



N° C3

Les données numériques : quels enjeux ?

26^{èmes} Rencontres du Risk Management | AMRAE 2018



Denis Coutrot

Data Science VP at Transdev

denis.coutrot@transdev.com

PARTIE 1

QU'EST-CE QUE LA DATA SCIENCE ?

QUE FONT LES DATA SCIENTISTS ?

Data science?



Data science?



[Big] Data science ?



An aerial photograph of a city street during the day. The street is filled with various vehicles, including several yellow taxis, white sedans, and a large white bus. A bus stop with a black canopy is visible on the sidewalk. People are walking along the sidewalks and crossing the street. The street has white dashed lines for traffic flow.

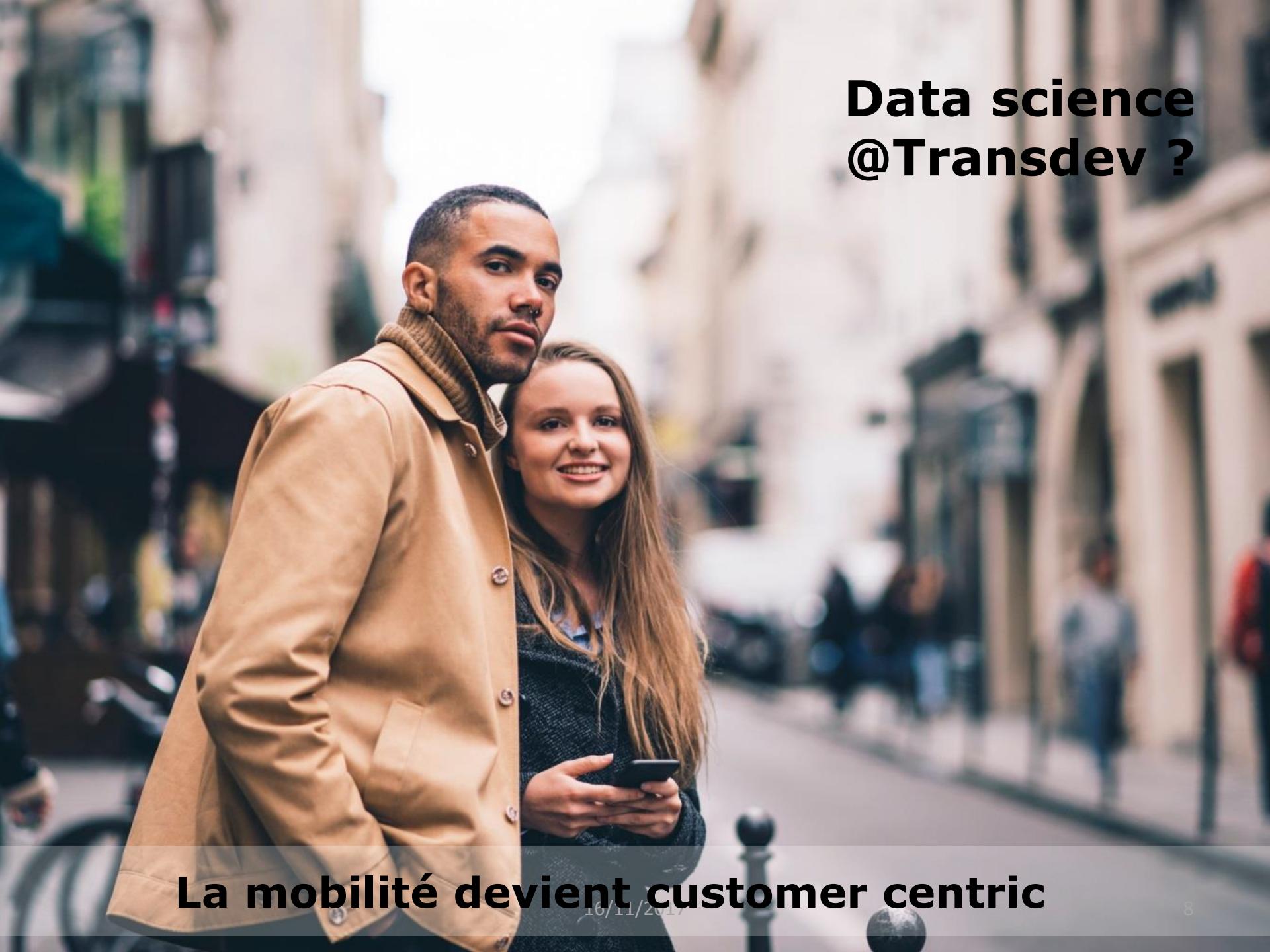
**Data science
@Transdev?**

La gestion de la mobilité est un challenge

A close-up, low-angle shot of the front of several red and yellow transit vehicles, likely buses or trams, parked in a row. The focus is on the headlight and front grille of the nearest vehicle.

**Data science
@Transdev?**

Des milliards de données chaque jour

A photograph of a man and a woman standing on a city street. The man, on the left, is wearing a tan trench coat over a dark turtleneck sweater and has his arm around the woman. The woman, on the right, is wearing a dark blue textured jacket over a light-colored collared shirt and is smiling while looking at a black smartphone she is holding in her hands. They are standing on a sidewalk next to a road with blurred buildings in the background.

Data science @Transdev ?

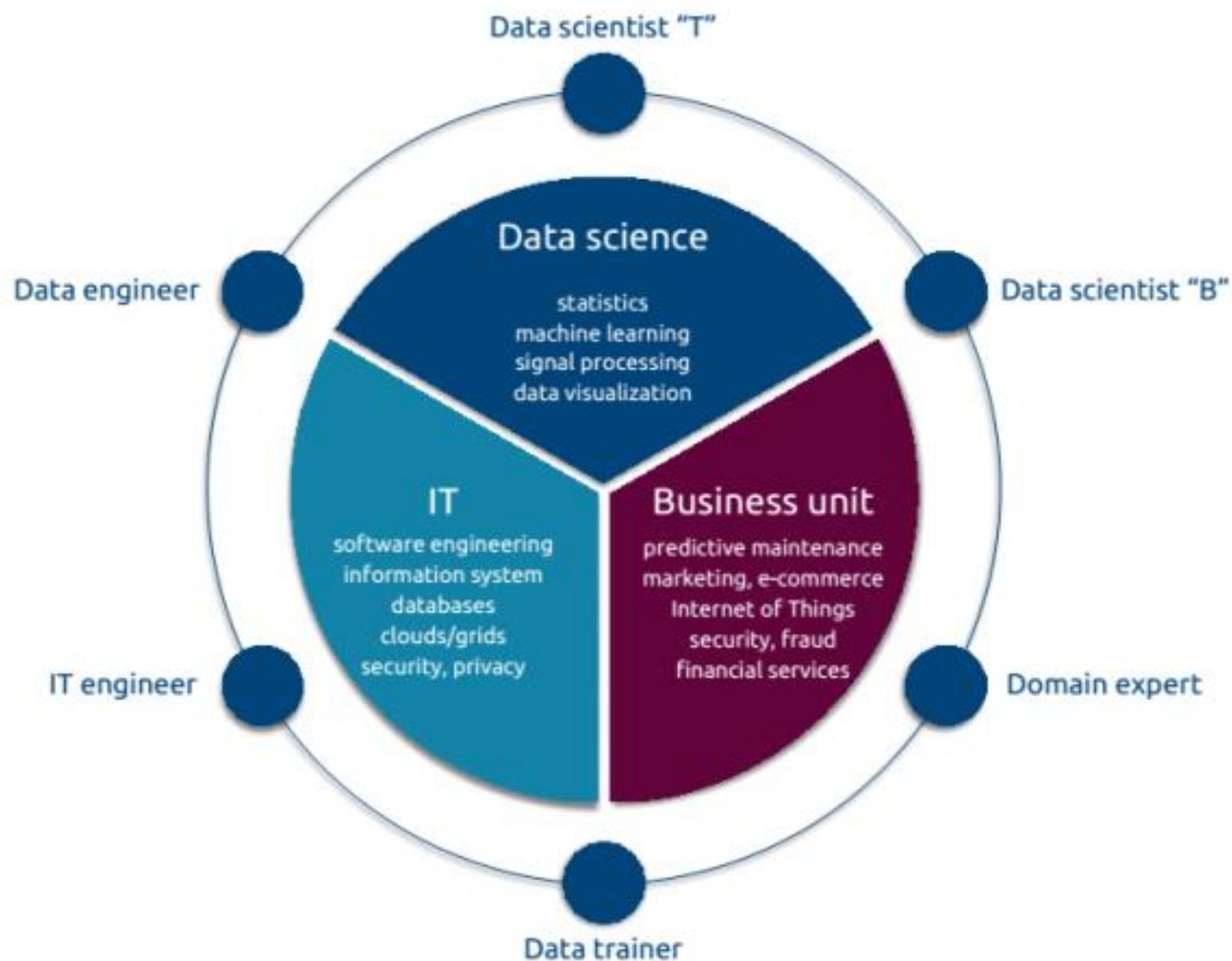
La mobilité devient customer centric

16/11/2017

La data: une transversale de l'entreprise



THE DATA SCIENCE ECOSYSTEM



A photograph showing three people near a black limousine. A man in a dark suit and sunglasses is opening the rear passenger door. A woman in a light-colored dress is smiling and about to enter the car. Another man in a blue suit and glasses stands behind her, also smiling. They appear to be in a professional or celebratory setting, possibly at an airport or hotel. The background shows a modern building with large windows.

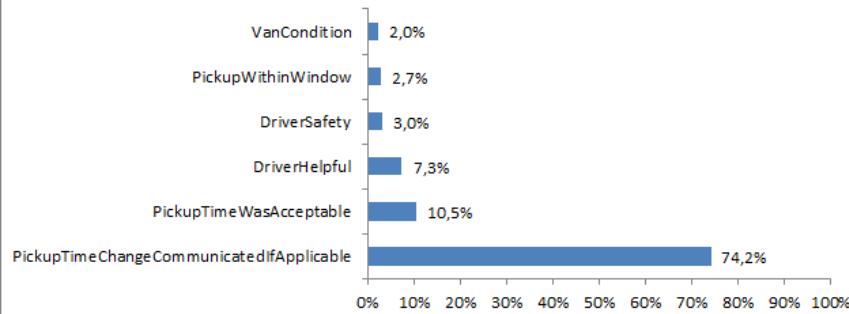
Cas: satisfaction client
SuperShuttle USA

Comprendre les attentes de nos clients

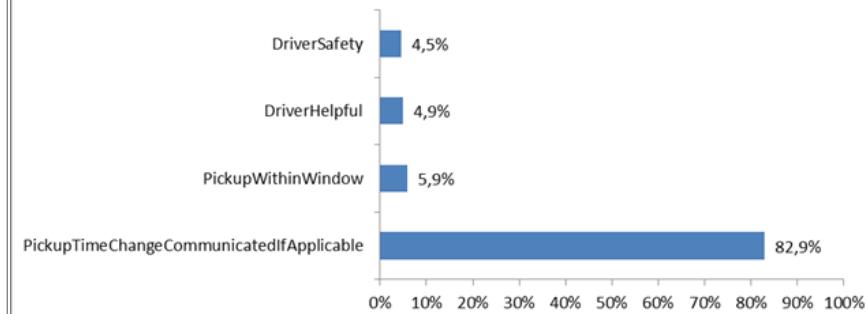
- Des niveaux de satisfaction très élevés / une concurrence féroce
- 1 272 847 questionnaires de satisfaction pour comprendre les + et les - du service
- Aller dans les détails grâce au Machine learning / Random forest
 - ... et 48h de run



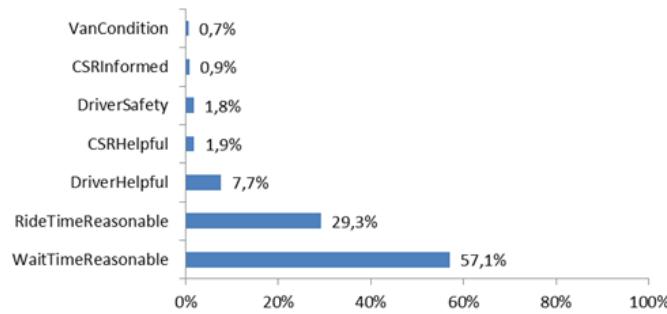
Features importance - Blue Van Inbound



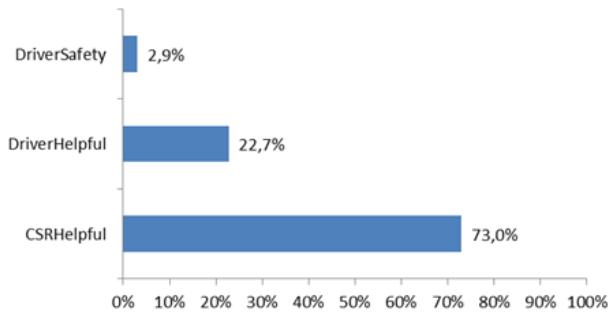
Features importance - Execucar Inbound



Features importance - Blue Van Outbound



Features importance - Execucar Outbound



Résultats

Organisation et méthodes



BI (Business Intelligence)

1. Business process & Reporting
2. Data collection to feed & measure a process and improve performance

Data Science

A business question looking for an answer through data exploration

Data / Big data

Few Data highly structured with high cost of production
ex: Financial data, tasks, mgt model

Large volume of measured data structured at low cost
ex: Km, liters, temperature

Unstructured data
ex: Text, Facebook comments, real time logs



Checken forecast

Alert – Control the business

Forecast en modellen

Understanding the business

Analyse tools en rapporten

Knowing the business

KPI-tools

Improving the business

Datamart

Definities, logica en aggregaties

Datavault

Brondaten en historie

Ethique et risques

Des données invasives

Accountability

(Not so) Open Data ?

Cyber sécurité



Paul Sterckx

AIG

Directeur des Risques Financiers et Responsabilités

PARTIE 2

DATA & ASSURANCE

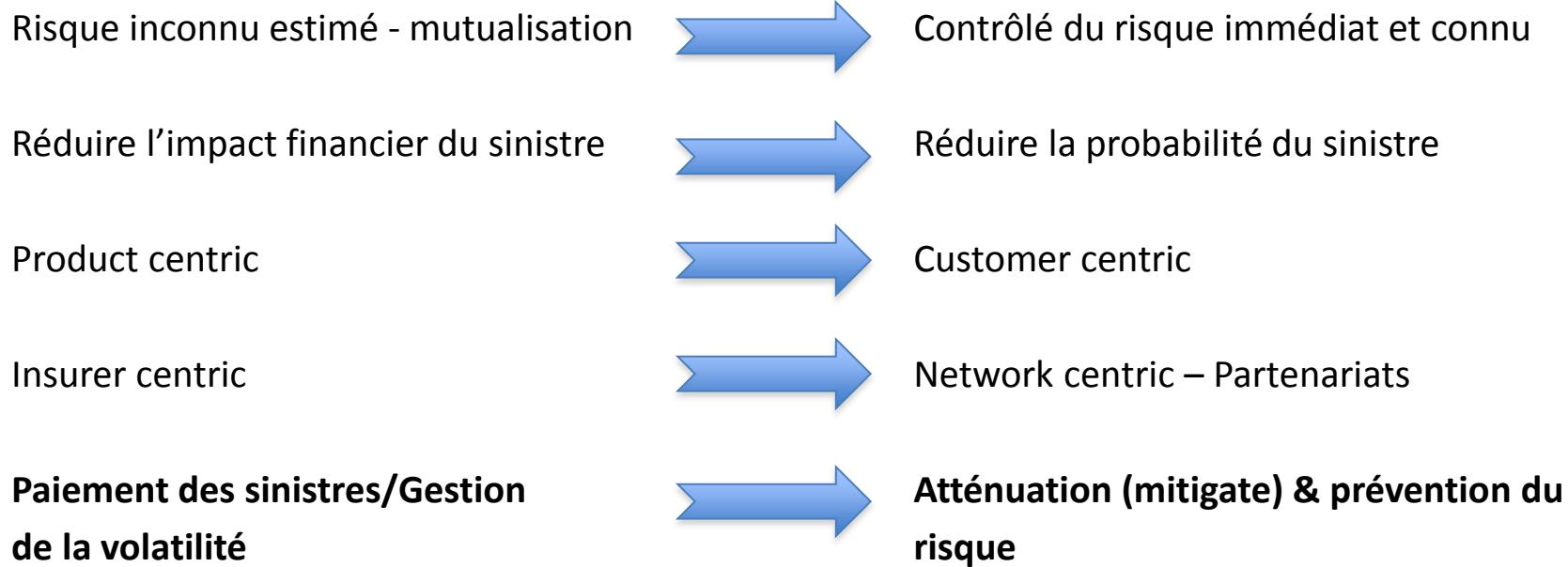
Internet des Objets – Faciliter la Vie



Internet des objets



« The Journey »



Les challenges

a) Le mariage entre big data et assurance:



Source: Utilisation de nouvelles sources de données par secteur: E&Y, IoT in Insurance

b) La disruption par de nouveaux entrants:

1. **Assurtech:**
 - i) 43% des compagnies intègrent dans leur stratégie les assurtechs MAIS,
 - ii) seulement 28% ont investi dans des partenariats avec ces startups;

2. Les géants du numérique → nouveaux concurrents?:
 - i) Amazon Protect → garantie de dommage accidentel et vol des produits achetés disponible en ligne;
 - ii) Contrats GAFA? > les géants du digital & assurance.

Les challenges (cont.)

c) Gouvernance des données: plusieurs enjeux

- i) la qualité des données utilisées - l'interprétation des données – Trump vs. Clinton
- ii) mesures de sécurité (cyber);
- iii) utilisation transparente des données afin de garder la **confiance** des clients ↔↔↔ l'assureur BIG Brother;
- iv) Respect des réglementations/lois définissant l'usage/le partage des données (ex.: GDPR);

d) La fin du modèle de **mutualisation (pooling)**?:

Plus d'informations sur le profil des assurés > ciblage des clients à faible risque?



Plus difficile de trouver une garantie pour certains segments de clients vulnérables (jeunes conducteurs (assurance auto)) → risque de baisse du niveau de mutualisation et solidarité entre assurés.

Les opportunités

- a) Les assureurs possèdent déjà la masse de données clients → **atout maître** → objets connectés + masse de données + exploitation de ces données = nouveaux produits/services **personnalisés** (blockchain);
- b) Création de **nouveaux profils/rôles** avec de **nouvelles compétences**: Directeur digital/ data scientiste;
- c) Nouveau modèle **relationnel** avec les **clients** > nouveaux canaux automatisés pour développer une relation continue;
- d) Un **nouvel éco-système**: **partenariat** avec des assurtechs ainsi qu'avec les acteurs de l'internet des objets > accepter qu'une partie du know-how se trouve en externe;
- e) **Optimisation** des opérations: nouveaux dispositifs (fitbit – boîtes noires (la télématique)) > nouvelles données > tarification plus adaptée et efficace + une meilleure gestion/prévention des risques;
- f) Meilleure **gestion des sinistres**: plus rapides – prioriser les sinistres selon leur complexité; meilleur « loss control »;
- g) **Segmentation** plus fine > actions targeting plus ciblées;

Les Risques

a) Le respect de la vie privée:



b) La Question de la Responsabilité - Le dilemme éthique:

Durant les secondes précédant un accident, un véhicule autonome devrait-il faire tout ce qu'il peut pour protéger ses passagers, même si cela implique de causer du tort à d'autres automobilistes ou piétons ?



Responsabilité de QUI? Le constructeur automobile ?



Qui est responsable pour un risque de catastrophe en cas de défaillance de ces objets?

Les Risques

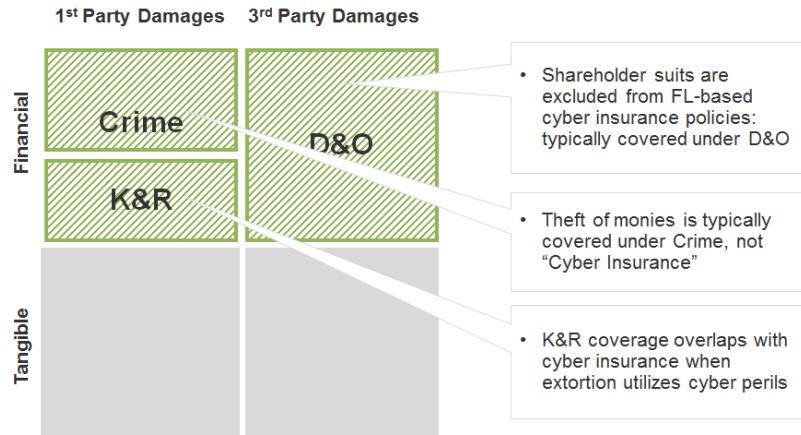
c) Le Cyber:

- De leur nature, les objets connectés et les données qu'ils produisent sont très susceptibles d'être l'objet d'une attaque cyber;
- L'impact d'un incident cyber sur le prix d'une action: une baisse en moyenne de 1,8%;
- 2/3 des entreprises victime d'un incident cyber > recul de leur cours boursier;

Plan d'action:

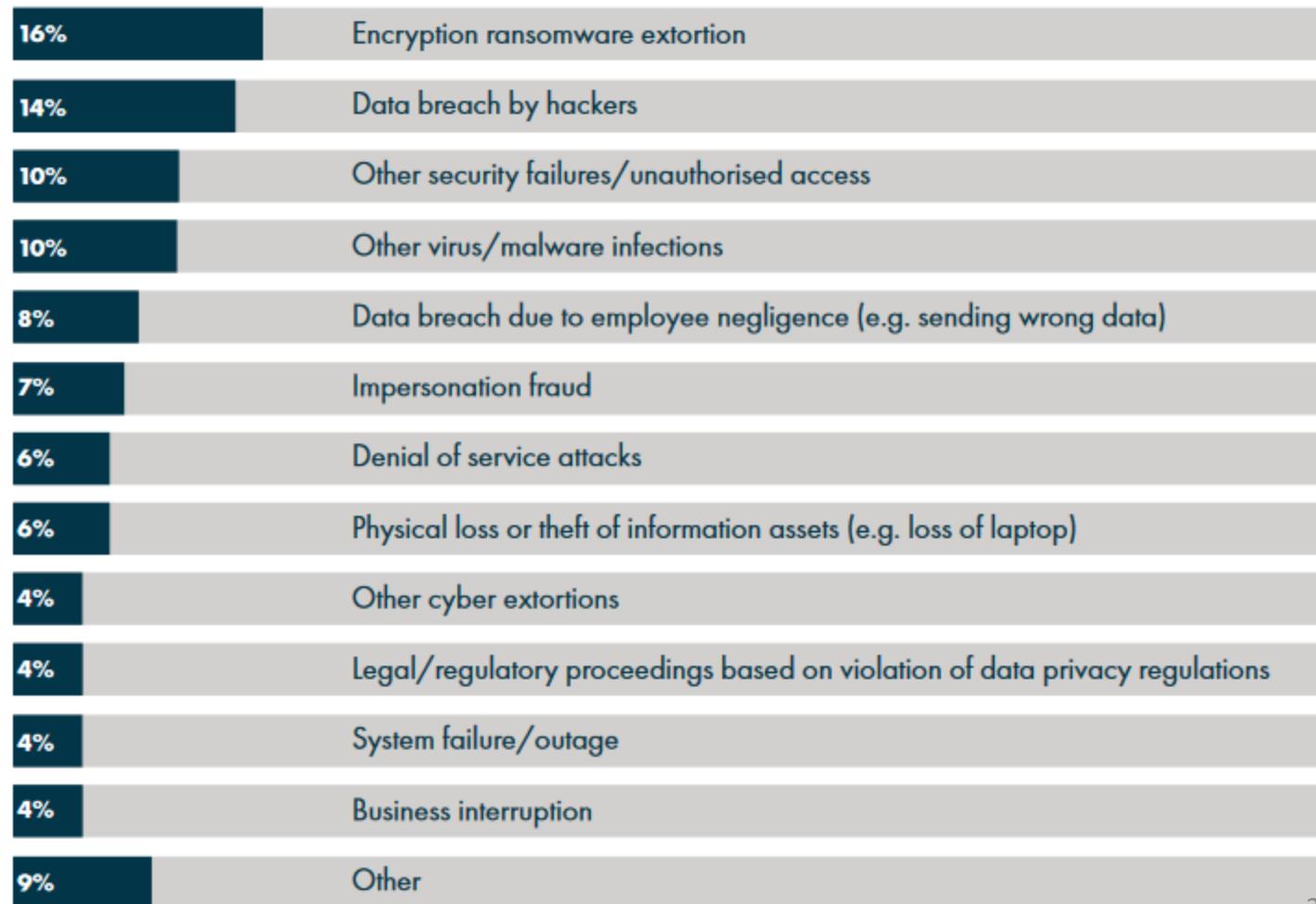
- a) Augmenter le budget sécurité cyber: actuellement à environ **5%** en moyenne du budget total IT;
- b) Vulgariser le risque cyber en tant que risque majeur auprès du Comex;
- c) Considérer le risque comme un risque opérationnel ET organisationnel: pas uniquement un risque technique.

Exemple d'un incident cyber impactant plusieurs polices d'assurance:



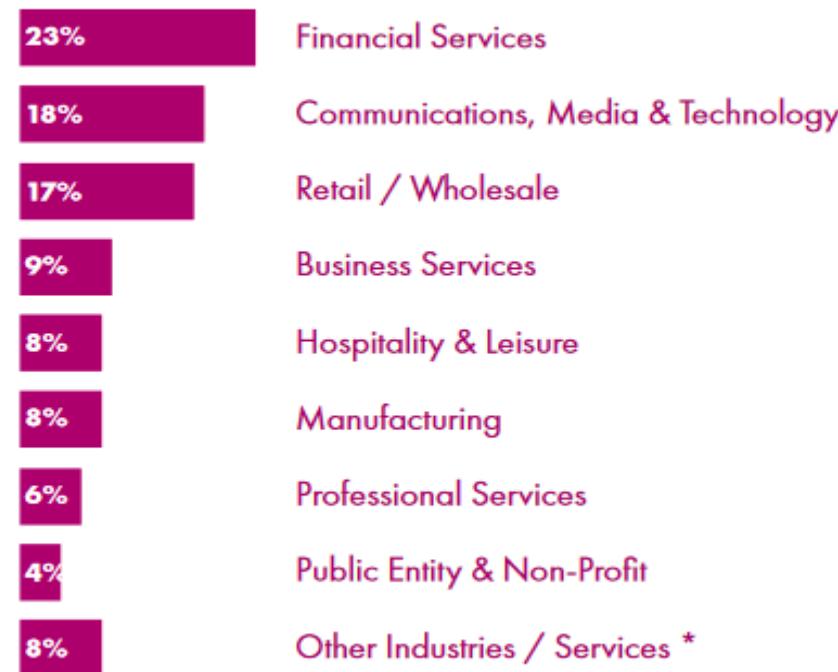
Cyber – Types de sinistres

Cyber claims received by AIG EMEA (2013-2016) - By type



Cyber – Segmentation par Industrie

Cyber claims received by AIG EMEA (2013-2016) - By industry



*Construction, Food & Beverage, Information Services, Other Services, Transportation,
Agriculture & Fisheries, Energy and Real Estate



Vice-Président du Club des Experts de la Sécurité de l'Information et du Numérique (CESIN)

Eric A. CAPRIOLI

Avocat au Barreau de Paris

Docteur en droit

Membre de la délégation française aux Nations Unies

PARTIE 3

Les données Numériques : quels enjeux juridiques ?

Le patrimoine informationnel, c'est quoi ?

Confidentialité : élément de la sécurité applicable de l'émission de la donnée jusqu'à son archivage (ISO 27001)



Données confidentielles par nature
secret professionnel, secret bancaire, secret des données de santé, etc.

Données confidentielles du fait de leur traitement
Données à caractère personnel

Données confidentielles du fait de leur valeur
Fichier client, savoir faire, etc.

Informations protégées
Droits de propriété intellectuelle

Quels sont les risques d'atteintes ?



Fuite critique massive : espionnage ciblé, vol organisé de supports... (ex : Advanced Persistent Threats/APT)



Fuite critique volontaire : détournement d'informations par un salarié, etc.



Fuite involontaire : session non fermée, ouverture de mails ou de pièces jointes contenant un malware, etc.

Quels sont les risques d'atteintes ?



Protection des DCP : loi « Informatique et Libertés »

- **Loi du 6 janvier 1978 modifiée dite « Informatique et Libertés »**
- **Consentement préalable**
- **Obligation pour le responsable de traitement d'assurer la sécurité des DCP** (art. 34)
 - ✓ Un sous-traitant qui doit présenter des garanties suffisantes (art. 35)
- Une obligation de **notification** des violations DCP à la charge des seuls **opérateurs de communications électroniques** (art. 34 bis)
- **Sanctions administratives** (CNIL) Jusqu'à 3.000.000 € depuis la loi pour une République numérique et **pénale** (articles 226-17 et 226-17-1 du Code pénal)



DLP : De la protection contre les fuites d'informations au contrôle de l'activité des salariés

CA de Paris, le 12 mai 2016 a invalidé le licenciement disciplinaire d'un salarié qui avait, en contradiction avec la charte informatique, adressé de son poste professionnel vers son email personnel 3 messages contenant 5 pièces jointes non chiffrées des documents professionnels classifiés « confidentiel ». Fuite découverte avec le DLP.

Le système n'avait pas seulement pour objet la sécurité du SI mais également le contrôle de l'activité des salariés et qu'il s'agissait d'un « *changement de finalité du contrôle caractérisant une modification substantielle du filtrage et de la collecte d'informations* », changement nécessitant une déclaration modificative auprès de la CNIL. En l'absence de cette déclaration et d'une information du comité d'entreprise, les données collectées au moyen de cet outil constituent des preuves illicites qui doivent être écartées des débats.

La charte informatique doit aussi prévoir l'utilisation de ce type d'outil

Règlement européen : les nouvelles obligations

- Une obligation de sécurité renforcée
 - ✓ Accountability
 - ✓ Politique de sécurité)
 - ✓ Analyse d'impact
 - ✓ La notification des violations de DCP pour tous les secteurs
- Des sous traitants impliqués dans les exigences de sécurité



Les mesures à prendre : 3 axes complémentaires

Technique/Sécurité
Organisation



Juridique ET
Conformité

Communication

Juridique : le contrôle de l'activité

- Le pouvoir de direction de l'employeur
- **Le respect des critères « informatique et liberté »** : loyauté, transparence et proportionnalité
- **Une information des personnels**
- Un **formalisme** à respecter
- **Assurer effectivement la sécurité du SI** : **Contrôle statistique vs. Contrôle nominatif**



A défaut, un contrôle et une sanction invalidée

Protection des bases de données par le CPI

- **Le producteur de la base de données est protégé par le droit d'auteur (art. L. 341-1 et suivants du CPI) :**
 - originalité de son organisation, classement, tri, extraction
 - Droit sui generis : moyens humains et financiers pour la développer.
- **Le producteur peut interdire (sauf exceptions)** : l'extraction, par transfert permanent ou temporaire de la totalité ou d'une partie qualitativement ou quantitativement substantielle du contenu d'une base de données ; La réutilisation, par la mise à la disposition du public de la totalité ou d'une partie qualitativement ou quantitativement substantielle du contenu de la base,
- **La base est le contenant qui doit être distingué des données contenues, lesquelles peuvent être (ou pas) protégées.**

En cas de violation : action en contrefaçon (L.343-4 CPI)

Sanctions : 3 ans de prison et 300.000 € d'amende (plus élevée si en bande organisée)

Une stratégie globale, pluridisciplinaire

Charte informatique

Contrat de travail

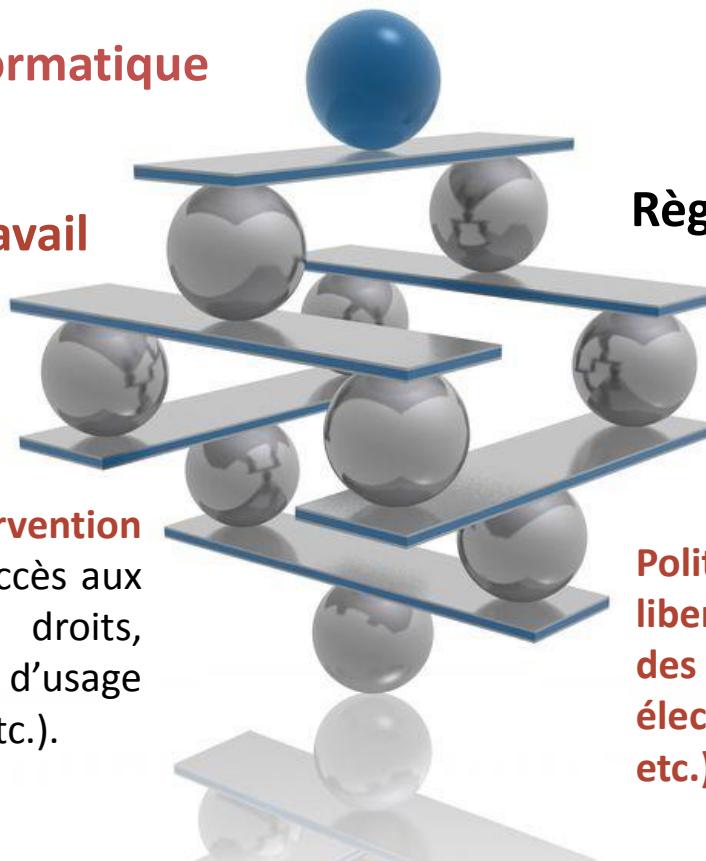
Procédures d'alerte

Procédures écrites d'intervention

(intervention sur poste, accès aux logs, transmission des droits, d'information en cas d'usage présumé non conforme, etc.).

Règlement intérieur

PSSI



Politiques (« informatique et libertés », confidentialité, gestion des logs, conservation des messages électroniques, Gestion des accès, etc.).

Et demain, l'atteinte au secret des affaires

- L'échec des multiples propositions du député Bernard Carayon
- Adoption de la directive le 8 juin 2016 sur la protection des savoir-faire et des informations commerciales non divulgués (secrets d'affaires) contre l'obtention, l'utilisation et la divulgation illicites.
- Les données protégées devront faire l'objet d'un marquage.
- En attente de la transposition

Transposition anticipée : Une tentative avortée avec la loi Macron



En synthèse



Et l'avenir ?



« Le monde de demain sera gouverné par des hommes qui s'appuieront sur des analyses fines et spécifiques de ce Big data. Potentiellement, celui d'après-demain pourrait être gouverné par des machines qui auront créé assez de data décisionnelle et de retour d'expérience étayé pour décider en autonomie. »

Les Echos

?

?

?

?

?

?

?

?

Avez-vous des questions ?

MERCI DE VOTRE ATTENTION !

**AVANT DE PARTIR , N'OUBLIEZ PAS DE REMPLIR
L'EVALUATION !**

Soit sur la feuille, à remettre à l'hôtesse à la sortie
Soit directement sur la **WEB APPLI**

Merci : vous participez à l'objectif ZERO PAPIER !

**Les slides seront en ligne dès la semaine prochaine sur
www.amrae.fr**