

# On collection development in the digital era

[vincent.lariviere@umontreal.ca](mailto:vincent.lariviere@umontreal.ca)

Vincent Larivière

[@lariviev](#)

[crc.ebsi.umontreal.ca](http://crc.ebsi.umontreal.ca)

Chaire de recherche du Canada sur  
les transformations de la communication  
savante

École de bibliothéconomie et des  
sciences de l'information

Université   
de Montréal



Observatoire  
des Sciences et des  
technologies

The logo features the word 'érudit' in a bold, black, sans-serif font, with a red accent on the 'é'.

# Outline

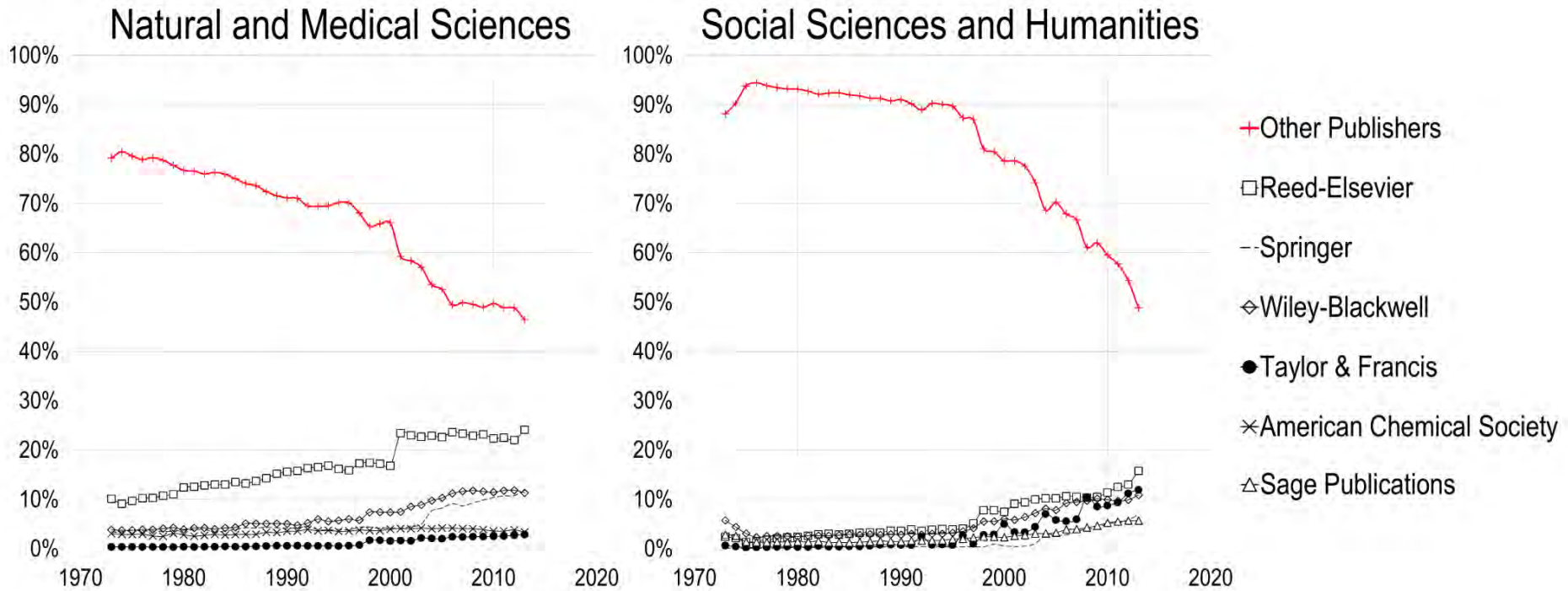
- Scholarly communication in the digital era
- Collection development (UdM data)
  - Data sources and indicators
  - Methodology
  - Results
- Limitations
- Outlook

# The digital era

- Early 1990s typically considered as turning point
- Democratization of use mid 1990s
- Characteristics of digital scientific information:
  - Easy to create, update, reuse, access, transmit
  - Require less space

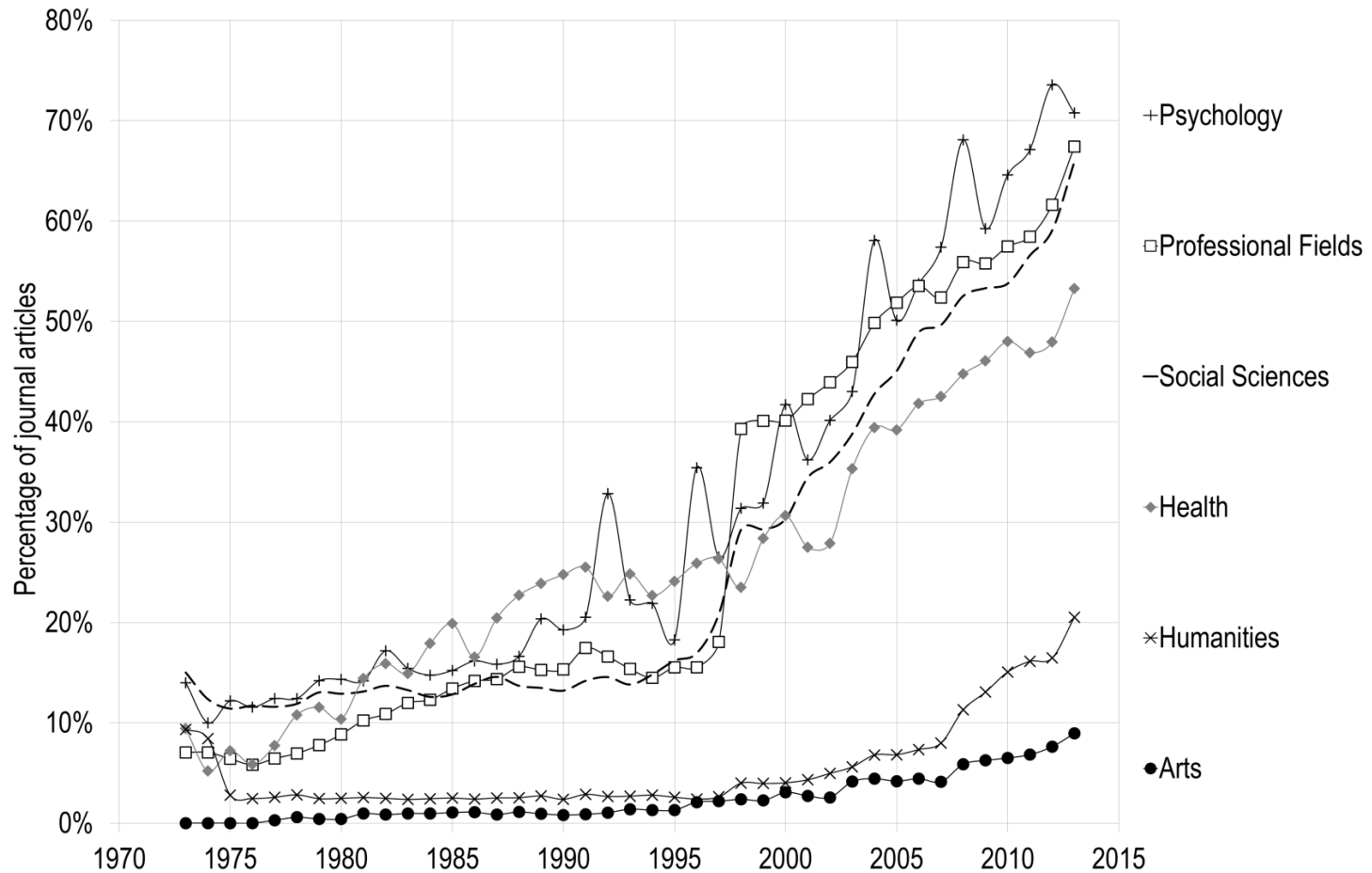
# Consolidation of the publishing industry

## Decline of « independent » scientific journals



# Consolidation of the publishing industry

## Control of top 5 big publishers in the SSH



# Collection development

- Increasing necessity!
- Decline of \$CAN (and invoices are in \$US)
- Cuts in the educational sector in Quebec
  - Cuts in the educational sector
- Quebec universities are said to be underfunded
  - Lower amount spent per student compared to the rest of Canada
- Many universities are unable to afford the big deals anymore
  - Analyses conducted at four Quebec universities
  - More than 20 Canadian universities will perform such analysis within the next few months

# Collection development

- Measure the usage of the electronic journal collection
  - Variations across disciplines, publishers.
- Isolate journals that are most important to the community
  - Keep usage as high as possible while lowering costs.
- Regain control of spending on scientific journals
  - To be able to buy books...
- We have access to more resources, but are we really using them?

# Indicators

- **Downloads:** indicates usage by the university community taken as a whole
  - Important weight to students
- **References made (citations):** Indicates usage by researchers
  - Important weight to professors
- **Survey:** Indicates *perceived* usage by members of the university community
  - Graduate students, postdocs and professors



# Data sources (UdeM)

- **ULRICH database:** All existing academic/peer-reviewed journals—basis of the consultation
- **Downloads (JR1):** Number of paper downloads for each journal for the 2010-2013 period
- **References made by the UdeM community:** Data from Thomson Reuters' Web of Science ( $\approx 12,000$  journals) for the 2010-2013 period
- **Survey** of the community (2,213 respondents)
  - For graduate students, postdocs, lecturers and professors
  - 15 most important titles: 10 for their research and teaching, 5 for their field (more general journals)

# Numbers on the UdeM collection

- Academic journals indexed in ULRICH (basis for the consultation): **108,716**
- Journal subscriptions: **≈50,000**
  - Academic journals for which we have download data: 27 000
  - Resources downloaded at least once in the last 4 years: 16 830
- Journals cited by UdeM researchers: **9,279**
- Journals mentioned in the consultation: **8,263**

**Usage is not evenly distributed across journals... and publishers!**

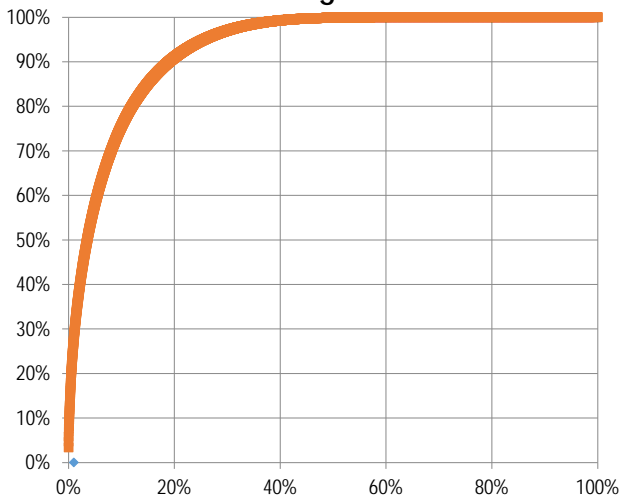
# Methodological choices

- Composite indicators are necessary arbitrary.
  - Why should download have a greater weight than citations? Or vice versa?
  - Hence, each indicator has the same weight
- Research (and download, citation) practices are field-dependent.
  - Hence, the analysis has to be performed at the level research fields
- Mentions in the survey all have the same value
  - The same weight is attributed to journals mentioned by graduate students, lecturers, professors, etc.

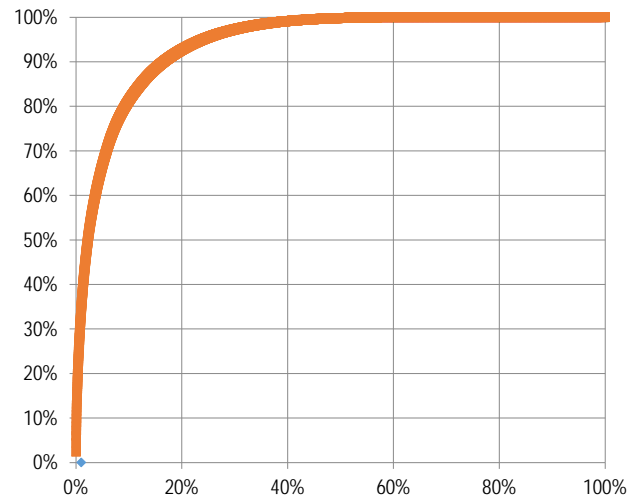
# Concentration of use

## Health

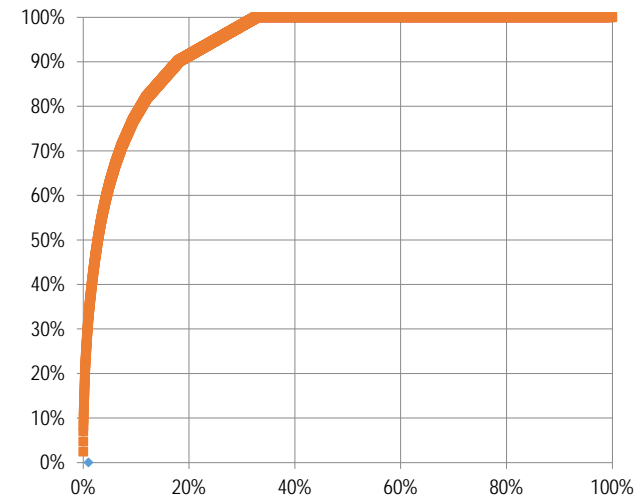
Téléchargements



Références



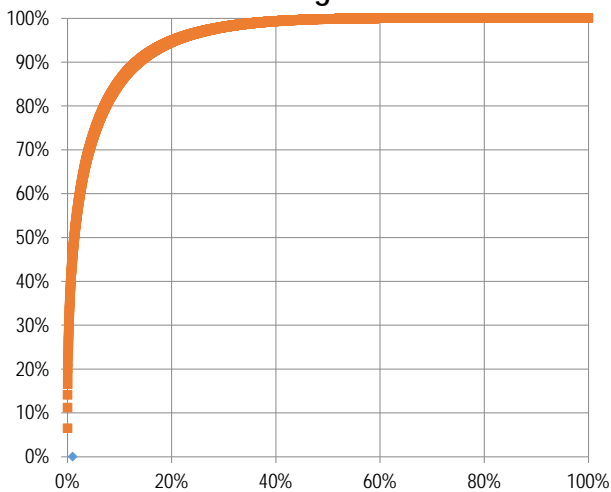
Mentions



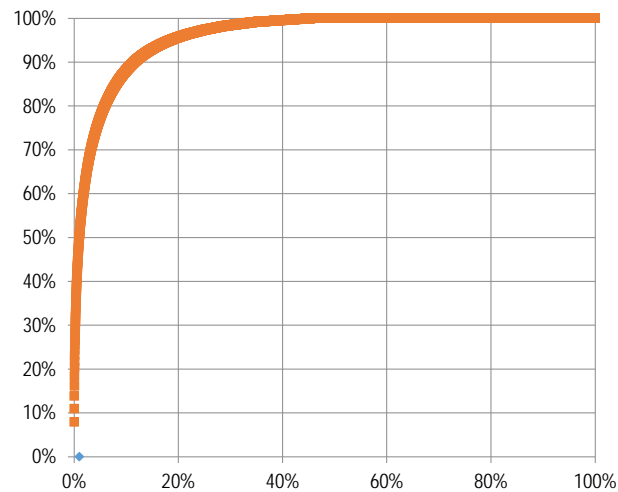
# Concentration of use

## Natural Sciences and Engineering

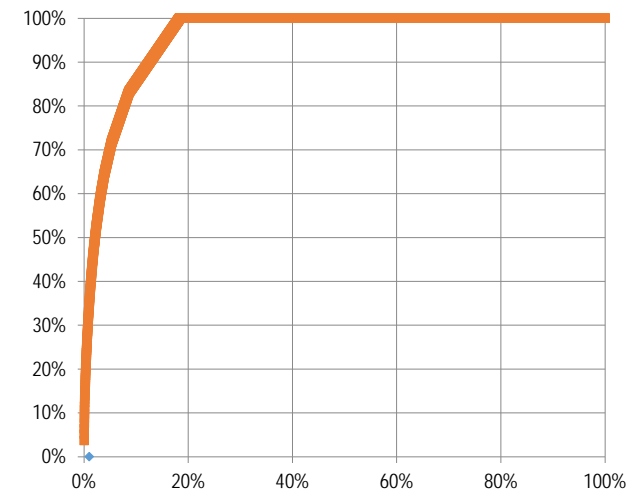
Téléchargements



Références



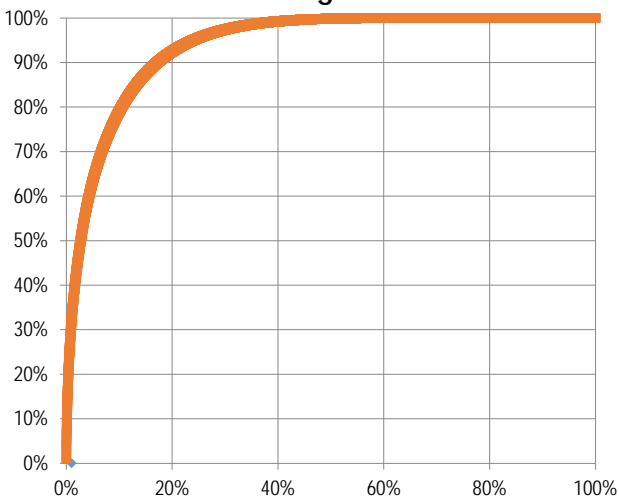
Mentions



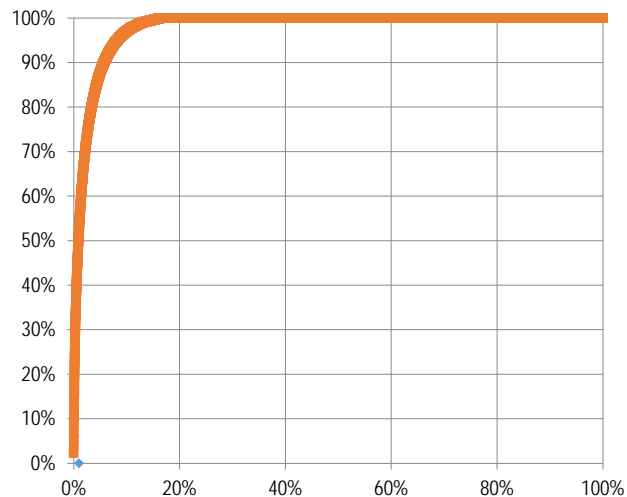
# Concentration of use

## Social Sciences and Humanities

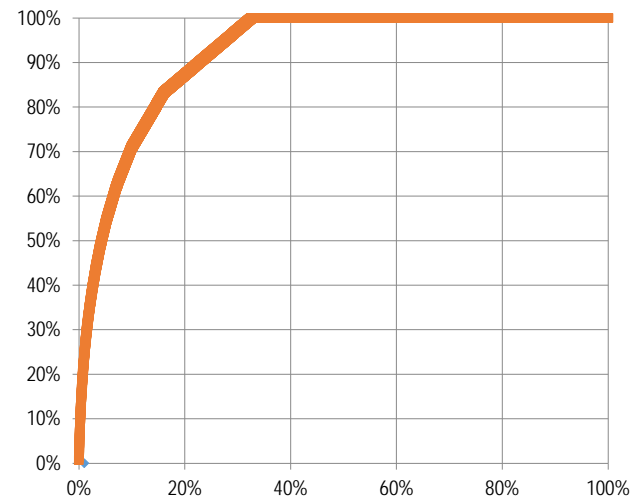
Téléchargements



Références



Mentions



# How to define core journals?

- To be considered as *core*, a journal has to be in the top 80% of downloads **OR** of citations received **OR** of mentions in the survey
- Hence, each periodical has three distinct chances of making it to the group of core journals
- By keeping 80% for any of the indicators, we end up keeping between 85% and 90% of usage

# Thresholds (annual)

---

Domain	Downloads	References	Mentions
Health	428,75	14,60	3
NSE	192,50	8,60	2
SSH	188,75	4,40	2

---



# Core journals

---

Domain	Core journals			Distinct core journals	All journals	% of Core Journals
	Downloads	References	Mentions			
NSE	397	336	584	779	6,483	12.2 %
Health	821	648	819	1,255	6,588	19.5 %
SSH	1,421	462	2,249	2,818	13,772	20.5 %
All Domains	2,639	1,446	3,652	4,852	26,843	18.1 %

---

# A few results on core journals

- Between 10% and 40% of journals from for-profit publishers
- Academic publishers generally higher than for profit publishers
- National journals/platforms are at the top
  - Mainly because of downloads

# Limitations

- Some journals do not have download data
- Humanities are at a disadvantage in terms of citations
  - And should be distinguished from the Social Sciences
- Survey's response rate was lower in some units
- Use of WoS data for citations
- Results serve as a basis for negotiation
  - Openness to negotiation varies by publisher...
- Importance of librarians for enrolling users in the process and for complementing the quantitative

# Outlook

## Since the digital era:

- Academe lost control on diffusion of knowledge
  - Except in the humanities
- Big publishers are reaping huge profits
  - Public funds diverted from academe
- Lost control on subscription costs
  - Affects book-centric disciplines
- Have access to more—but we do not use it!
  - Difficult to get out of the big deals!

# Why do we need publishers?

- Diffusion of research results?
- Peer review?
- Archive knowledge?
- Provide with a hierarchy of discoveries... and of authors, institutions, countries
  - Symbolic capital
- Publishers do not have value as such—their importance comes mainly from the (typically voluntary) work editors, authors and referees.

# Questions?

Vincent Larivière

[vincent.lariviere@umontreal.ca](mailto:vincent.lariviere@umontreal.ca)

[@lariviev](#)

[crc.ebsi.umontreal.ca](http://crc.ebsi.umontreal.ca)

Université   
de Montréal

Chaire de recherche du Canada sur  
les transformations de la communication  
savante  
École de bibliothéconomie et des sciences  
de l'information



The logo for Érudit consists of the word 'érudit' in a bold, black, sans-serif font. The letter 'é' is red, while the remaining letters are black.