



**RENCONTRES**  
INTERPROFESSIONNELLES  
DES ACTEURS DU LIVRE  
10<sup>e</sup> ÉDITION

# Tirer parti de l'intelligence artificielle

Une autre perspective pour le livre et la lecture ?

27 mars 2018, Montréal

# 1. Argument

- L'IA est "l'électricité du XXIe siècle" (Andrew Ng)

+ Coût quasi nul

- Nécessite une grande quantité de données.

= Contrairement à la plupart des industries culturelles, le milieu du livre récolte ses données depuis longtemps

➔ Nécessité de travailler de concert à 1) la valorisation des données existantes  
2) la structuration des données futures

## 2. Contexte économique

- **Investissements financiers et humains massifs pour l'IA**
  - Fonds privés : +100 millions pour Element.ai
  - Fonds publics : +100 millions pour la stratégie IA (ex. IVADO)
- ➔ Montréal capitale de l'IA
- **Monde du livre**
  - Économie de la recommandation qui s'informatise
  - Selon F. Benhamou : économie de biens ➔ économie de services fondée sur le partage, moins d'appétit pour la propriété
- **Stratégies** numériques de communication/marketing et d'édition (Alto, La Pastèque)
- **Objectif** : viser mieux pour vendre plus ? Une recommandation ciblée

# 3. Qu'est-ce que l'intelligence artificielle (IA) ?

- Simulation d'une activité normalement réservée à l'être humain (**discrétisation**, cf. Leibniz)
- Différence IA et **algorithmes** ? = Synonymes
- Applications macro : voiture autonome, médecine, chatbot, optimisation énergétique, élections...
  - ➔ *comprendre* une masse de données représentative d'un ensemble  
« voir ce qu'on ne voit pas »
- Application micro : recommandation prédictive et publicité ciblée, traduction, adaptation à l'individu (assistant personnel)

**Risques** : croire en 1) la neutralité de l'algorithme  
2) l'absolu de la donnée

Ex : TAI, concours de beauté, cf. *Weapons of Math Destruction* (O'Neil, 2016)

# 4. Données et métadonnées

## Données



## Métadonnées

- Photo prise le 20 mars 2018
- Géolocalisée au Lac Saint-Jean
- 4 personnes identifiées
- 2 filles, 2 garçons
- Sourient → **corrélation** : famille ?
- Etc...

- Métadonnées = données qui décrivent d'autres données. Big data = métadonnées + données
  - **Caractéristiques** : les métadonnées sont modulables, internes ou externes, descriptives et structurelles, administratives et techniques
  - **Objectifs** des métadonnées : organiser les connaissances pour exploiter le document
- L'IA **transforme** la donnée en métadonnée

## 4. Données et métadonnées

- **Monde du livre :**
  - **Métadonnées sur l'acheteur :** âge, catégorie socio-professionnelle, nombre d'enfants, fréquence et historique d'achats, préférences (*likes* Facebook, centres d'intérêt), statistiques de lecture (temps de lecture quotidien, relectures, vitesse, etc.), ...
  - **Métadonnées sur le livre :** aujourd'hui Google Books, mais demain ?
  - **Objectif :** moissonner ces données (les récolter), les croiser, tirer des corrélations. Ex. bases de données décisionnelles, cf. ex-pdg Hachette

# 5. IA = données

- **Objectif** : joindre données massives et IA
- Contexte favorable (création massive de données, collecte constante) 2,5 quintillions d'octets créés chaque jour ; 2 zettabytes/an en 2019 pour tout le trafic web.  
1Go = 1 film → 200 petabytes = tout ce qui a jamais été imprimé → 5 exabytes = l'ensemble des mots prononcés depuis le début de l'humanité → 1 zettabyte = un milliard de milliards d'octets, soit tout ce qui a été mis en ligne l'année 2010 (*Analyse MFG Labs*)
- Accélération demain avec l'internet des objets  
→ « L'intelligence artificielle fait émerger l'information à partir de données éparses, elle donne un sens aux données. » Valérie Bécaert, directrice IVADO
- Les GAFAM proposent en accès libre leurs algorithmes d'IA : **pourquoi ?**

## 6. Les données existent dans le milieu du livre

- Acteurs privés : récolte depuis les débuts (Amazon, Kobo), ou de façon plus récente (grandes chaînes, librairies indépendantes)
- Acteurs publics ou semi-publics : IVADO, BTLF (Sté de gestion à financement public)

➔ **Constat** : la récolte indépendante **ne suffit pas**, les données des PME sont largement insuffisantes

« Pour tirer profit de l'IA, il faut d'abord s'attaquer aux données. » J-F Gagné, Element.ai

**Objectif** : Valoriser plutôt que créer

**Comment ?** - Soigner les métadonnées : récolter et structurer

- Développer un travail commun, car la quantité de données fait la force



## 6. Les données existent dans le milieu du livre

- **Pour le livre numérique :**
    - Possibilité de récolter tous les types de données, notamment de lecture (ex. Amazon paie à la quantité de lecture, Kobo) → phénomène de privatisation
  - **Pour le livre papier :**
    - Impossibilité de récolter des données de lecture du client
    - Possibilité de récolter des données de lecture sur le livre (ex. Alibaba, IA score 82,44 au *Stanford Question Answering Data Set* c. meilleur résultat humain 82,304)
    - Possibilité de récolter des données de vente (avant) et des avis (après)
  - **Défi :** Comment collecter ces deux types de données ? Comment les croiser ?
- Importance d'un travail concerné et concerté pour identifier, collecter, conserver, traiter

# 7. L'IA : prédiction et production

- **Contexte :**
  - Surabondance (356 millions de livres, chiffres *GFK 2018*)
  - Implication des éditeurs dans la recommandation (normalement réservée aux libraires)
- **Non IA : recommandation humaine**
  - Sites communautaires spécialisés (plateforme de lecture sociale des LIQ)
  - Blogs spécialisés (Le Blog des livres, Tamaculture, Bookivores, Stalker...)
  - YouTubeurs (Miss book, Le marque page...)
- ➔ Diversité des propositions, modèles fondés sur la personnalité de l'influenceur

# 7. L'IA : **prédiction** et production

## IA : recommandation machine

- **Prédiction macro** : en fonction des autres
    - “D’autres lecteurs ont aimé” / “Vous aimerez peut-être”
    - IA : Amazon et la recommandation de lecture : on table sur un comportement *futur*. Politique agressive d’acquisition des données (ex : le rachat de Goodreads, qui passe à 55 millions de membres)
  - **Prédiction micro** : en fonction de soi
    - Contexte de **longue traîne** → donner de la place aux livres qui passent sous le radar
    - Correspondre au plus près au lectorat
- Modèles fondés sur la personnalité du lecteur

# 7. L'IA : **prédiction** et production

## **Pour le libraire ? (ex. Flaubert)**

- Hypomnémata (cf. Stiegler) : Mémoire humaine / mémoire numérique
- IA sur la lecture + IA sur les métadonnées
- ➔ Moteur de recherche *ad hoc* qui fouille à la fois dans la masse des livres existants + *dans* le contenu des livres pour proposer des recommandations précises
- **Nouveau métier** (cf. F. Benhamou) : une table du libraire, faite de livres récents sur le sujet + de livres du passé (éventuellement en impression à la demande)
- ➔ Le libraire comme influenceur doté de nouveaux outils.

## 8. L'IA : prédiction et **production**

Les productions écrites par des humains

- **Sans IA :**
  - Sites spécialisés : ex. Wattpad (écrire et se faire recommander quoi améliorer par le public), cf. Anna Todd → une influence sur l'autoédition avant tout.
  - Mécénat : ex. Tipeeee.com
  - YouTube : ex. Nine Gorman – 66 000 abonnés
- **Avec IA :** ex. *The Best Seller Code* – critères très fins (une héroïne, un chien, etc.) – modèle de production *Netflix* (House of Cards)

# 8. L'IA : prédiction et **production**

Les productions écrites par des machines

- Musique, peinture, vidéo, cuisine
- Littérature ?
  
- **Avec IA, le livre hors de la littérature :**
  - Livres d'art
  - Philip Parker (200 000 livres)
  - Entreprises (Narrative Science)
  - Journalisme (Libération)
  - Arnaques / blanchiment

## 8. L'IA : prédiction et **production**

Les productions écrites par des machines

- Concours littéraires (Shinichi Literary Award)
  - Collections d'édition (Using Electricity)
  - Sites dédiés (Literai.com)
- 
- Importance d'une narration « de bon sens », mais textes poétiques faciles à produire

➔ Vers un **outil d'écriture** ?



**RENCONTRES**  
INTERPROFESSIONNELLES  
DES ACTEURS DU LIVRE  
10<sup>e</sup> ÉDITION

# Tirer parti de l'intelligence artificielle

Une autre perspective pour le livre et la lecture ?

Merci !