To Open or Not to Open...

M. Tamer Özsu
University of Waterloo

About Me

- Professor of Computer Science, University of Waterloo (2000 - present); University of Alberta (1984-2000)

- Director, Cheriton School of Computer Science, University of Waterloo (2007-2010)

- ACM Publications Board Member (2004- present), CRA Board Member (Canadian rep.), ACM SIGMOD Chair (2001-2005), VLDB Endowment Trustee (1996-2002)

- EIC of The VLDB Journal (1997-2005), Editorial Boards of a number of journals (I think about 6 - past and present)

- PC Chair of a number of conferences, PC member of a number of conferences

Acknowledgments

- Thanks to
  - Ronald Boisvert
  - Jack Davidson
  - Bernard Rous
  - Scott Delman

for comments on these slides

Disclaimer

- Much of my thinking on these topics is influenced by discussions with colleagues on ACM Publications Board, but I do not speak for ACM or the Publications Board.

Open Access

- Green OA
  - Self-archiving in a personal or an institutional or a central repository after actual publication
  - Usually final draft copy
- Gold OA
  - Actual published copy is opened for access at the publisher's site
  - May or may not require payment (subsidized)
- Hybrid OA Journals
  - Author offered traditional subscription model or author pays model
  - Author or institution pays per article ($750-3,000)
- Author Pays
  - Only option is author pays or institution pays option
- A spectrum, but most people mean Gold OA when they refer to OA

Pros & Cons

- Subscription model
  - + Publisher's natural incentive = best quality product
  - − Commercial monopoly can lead to price gouging

- Green OA Model
  - + Publications more visible, more cited
    - Actual evidence is mixed
  - + Taxpayers who subsidize research have access
    - More compelling in other fields
  - + Easier on library budgets
    - More compelling in other fields
  - − "Open versions" difficult to locate

- Gold OA Model
  - + Same as green OA
  - − How to sustain publication?
    - No sustainable business model for a publishing ecosystem (if we want one, that is…)

- Author Pays or Hybrid Model
  - + Same as gold & green OA
  - − Publisher's natural incentive = more papers ➔ predatory publishing
  - − Unfair
  - − Need for financial sponsorship
  - − More risky for publishers
    - Pricing up-front for long-term sustainability

Copyright

- Legally giving the creator of an original work exclusive rights to that work for (usually) a limited time.
- Started in Europe in 15th century
  - Senate of Venice (1469): exclusive right for five years given to John of Spira to print the epistles of Cicero and of Pliny
- England is important
  - Licensing Act: Licenses to publishers, protecting publishers (ended 1694)
  - Statute of Anne (1710)
    - "An Act for the Encouragement of Learning, by Vesting the Copies of Printed Books in the Authors or Purchasers of such Copies, during the Times therein mentioned"
    - Rights to the creators for a period of 14 years renewable for another 14 of creator alive, then falls into public domain
- First federal US copyright law: 1790
  - Directly follows Statute of Anne

Copyright – Current State

- Duration: typically the lifetime of the creator plus some years (typically 50-100) or a fixed time if the object was created anonymously or by a corporation

- Exclusive rights
  - to produce copies or reproductions of the work and to sell those copies (including, typically, electronic copies)
  - to import or export the work
  - to create derivative works
  - to perform or display the work publicly
  - to sell or assign these rights to others
  - to transmit or display by radio or video

- Fair use exceptions – no clear definition
  - the purpose and character of your use
  - the nature of the copyrighted work
  - what amount and proportion of the whole work was taken, and
  - the effect of the use upon the potential market for or value of the copyrighted work

- Automatic, no registration needed
  - In some jurisdictions, copyright registration is provided by law
- Holder can enforce and seek damages
  - Enforcement cost borne by the holder

What do we want?

- Authors:
  - Visibility (wide circulation)
  - Easy access
  - Impact (citations)
  - Protection of intellectual property
  - Healthy publishing environment

- Readers/Information consumers
  - Easy & affordable access
  - Trust in the accessed material
  - Sustainability of research archive (long term accessibility)
  - Value-added services (reference links, citation links, etc)

- Society/General public:
  - Easy access
  - Efficient and effective use of public resources
  - Healthy environment for research & discovery

What is a reasonable/good model?

Gold OA/Author pays with author holding the copyright are not it…

- Gold OA: no reasonable business model exists today

- What is the cost of producing a single article in a journal?
  - One study cites £ 7,000
  - High variability in estimates

- Author pays does not scale

- Most of the gold/author pays OA push is from North America and Europe with heavy research funding

Segregate "Publishers"

- Commercial publishers from professional society publishers

- Serious publishers from predatory publishers

Usual warning: These are generalizations; treat with caution.

Fair Access Model – Possible Solution for CS

Professional societies become preferred publishers of choice ➜ We (the members) are the publishers

- Keep production costs low
    - Volunteer effort
- Keep subscription rates low
    - Enable wide access
- Liberal access rights
    - Support Green OA
    - ACM Authorizer service
- Move to a license system (replacing copyright) with exclusive rights granted to publisher
    - Maintain long-term viability of the archive and the society
    - Value-added services built on top of the archive
    - To protect IP rights

Authorizer Service

Conference, Symposia, Workshop Papers

- J. Cheng, Y. Ke, S. Chu, and M. T. Özsu. "Efficient Core Decomposition in Massive Networks," In *Proc. 27th IEEE International Conference on Data Engineering*, 2011, pages 51-62.
  - PDF version
  - Official version
- L. Zou, L. Chen, M. T. Özsu, and D. Zhao."Dynamic Skyline Queries in Large Graphs," In *Proc. 15th Int. Conf. on Database Systems for Advanced Applications* (DASFAA'10), April 2010, Part II, pages 62-78.
  - PDF version
  - Official version
- Mining data streams with periodically changing distributions
  Yingying Tao, M. Tamer Özsu
  CIKM '09 Proceedings of the 18th ACM conference on Information and knowledge management, 2009

  Downloads (6 Weeks): **7**, Downloads (12 Months): **82**, Downloads (cumulative): **310**

- Mining frequent itemsets in time-varying data streams
  Yingying Tao, M. Tamer Özsu
  CIKM '09 Proceedings of the 18th ACM conference on Information and knowledge management, 2009

  Downloads (6 Weeks): **5**, Downloads (12 Months): **49**, Downloads (cumulative): **179**

- Efficient decision tree construction for mining time-varying data streams
  Yingying Tao, M. Tamer Özsu
  CASCON '09 Proceedings of the 2009 Conference of the Center for Advanced Studies on Collaborative Research, 2009

  Downloads (6 Weeks): **3**, Downloads (12 Months): **36**, Downloads (cumulative): **152**

- A partial-order based active cache for recommender systems
  Umar Qasim, Vincent Oria, Yi-Fang Brook Wu, Michael E. Houle, M. Tamer Özsu
  RecSys '09 Proceedings of the third ACM conference on Recommender systems, 2009

  Downloads (6 Weeks): **2**, Downloads (12 Months): **27**, Downloads (cumulative): **162**

- L. Zou, L. Chen, and M. T. Özsu,"Distance-join: Pattern match query in a large graph database," In *Proc. 35th Int. Conf. on Very Large Data Bases* (VLDB'09), August 2009, pages 886-897.
  - PDF version
  - Official version