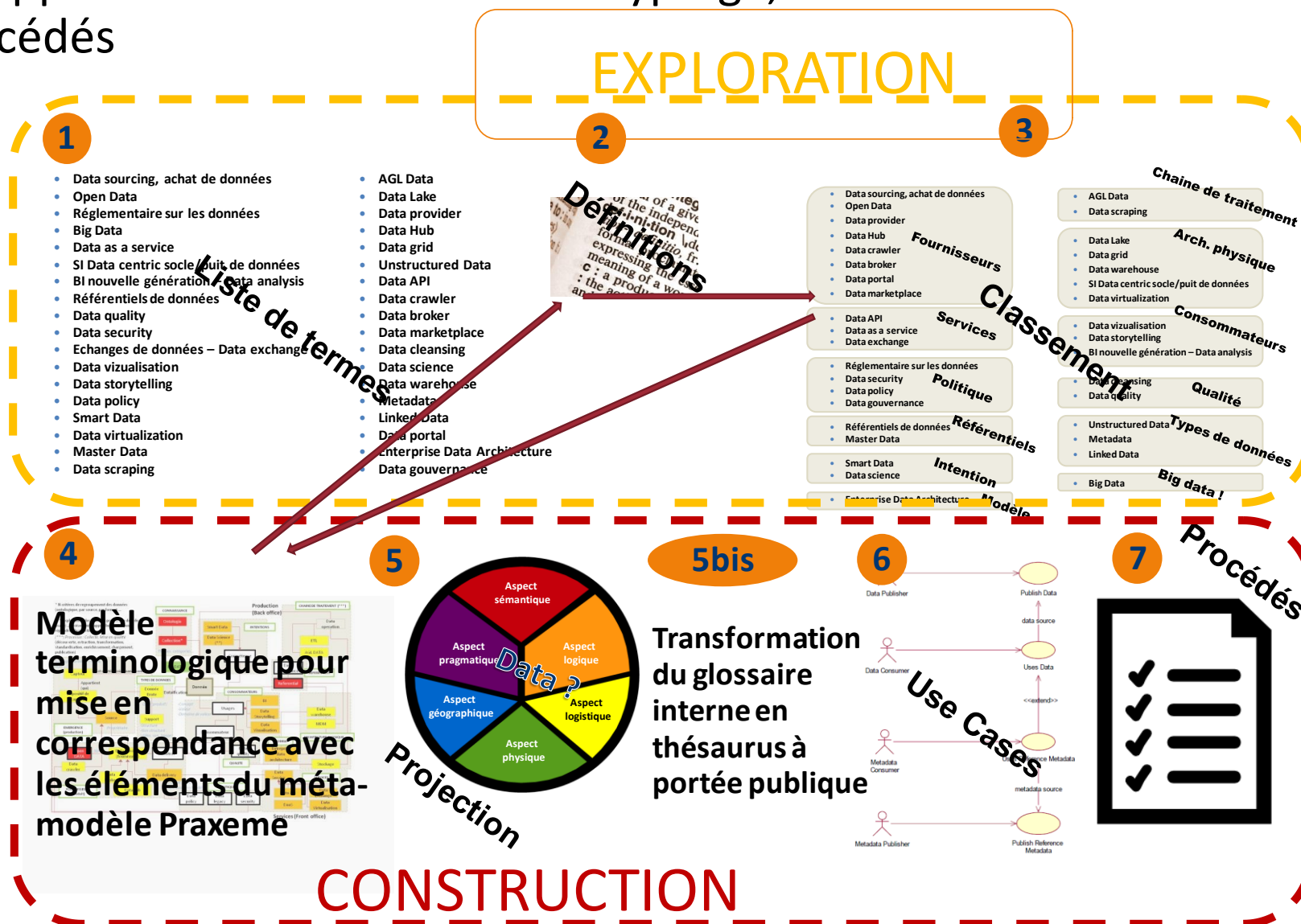


terminologie, ontologie, sémantique ...

PxData : idée et ambition de la contribution

■ Les forces de Praxeme – le chantier PxData

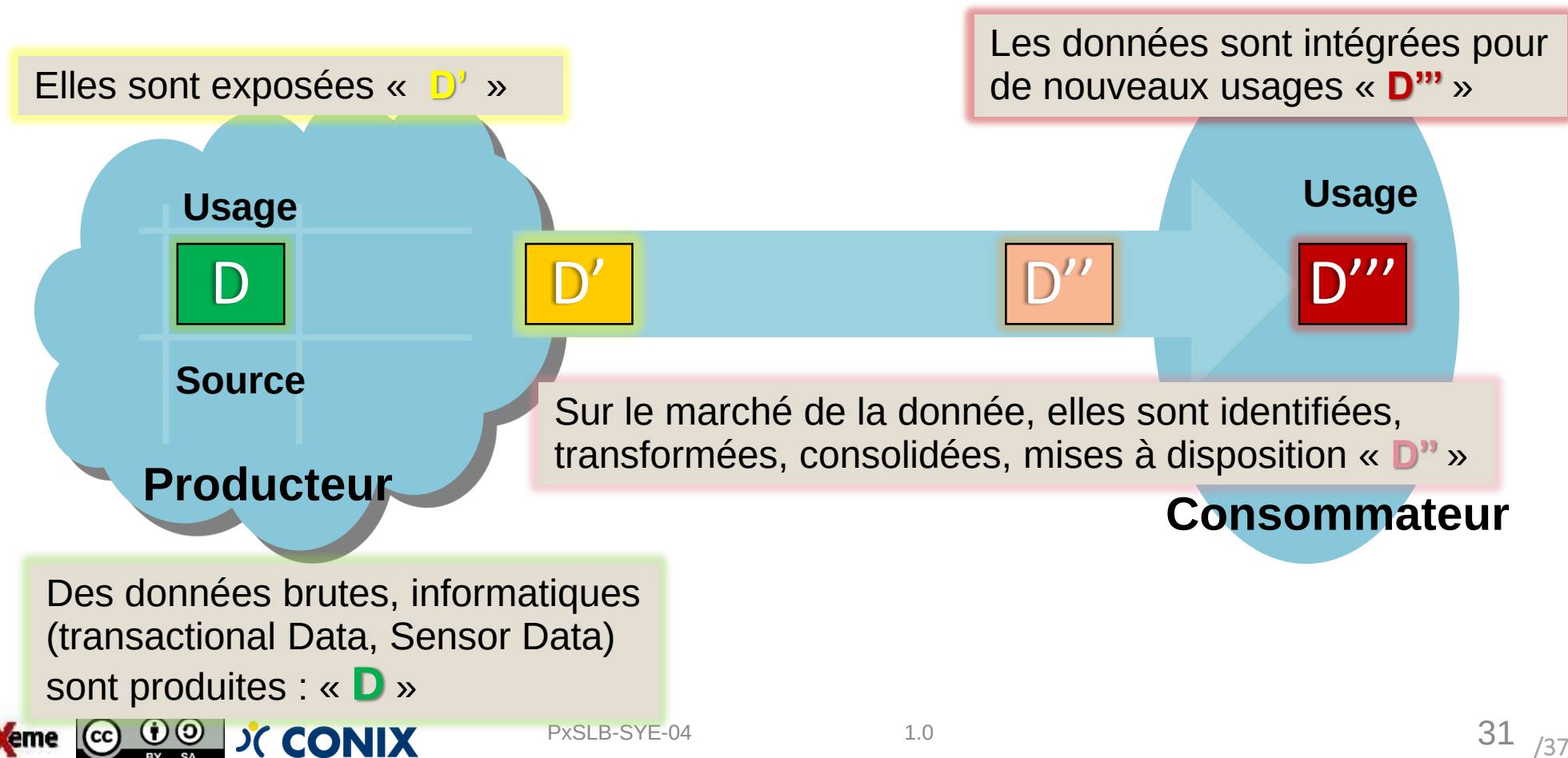
- 3) Opportunités des X-Data : Décryptage, Formalisation – modélisation, Procédés



PxData : les 1^{er} travaux – analyse du Data World

■ Les forces de Praxeme – le chantier PxData

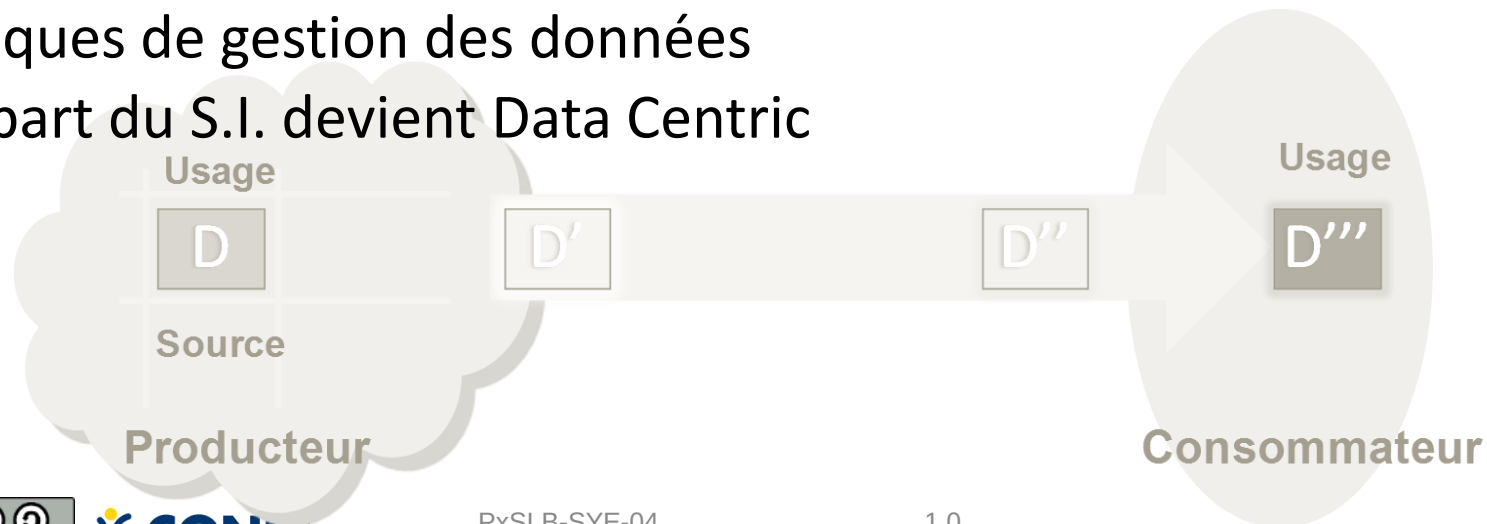
- 3) Opportunités des X-Data
- Analyse des termes : Data Provider, Data Broker, Data Producer, DataHub, OpenData, Data sourcing, Data crawling, Data sourcing, Data Marketplace, Data portal, Data collector, DaaS, Data aggregators, Data Management Platform
- Emergence de l'idée du modèle de la supply chain de la donnée



PxData : les 1^{er} travaux – analyse du Data WorldPx

■ Les forces de Praxeme – le chantier PxData

- 3) Opportunités des X-Data
- Modèle de transformation : intégration d'une supply chain de la donnée
- Avec :
 - Une source et un processus d'acquisition (rôle de producteur)
 - Un mode de distribution (rôle de fournisseur)
 - Un mode de mise à disposition (rôle de datamédiaire)
 - Un mode de consommation (rôle d'acquéreur)
- La donnée comme ressource autonome, dynamique, changeante et mobile
- Qui profite à une consommation agile découplée des dispositifs classiques de gestion des données
- Une part du S.I. devient Data Centric

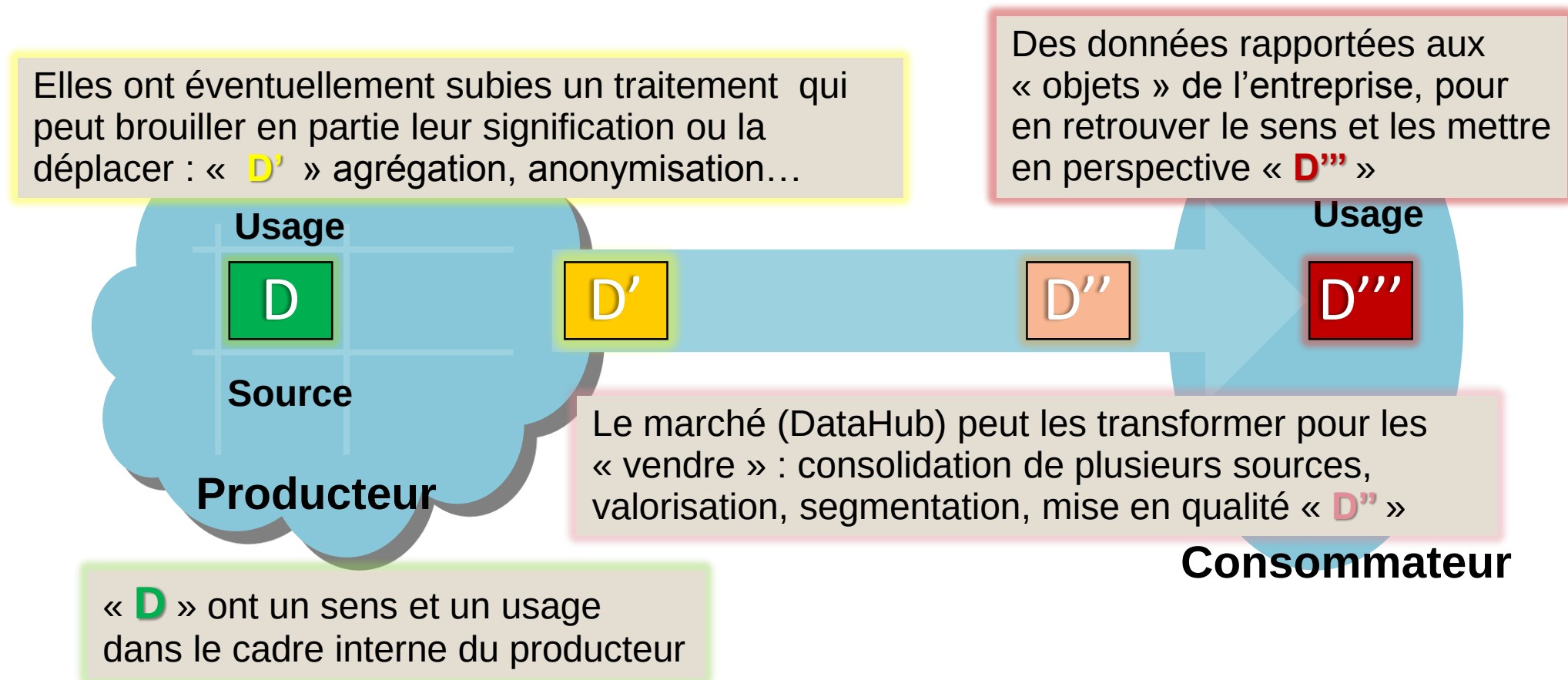


PxData : les 1^{er} travaux – analyse du Data World

■ Les forces de Praxeme – le chantier PxData

■ 3) Opportunités des X-Data

- Un point clé : un nécessaire processus de sémantisation (le centre de l'offre Praxeme) pour maîtriser les dérives sémantiques

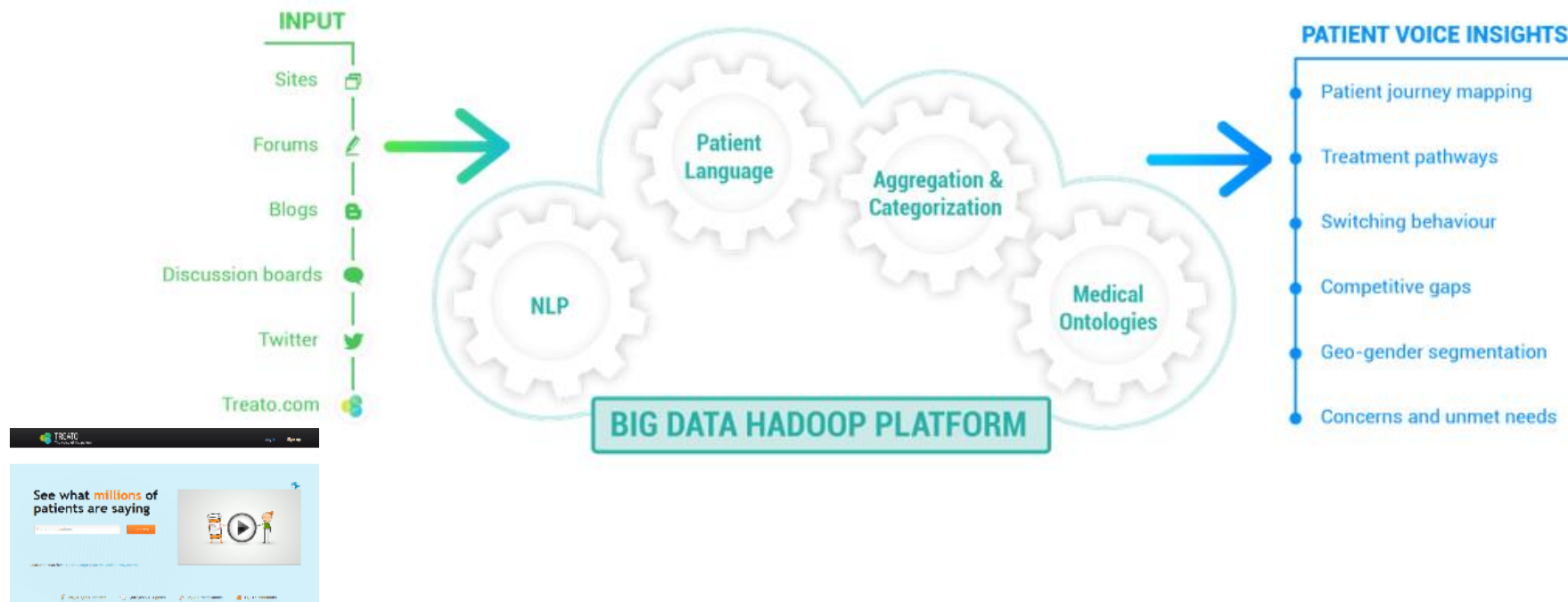


PxData : les 1^{er} travaux – analyse du Data World

■ Les forces de Praxeme – le chantier PxData

■ 3) Opportunités des X-Data

- Un point clé : un nécessaire processus de sémantisation (le centre de l'offre Praxeme) pour maîtriser les dérives sémantiques



Understanding patient language

Healthcare is one of the most complex and sensitive fields discussed online. Patients have their

Comprehensive medical ontology

Treato draws from extensive medical dictionaries and patient language

Making big data seem small

Treato uses a state-of-the-art, open-source Hadoop distributed computational platform to provide

<http://corp.treato.com/technology.html>

PxData : les 1^{er} travaux – analyse du Data World

■ Quels procédés utiles ? Pistes pour 2015

- Compléments à apporter aux procédés existants ou identifiés
 - Cf. le catalogue
 - Extension et universalisation du modèle sémantique
 - Intégration de données « externes » - de la supply chain
- Procédé d'élaboration d'une Data policy
 - Formulaire du dossier « Data policy » (Politique de l'information)
 - Décisions à prendre (options à envisager par domaine : acquisition (ponctuelle ou continue), valorisation (monétisation...), partenariat, pattern de la Data supply Chain, encadrement par le juridique, valeurs à respecter, infrastructure, impact sur les processus, sur la conception...)
 - Avec son mode d'emploi
- Procédé d'évaluation des offres de fournisseurs de données
- Procédé d'élaboration d'une Data governance
- Procédé modèle métrologique (lien ADP) – conception des indicateurs
- ...
 - Au service des nouveaux métiers (exemple Chief Data Officer, Data Scientist...)
 - Au service des transformations : rôle des données

Conclusion : PxData la synthèse



■ Rappel de l'objectif

- Big data, X-data... le phénomène Data a-t-il besoin d'une méthode ?

À retenir

- Plus que du buzz, plus qu'un effet de mode, on a basculé dans le Data World
- Par nature c'est un facteur incontournable de transformation des entreprises et de leur Système d'Information
- L'approche Data se voit renforcée dans les entreprises et devient centrale
- Qui dit transformation dit intérêt d'une méthode
- Praxeme est nativement orientée Data et couvre déjà une large part des enjeux de transformation
- PxData a vocation à étendre la couverture de Praxeme par :
 - La prise en compte de nouveaux modèles de transformation centrés données : idée de supply chain de la donnée
 - La définition de procédés à destination des acteurs de transformation : chief data officer, data architect, data scientist, data manager...

Etes vous confronté au Data world ?

Quelles difficultés rencontrées ?

Le Data world est-il efficacement exploité?

Qu'attendriez vous d'une méthode ?

Quelles seraient les priorités ?