

# 03-01

## Systemes de recommandation I

420-A58-SF — Algorithmes d'apprentissage non supervisé — Hiver 2023  
Spécialisation technique en intelligence artificielle — M. Swawola, M.Sc.

---

**NOUS ÉCLAIRON.**  
**VOUS BRILLEZ.**

---

FORMATION CONTINUE  
ET SERVICES AUX ENTREPRISES



# Sommaire

1. Introduction
2. Représentation du modèle
3. Lectures et références

# Sommaire

1. Introduction
2. Représentation du modèle
3. Lectures et références

# Introduction

## ■ Utilisateur “Mike”

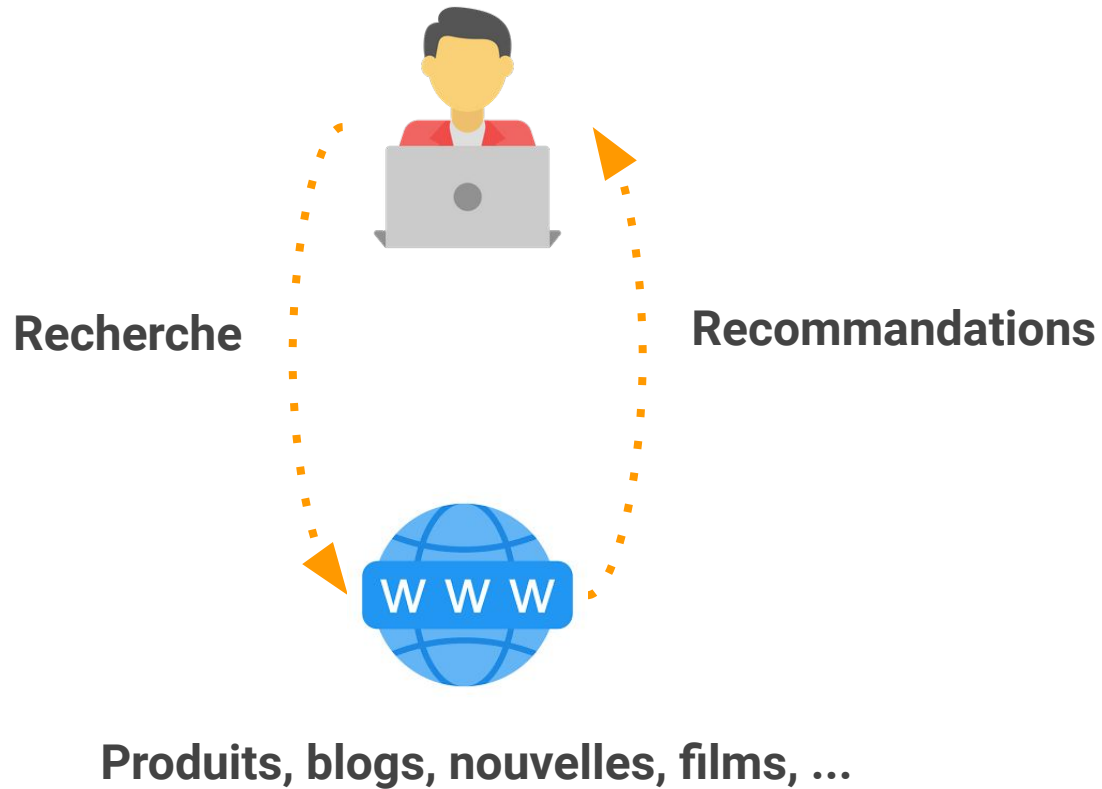
- Aime les films avec **Steven Seagal**
- Adore les films avec **Jean-Claude Van Damme**

## ■ Utilisateur “Bob”

- Recherche les films avec **Steven Seagal**
- **Se fait recommander les films avec JCVD**



# Introduction



## Quelques exemples

amazon



last.fm



NETFLIX

The  
Economist



You Tube

Linked in



# De la rareté ...

- Dans la distribution traditionnelle, l'espace est **limité**
  - Surface du magasin, nombre de rayons, hauteur des étagères, etc...
- Cette limite est également valable pour les cinémas, les chaînes de télévision, les pharmacies, etc...



Un magasin "brick-and-mortar" fermé ...

## ... à l'abondance

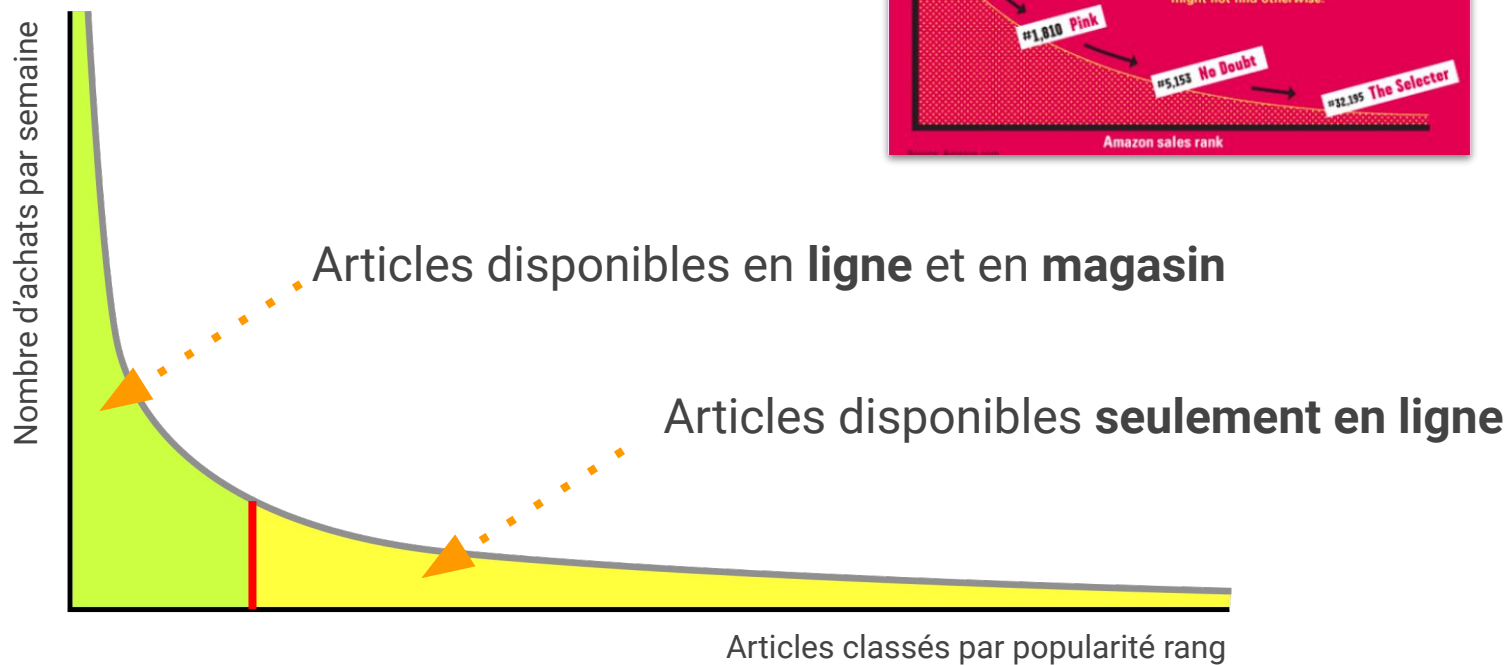
- Le Web a permis une diffusion à coût quasi-nul **d'information** sur les tous les produits
  - Nous sommes passés **de la rareté à l'abondance**
  - Phénomène de **longue traîne / long tail** (Chris Anderson, 2004) [3]



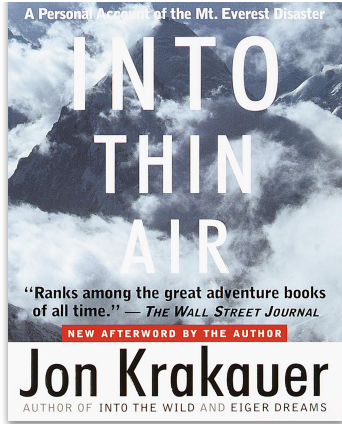
Tout est à portée ~~de la main~~ d'un clic !



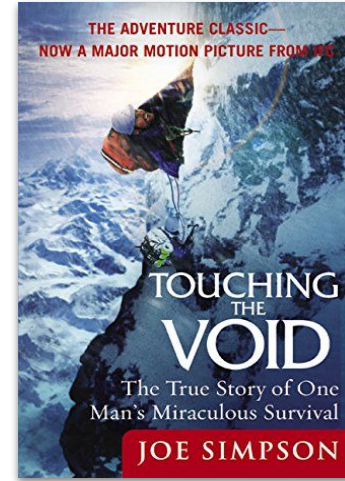
# La longue traîne



# La longue traîne - Anecdote



Paru en 1997



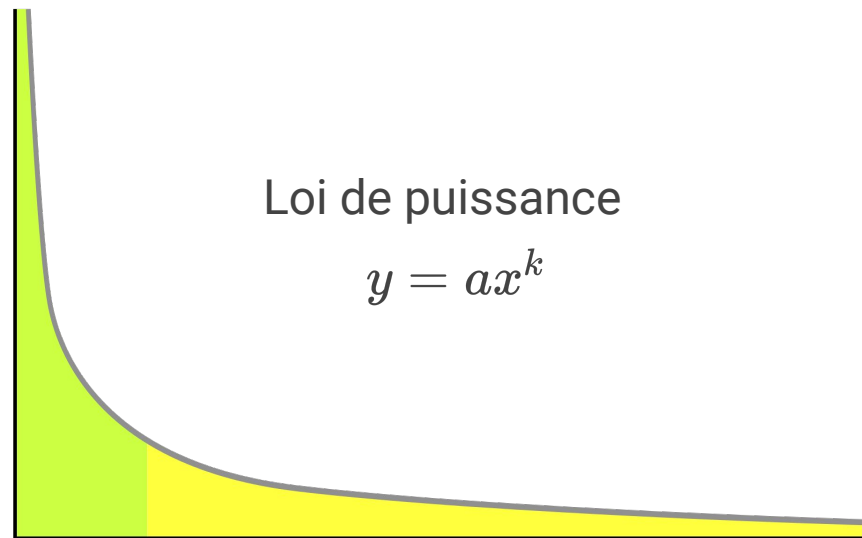
Paru en 2003

Devenu un bestseller après 2003 !

# La longue traîne - Applications

## ■ Exemples

- Livres, films, musique
- Articles de nouvelles
- Achats
- Contacts (recommandation d'amis sur Facebook, LinkedIn et Twitter)
- Etc...



# Les différents types de recommandation (1/3)

## ■ Éditoriale et préparée manuellement

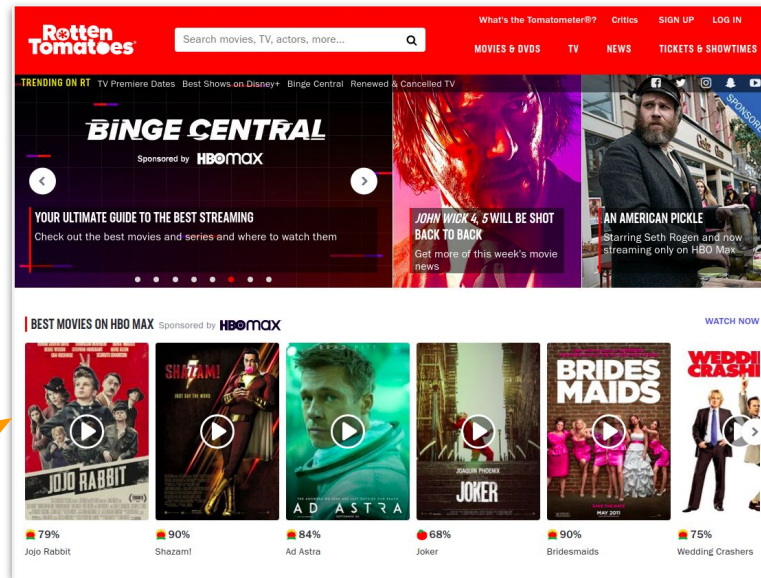
- Le choix de la rédaction
- Les 16 livres à lire absolument en 2019 selon Bill Gates
- Etc...



# Les différents types de recommandation (2/3)

## ■ Listes agrégées

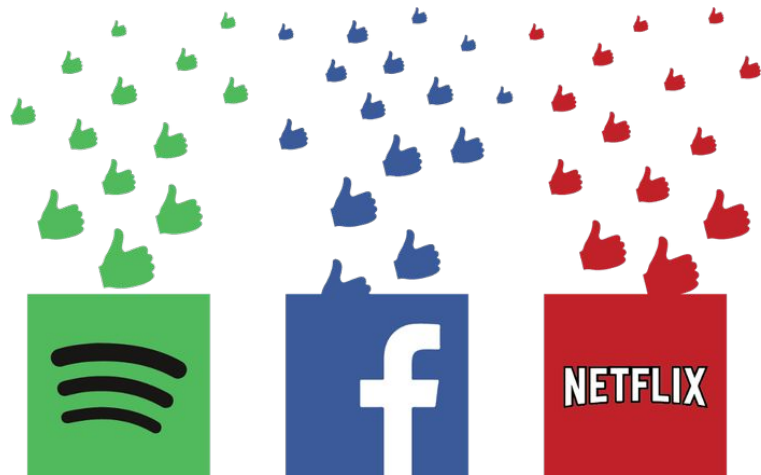
- Top 10 des séries télévisées
- Articles les plus populaires
- Jeux les plus téléchargés
- Etc ...



# Les différents types de recommandation (3/3)

## ■ Recommandation personnalisée à chaque utilisateur

- Contenu
- Achat
- Connections sur les réseaux sociaux
- Etc ...



## ■ Système de recommandation / Recommender Systems

→ **Sujet de ce cours !**

# Les différents types de recommandation (3/3)

## ■ Recommandation personnalisée pour chaque utilisateur

- Contenu
- Achat
- Connexions sur les réseaux sociaux
- Etc ...



## ■ Système de recommandation / Recommender Systems

→ **Sujet de ce cours !**

# Les différents types de recommandation (3/3)

- Recommandation personnalisée à chaque utilisateur

- Contenu
  - Achat
  - Connexions sur les réseaux sociaux
  - Etc ...
- Les systèmes de recommandation sont basés sur des données: les “notes” attribuées aux produits par les utilisateurs**



- **Sujet de ce cours !**



# Notation

## ■ Explicite

- Demander aux utilisateurs de noter / Liker les produits

## ■ Implicite

- Apprendre à partir des actions des utilisateurs (achat, ...)
- Quelle signification pour les faibles notes ?



**0 → 5 étoiles**

# Les différentes approches

- Il existe différentes approches adressant la problématique de recommandation
  - Systèmes de recommandation basés sur le contenu
  - Filtrage collaboratif
  - Systèmes hybrides
  - Systèmes de recommandation basés sur les facteurs latents

# Sommaire

1. Introduction
2. Représentation du modèle
3. Lectures et références

# Représentation du modèle

- $J$  est l'ensemble des **utilisateurs** / consommateurs / etc...
- $I$  est l'ensemble des **produits** / items / séries / articles / etc...
- **Fonction d'utilité**  $y: J \times I \rightarrow Y$ 
  - $Y$  est l'ensemble des notes (ensemble totalement ordonné)
  - Par exemple: de **0 à 5 étoiles** ou de **0 à 100%**

# Matrice d'utilité

- Fonction d'utilité  $y: J \times I \rightarrow Y$

	$j^{(1)}$	$j^{(2)}$	$j^{(3)}$	$j^{(4)}$
$i^{(1)}$	$y^{(1,1)}$	$y^{(1,2)}$	$y^{(1,3)}$	$y^{(1,4)}$
$i^{(2)}$	$y^{(2,1)}$	$y^{(2,2)}$	$y^{(2,3)}$	$y^{(2,4)}$
$i^{(3)}$	$y^{(3,1)}$	$y^{(3,2)}$	$y^{(3,3)}$	$y^{(3,4)}$
$i^{(4)}$	$y^{(4,1)}$	$y^{(4,2)}$	$y^{(4,3)}$	$y^{(4,4)}$

Utilisateurs

Items

Matrice d'utilité

$$Y = \begin{bmatrix} y^{(1,1)} & y^{(1,2)} & y^{(1,3)} & y^{(1,4)} \\ y^{(2,1)} & y^{(2,2)} & y^{(2,3)} & y^{(2,4)} \\ y^{(3,1)} & y^{(3,2)} & y^{(3,3)} & y^{(3,4)} \\ y^{(4,1)} & y^{(4,2)} & y^{(4,3)} & y^{(4,4)} \end{bmatrix}$$

# Matrice d'utilité

- Fonction d'utilité  $u: J \times I \rightarrow Y$

	$j^{(1)}$	$j^{(2)}$	$j^{(3)}$	$j^{(4)}$
$i^{(1)}$	$y^{(1,1)}$	$y^{(1,2)}$	$y^{(1,3)}$	$y^{(1,4)}$
$i^{(2)}$	$y^{(2,1)}$	$y^{(2,2)}$	$y^{(2,3)}$	$y^{(2,4)}$
$i^{(3)}$	$y^{(3,1)}$	$y^{(3,2)}$	$y^{(3,3)}$	$y^{(3,4)}$
$i^{(4)}$	$y^{(4,1)}$	$y^{(4,2)}$	$y^{(4,3)}$	$y^{(4,4)}$

Utilisateurs

En réalité, la matrice d'utilité  $Y$  est une matrice creuse

Items

Matrice d'utilité


$$Y = \begin{bmatrix} y^{(1,1)} & y^{(1,2)} & y^{(1,3)} & y^{(1,4)} \\ y^{(2,1)} & y^{(2,2)} & y^{(2,3)} & y^{(2,4)} \\ y^{(3,1)} & y^{(3,2)} & y^{(3,3)} & y^{(3,4)} \\ y^{(4,1)} & y^{(4,2)} & y^{(4,3)} & y^{(4,4)} \end{bmatrix}$$

# Matrice d'utilité

- Fonction d'utilité  $y: J \times I \rightarrow Y$

	$j^{(1)}$	$j^{(2)}$	$j^{(3)}$	$j^{(4)}$
$i^{(1)}$	$y^{(1,1)}$	$y^{(1,2)}$	$y^{(1,3)}$	$y^{(1,4)}$
$i^{(2)}$	$y^{(2,1)}$	$y^{(2,2)}$	$y^{(2,3)}$	$y^{(2,4)}$
$i^{(3)}$	$y^{(3,1)}$	$y^{(3,2)}$	$y^{(3,3)}$	$y^{(3,4)}$
$i^{(4)}$	$y^{(4,1)}$	$y^{(4,2)}$	$y^{(4,3)}$	$y^{(4,4)}$

Matrice d'utilité


$$Y = \begin{bmatrix} y^{(1,1)} & y^{(1,2)} & y^{(1,3)} & y^{(1,4)} \\ y^{(2,1)} & y^{(2,2)} & y^{(2,3)} & y^{(2,4)} \\ y^{(3,1)} & y^{(3,2)} & y^{(3,3)} & y^{(3,4)} \\ y^{(4,1)} & y^{(4,2)} & y^{(4,3)} & y^{(4,4)} \end{bmatrix}$$

# Matrice d'utilité

## ■ Fonction d'utilité $y: J \times I \rightarrow Y$

	$j^{(1)}$	$j^{(2)}$	$j^{(3)}$	$j^{(4)}$
$i^{(1)}$	?	?	?	$y^{(1,4)}$
$i^{(2)}$	?	$y^{(2,2)}$	?	?
$i^{(3)}$	$y^{(3,1)}$	?	?	$y^{(3,4)}$
$i^{(4)}$	?	?	$y^{(4,3)}$	$y^{(4,4)}$

- La plupart des utilisateurs n'ont pas noté d'items
- **Démarrage à froid** (cold start): nouveaux utilisateurs sans historique et nouveaux items sans notes

Matrice d'utilité



$$Y = \begin{bmatrix} ? & ? & ? & y^{(1,4)} \\ ? & y^{(2,2)} & ? & ? \\ y^{(3,1)} & ? & ? & y^{(3,4)} \\ ? & ? & y^{(4,3)} & y^{(4,4)} \end{bmatrix}$$



# Exemple: prédiction de la notation de séries

- On suppose une notation de 0 à 5 (étoiles)

	Alice (1)	Bob (2)	Mike (3)	Alex (4)
Breaking Bad				
Narcos				
The Mandalorian				
The Book of Boba Fett				
Andor				

- $n_s$  est le nombre de séries  $\rightarrow n_s = 5$
- $n_u$  est le nombre d'utilisateurs  $\rightarrow n_u = 4$

# Exemple: prédiction de la notation de séries

- On suppose une notation de 0 à 5 (étoiles)

	Alice (1)	Bob (2)	Mike (3)	Alex (4)
Breaking Bad	5			
Narcos	4			
The Mandalorian	1			
The Book of Boba Fett	?			
Andor	?			

# Exemple: prédiction de la notation de séries

- On suppose une notation de 0 à 5 (étoiles)

	Alice (1)	Bob (2)	Mike (3)	Alex (4)
Breaking Bad	5	5		
Narcos	4	5		
The Mandalorian	1	?		
The Book of Boba Fett	?	1		
Andor	?	0		

# Exemple: prédiction de la notation de séries

- On suppose une notation de 0 à 5 (étoiles)

	Alice (1)	Bob (2)	Mike (3)	Alex (4)
Breaking Bad	5	5	1	
Narcos	4	5	?	
The Mandalorian	1	?	5	
The Book of Boba Fett	?	1	2	
Andor	?	0	5	

# Exemple: prédiction de la notation de séries

- On suppose une notation de 0 à 5 (étoiles)

	Alice (1)	Bob (2)	Mike (3)	Alex (4)
Breaking Bad	5	5	1	?
Narcos	4	5	?	0
The Mandalorian	1	?	5	4
The Book of Boba Fett	?	1	2	5
Andor	?	0	5	?

# Exemple: prédiction de la notation de séries

- On suppose une notation de 0 à 5 (étoiles)

	Alice (1)	Bob (2)	Mike (3)	Alex (4)
Breaking Bad	5	5	1	? → 0
Narcos	4	5	? → 1	0
The Mandalorian	1	? → 1.5	5	4
The Book of Boba Fett	? → 1	1	2	5
Andor	? → 0.5	0	5	? → 4.5

Recommander "Andor" à Alex



# Exemple: prédiction de la notation de séries

- On suppose une notation de 0 à 5 (étoiles)

	Alice (1)	Bob (2)	Mike (3)	Alex (4)
Breaking Bad	5	5	1	? → 0
Narcos	4	5	? → 1	0
The Mandalorian	1	? → 1.5	5	4
The Book of Boba Fett	? → 1	1	2	5
Andor	? → 0.5	0	5	? → 4.5

- $r^{(i,j)} = 1$  si l'utilisateur  $j$  a noté la série  $i$
- $y^{(i,j)}$  = note donnée par l'utilisateur  $j$  sur la série  $i$  (si et seulement si  $r^{(i,j)} = 1$ )

# Exercice

- En considérant la notation utilisée dans ce cours, quelles valeurs prennent les variables  $n_s$  et  $n_u$  ? Que valent  $r^{(2,3)}$  et  $y^{(2,3)}$  ?

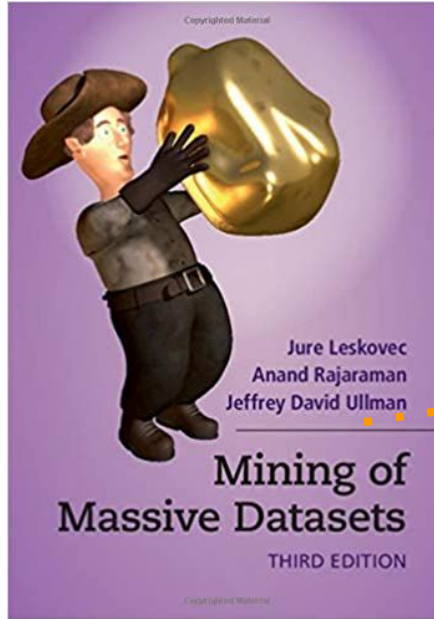
	Utilisateur 1	Utilisateur 2	Utilisateur 3	Utilisateur 4
Série 1	?	?	5	2
Série 2	3	0	?	1

- Réponses:
  - $n_s = 2$  et  $n_u = 4$
  - $r^{(2,3)} = 0$  et  $y^{(2,3)} = \text{indéfini}$



# Sommaire

1. Introduction
2. Représentation du modèle
3. Lectures et références



9 Recommendation  
Systems  
p. 319-353

Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman, **Mining of Massive Datasets,  
3rd edition**

# Références

- [1] CS229: Machine Learning - Stanford University
- [2] [Mining of Massive Datasets, 3rd edition](#)
- [3] [The Long Tail, Chris Anderson, 2004](#)