

**Séminaire de connaissance réciproque**  
**Enjeux du numérique – Déluge de données**

**Atelier n°1/2**  
**Les données**

**Les 01 et 02 février 2018 – ENSG**

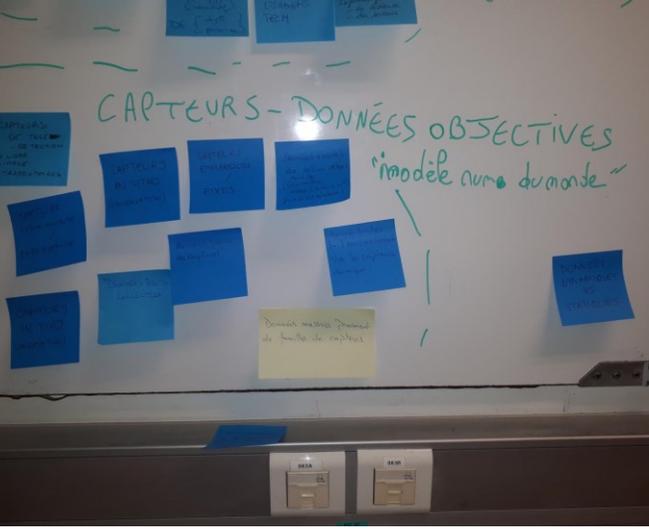
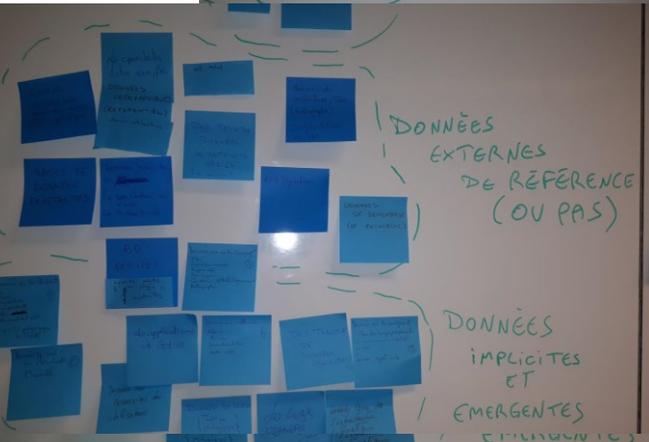
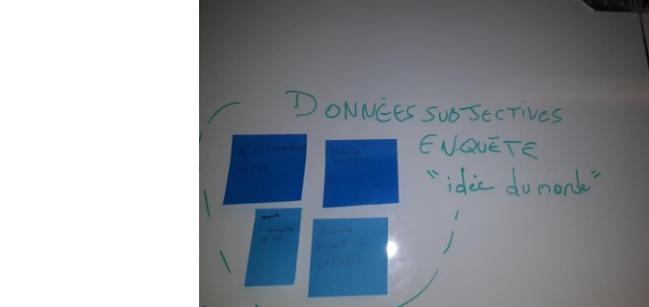
# Focus : Les données



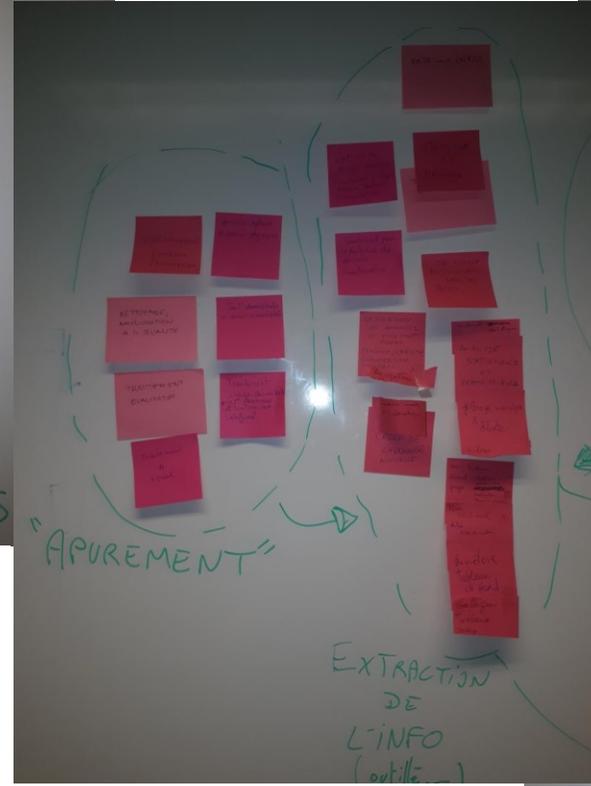
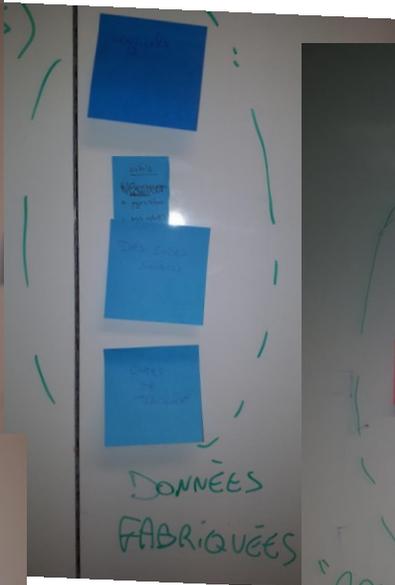
(L'atelier 2 à l'ESIEE « numérique - organisation, process, outils »)

# Travail collectif : ENSG (5), ESIEE (1), Ifsttar (11), UPEM (3)

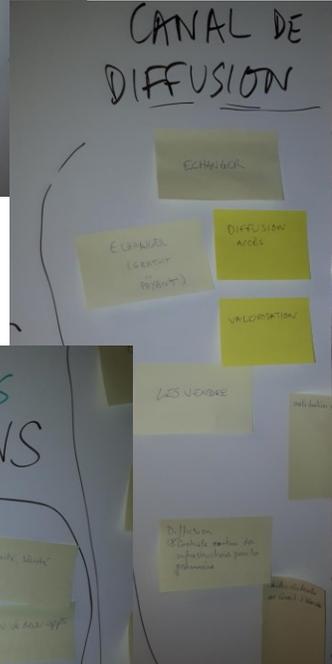
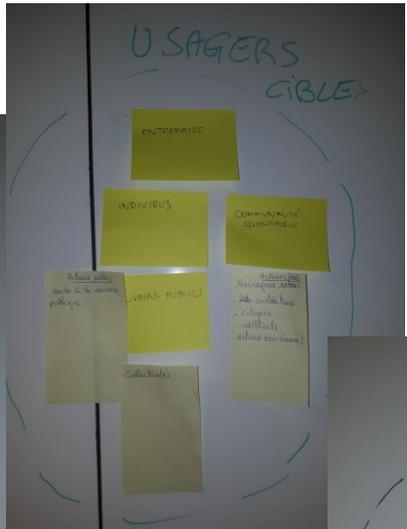
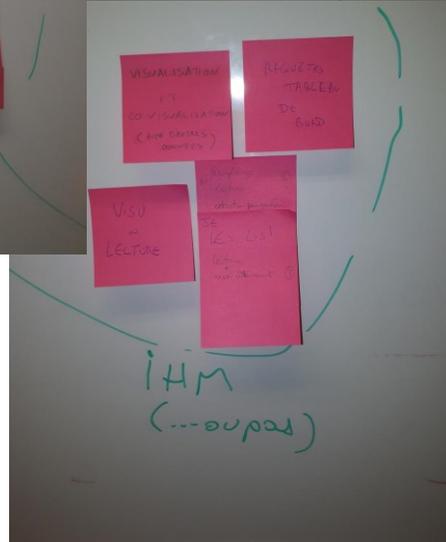
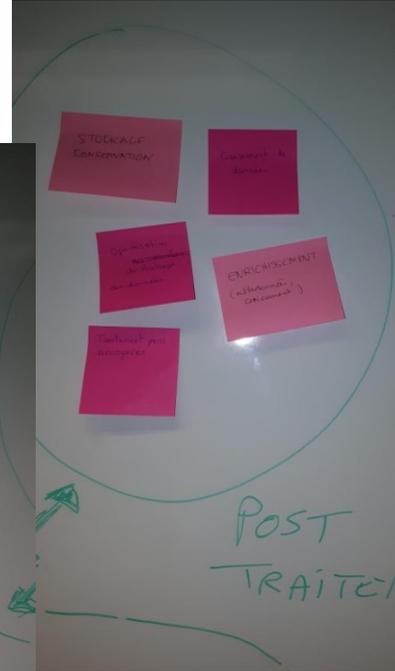
<https://mycore.cnrs.fr/index.php/s/X0qvYkhxSHvPTEC>



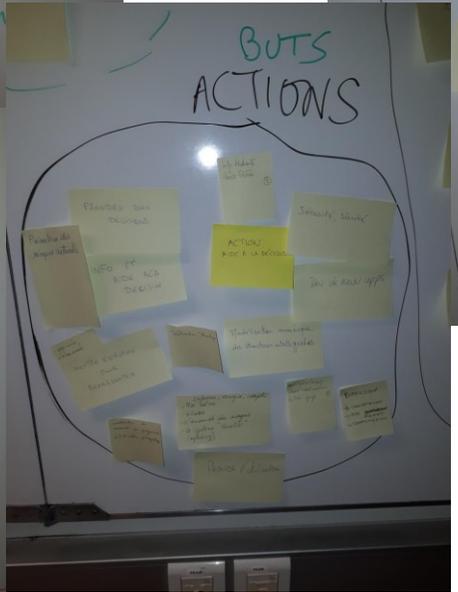
SOURCES



TRAITEMENTS



USAGES



# Brainstorming – Source, traitements, finalités

## Sources

- Données **subjectives**, données d'enquête - l'idée qu'on se fait du monde
- Données **objectives**, données capteurs - le modèle numérique du monde
- Données **fabriquées**, données issues de codes sources - une création numérique (idéalement) conforme au monde
- Les **codes sources** sont aussi des données
- Données **externes**, y compris les **données de référence**
- Données **implicites et émergentes** (données de gestion périphériques)

## Traitements

- **Nettoyage** ou « apurement » des données : objectif qualité
- **Extraction de l'information** (avec des outils... ou avec les yeux!)
- **Post-traitements** (avec possible retour à une étape d'extraction de l'information)

- Usagers cibles

- Actions - Décisions

- Canaux de diffusion

\* Ici, les contributions étaient **tellement variées** qu'il a fallu **beaucoup agréger** !

## Destinations, usages

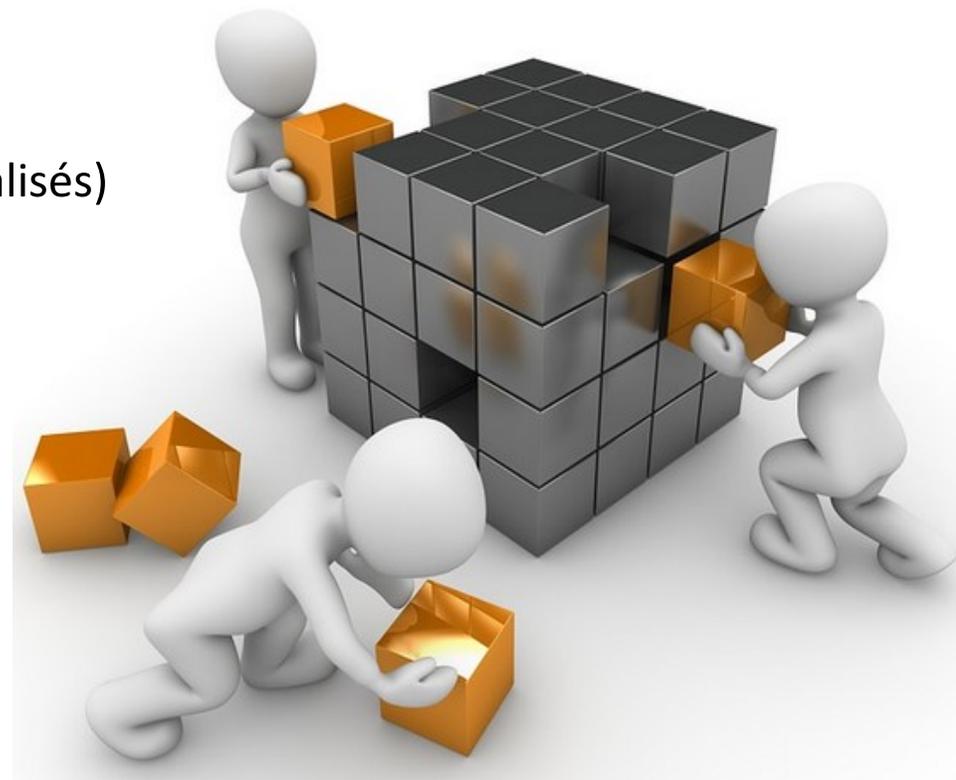
# Ce qui nous satisfait avec nos données

- **Des outils fonctionnels**
  - Partage
  - Numérisation / capacitation à a collecte
  - Moyens fournis : stockage
- **Des méthodes appropriées**
  - Formalisation commune
  - Volonté des chercheurs à mieux gérer les données
- **Une implication institutionnelle**
  - Politique
  - Communication interne
  - Portage du sujet « données » via un travail pluridisciplinaire



# Ce dont on aurait besoin pour nos données

- **Des outils à améliorer**
  - Infra de stockage et d'accès
  - Outils de partage pour gros volume, sécurisé, souverain
  - Archivage, sécurisation à long terme
- **Des méthodes à généraliser**
  - Guides de bonnes pratiques (y compris le code)
  - Normalisation, standardisation (formats communs et standards)
  - Transparence (sur les données existantes et sur les traitements réalisés)
  - Généraliser l'usage de plan de gestion de données (formaliser)
  - Exploitation plus systématique des données (des étudiants...)
  - Temporalité et gestion temporelle des versions des données
  - Partage de données, de code
- **Une prise de conscience institutionnelle à développer**
  - Du temps !!! Des sous!!! Des gens!!!
  - Politique de gestion des données
  - Compétence « data » d'un comité d'éthique
  - Manque de connaissances, de compétences
  - Prendre en compte la variété de la réglementation à l'échelle internationale
  - Accompagner la diffusion



# Ce qui nous bloque

- **Des outils manquants**
  - Intégration des outils au SI
  - Peu de doc décrivant les SI
  - Archivage intégré au SI
  - Espace de stockage
  - Alternatives aux outils Dropbox, Google et autre
- **Des méthodes à promouvoir**
  - Qualifier la qualité des données
  - Formats ouverts au lieu de propriétaires
  - Informer pour Dropbox, Google et autres risques
  - Efforts pour respecter la législation contraignante pour protéger les droits des personnes
  - Pas d'anticipation sur la gestion des données, pas de plan de gestion
- **Des verrous institutionnels**
  - Connecteurs entre silo (inter et intra établissement(s))
  - Moyens
  - Formation tout au long de la carrière : Augmentation des compétences techno
  - Besoin de conseil, d'accompagnement
  - Formation pour les chercheurs sur leurs droits et obligations
  - Résistance au changement et aux procédures



# Propositions d'actions phares pour une stratégie numérique UCIBLE



- Diagnostiquer / formaliser le traitement des données
  - Etat des lieux / cartographie des données / Classification
  - Etat de l'art et parangonnage
- Construire une stratégie numérique
  - Définir une politique numérique commune (alignée avec les autres stratégies : valo...)
  - Identifier des chantiers numériques (par exemple, learning analytics, suivi des alumni...)
  - Mettre en œuvre des méthodes innovantes (par exemple, machine learning...)
- Renforcer la politique de responsabilité sur les données
  - Comité d'éthique UCible compétent sur les données
  - Connaissances juridiques (données personnelles, PI...)
- Outiller
  - Offre interne, ouverte vers l'extérieur
  - Pour le partage, pour le chiffrement, pour la pseudonymisation, archivage...
- Former, informer, accompagner : action en continu
- Mesurer et auditer en continu la mise en œuvre de la stratégie

→ Aboutissement : un schéma directeur du numérique

Nous vous remercions pour  
votre participation et vos contributions

Prochaines rencontres « Enjeu du numérique – Déluge des données »

**Numérique – organisation, process, outils**

- Atelier 2 à l'ESIEE Marne-la-Vallée, les 05 et 06/04/2018

# Conception de l'atelier

- Arnaud Bonnard – Ifsttar / DS
- Pascal Romon - UPEM

## Illustrations sous licence CC

- Ville binaire <https://www.flickr.com/photos/opendemocracy/1891638026>
- Travail en groupe <https://pixabay.com/fr/salle-de-classe-1297779/>
- Cube de données <https://pixabay.com/fr/matrice-r%C3%A9seau-%C3%A9change-de-donn%C3%A9es-1027571/>
- Mal de tête <https://pixabay.com/fr/maux-de-t%C3%A4te-mal-de-t%C3%A4te-la-douleur-1910644/>