User Model In a Box: Cross-System User Model Transfer for Resolving Cold Start Problems

Chirayu Wongchokprasitti[1], **Jaakko Peltonen**[2,3],
Tuukka Ruotsalo[2], Payel Bandyopadhyay[4],
Giulio Jacucci[4], Peter Brusilovsky[1]

[1]School of Information Sciences, University of Pittsburgh
[2]Helsinki Institute for Information Technology HIIT, Aalto University
[3]School of Information Sciences, University of Tampere
[4]Helsinki Institute for Information Technology HIIT, Department of Computer Science, University of Helsinki

In Brief:

- We compare methods of **cross-system user model transfer**

In Brief:

- We compare methods of **cross-system user model transfer**

- Two **large real-life systems:**

SciNet for scientific document search

CoMeT for managing scientific talks

In Brief:

- We compare methods of **cross-system user model transfer**

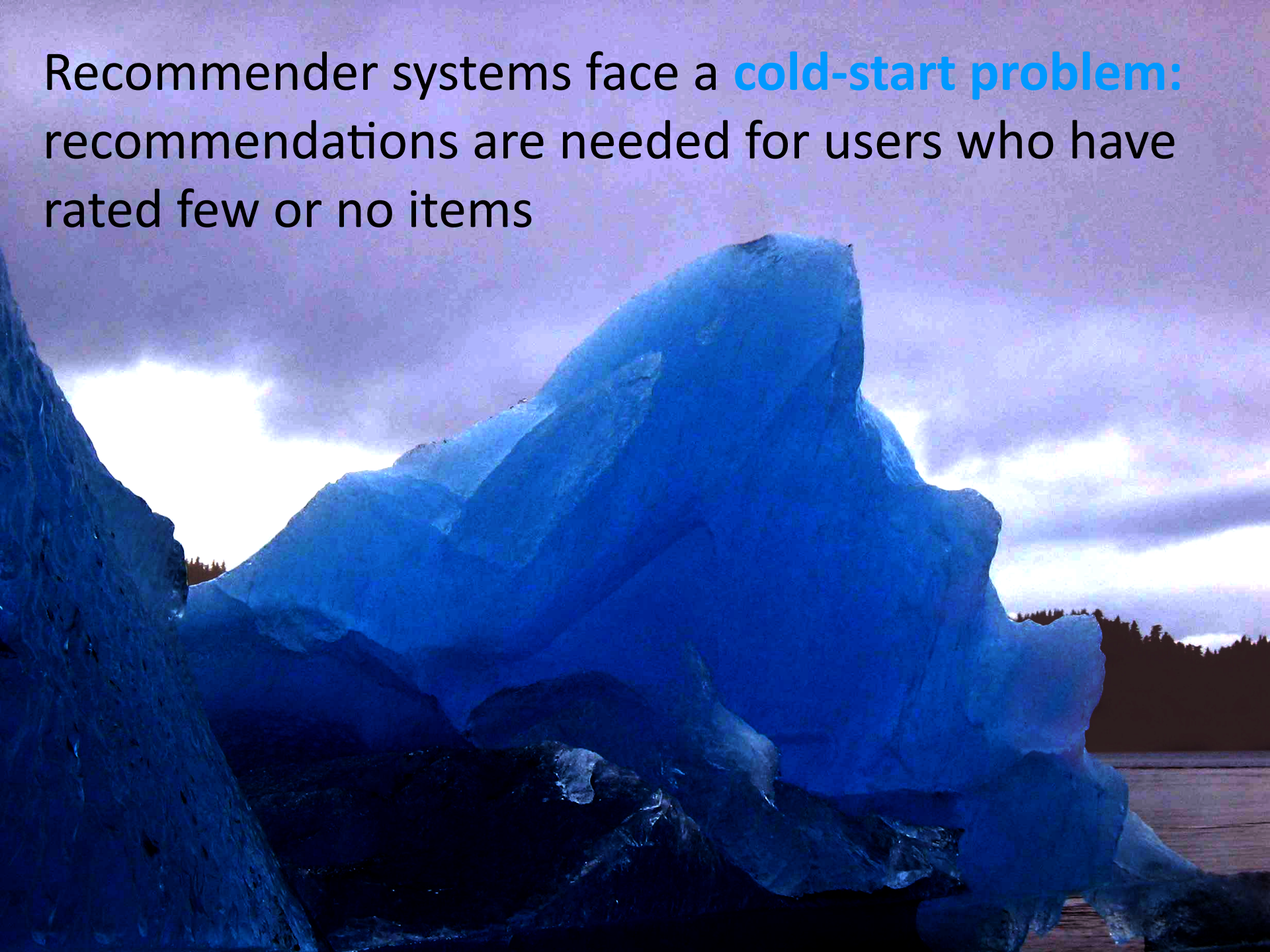
- Two **large real-life systems:**

  SciNet for scientific document search

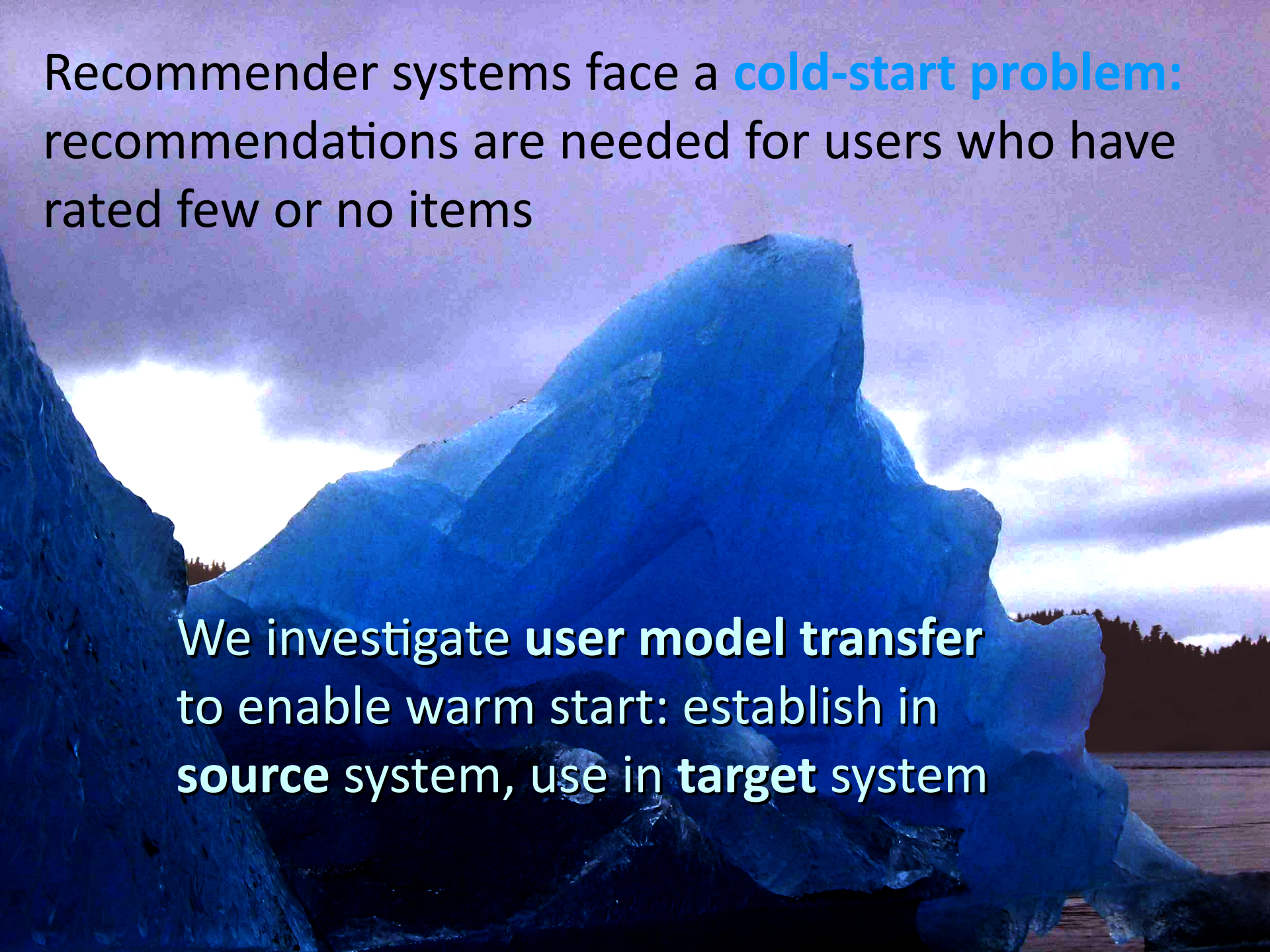
  CoMeT for managing scientific talks

- transfer of novel **explicit open user models** (curated by user during information seeking) strongly improves **cold-start talk recommendation**

Recommender systems face a **cold-start problem:** recommendations are needed for users who have rated few or no items



Recommender systems face a **cold-start problem:** recommendations are needed for users who have rated few or no items



We investigate **user model transfer** to enable warm start: establish in **source** system, use in **target** system

Cross-system/domain recommendation has grown in popularity, but still few studies exploring real information transfer (lack of paired users across systems).

Major focus has been on approaches not assuming common users. Major approaches: collaborative filtering or content-based.

Results mixed, especially content-based has been hard. Focus has been on settings having shared semantic features (social tags, Wikipedia).

We expand earlier research by exploring transferability of **open user models** across related but different domains.

Users of the source system can **explore** and **curate** their model by **visual interaction**.

→ better quality user models, valuable for cross-system transfer

1st work exploring transferability of open user models.

Contributions:

1) cross-system transfer of open user models greatly improves cold-start recommendation

2) we investigate **ways of transferring** open user models, as well as transfer of more traditional implicit and explicit document information.

Open user models bring greatest benefit. We explain it by analysis of cross-system similarities of the different information types.

Academic Information Setting

- Academic users attend **research talks**.
- **A talk management system** can recommend interesting talks given the user's preference.
- Relatively many talks but few bookmarks and ratings (Farzan et al., 2008)
- New users face the cold start problem
- Academic users also **search for scientific documents** in a scientific search system. Can its user model help talk recommendation?

Target system: **CoMeT** system for talk management and recommendation

« Day **Week** Month »

Week 5 of March: March 24 - 30, 2013

Monday, Mar 25

4 bookmarks
44 views

New Algorithms for Nonnegative Matrix Factorization and Beyond **Bookmarked** Unbookmark
By: Ankur Moitra Subscribe , School of Mathematics Institute for Advanced Study Princeton University at: 10:00 AM - 11:30 AM
Location: 6115 Gates and Hillman Centers
Keywords: algorithm factorization machine learning
Posted to groups: Big Data Intelligent Systems Program Machine Learning Group PAWS Group
Bookmarked by: Yun Huang chirayu Peijun Ren zhiqiao

5 bookmarks
74 views

One-Way Mirrors and Weak-Signaling in Online Dating:A Randomized Field Experiment **Recommended**
Bookmark
By: Jui Ramaprasad Subscribe , McGill University, Montreal, Quebec at: 12:00 PM - 1:20 PM
Location: Hamburg Hall 1502
Posted to groups: Human-Centered Computing Social Computing
Bookmarked by: Shenghua Zhang Jie Wang kainan Jin Gan cfan
Series: Tepper IS Seminar **Subscribed** Unsubscribe

21 views

Theoretical Connections Between Convex Optimization and Active Learning **Recommended**
Bookmark
By: Aaditya Ramdas Subscribe , Machine Learning Department at: 12:00 PM - 1:00 PM
Location: Gates Center 6115
Posted to groups: Intelligent Systems Program Machine Learning Group
Series: Machine Learning Lunch seminar at Carnegie Mellon **Subscribed** Unsubscribe

34 views

Higgs Results: Unveiling the Mystery of Mass Bookmark
By: Christoph Paus Subscribe , MIT at: 4:30 PM - 5:30 PM
Location: 7500 Wean Hall, CMU
Series: Physics CMU Seminar Subscribe

Tuesday, Mar 26

55 views

Equating Observed-Scores: The Percentile Rank, Gaussian Kernel, and IRT Observed-Score Equating Methods **Recommended** Bookmark
By: Alina A. von Davier Subscribe , Educational Testing Service at: 9:30 AM - 5:00 PM
Location: Carnegie Mellon University, University Center, Rangos 2

< March 2013 >

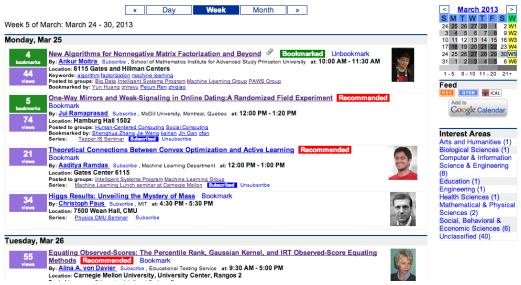
S	M	T	W	T	F	S	W
24	25	26	27	28	1	2	W1
3	4	5	6	7	8	9	W2
10	11	12	13	14	15	16	W3
17	18	19	20	21	22	23	W4
24	25	26	27	28	29	30	W5
31	1	2	3	4	5	6	W6

1 - 5 6 - 10 11 - 20 21+

Feed
RSS ATOM ICAL
Add to Google Calendar

Interest Areas
Arts and Humanities (1)
Biological Sciences (1)
Computer & Information Science & Engineering (8)
Education (1)
Engineering (1)
Health Sciences (1)
Mathematical & Physical Sciences (2)
Social, Behavioral & Economic Sciences (6)
Unclassified (40)

Target system: **CoMeT** system for talk management and recommendation



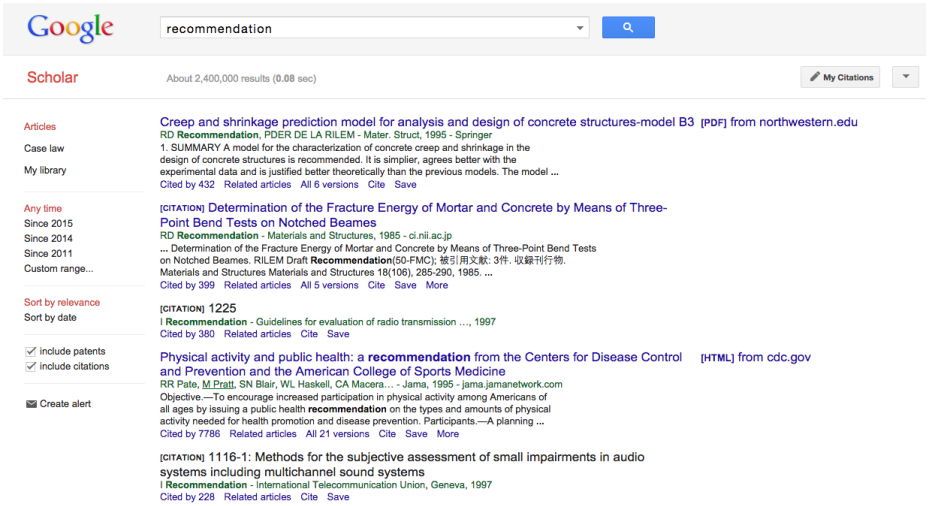
System for sharing information about research talks at Carnegie Mellon University and University of Pittsburgh.

- Available online: halley.exp.sis.pitt.edu/comet/
- Collaborative tagging system: anyone can announce, find, bookmark, and tag talks.
- Has **content-based recommender** - builds interest profile of individual users, recommends new talks to users immediately when posted.

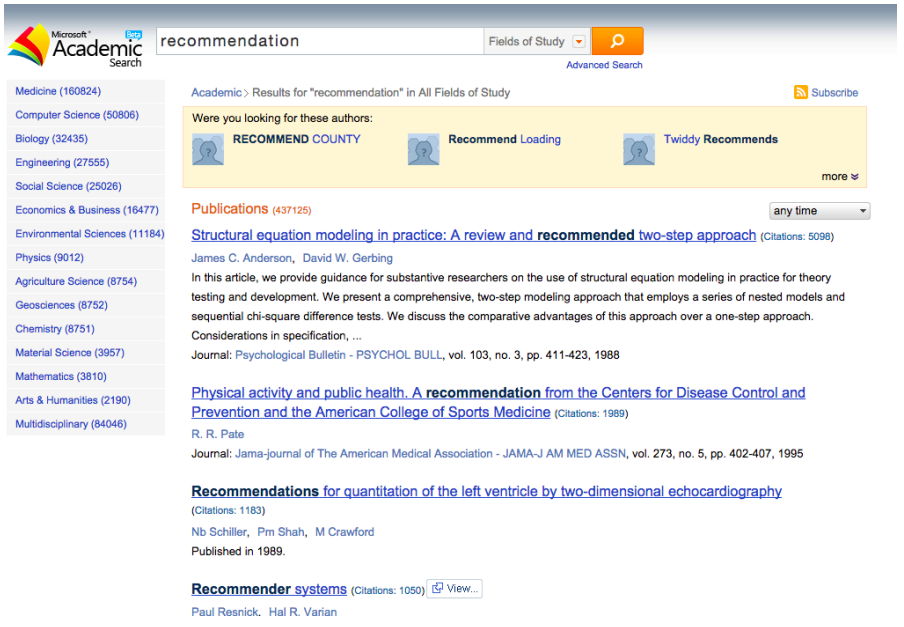
Academic Information Setting

- We use CoMeT as the **target system**
- Academic users also **search for scientific documents** in a scientific search system. Can its user model help talk recommendation?

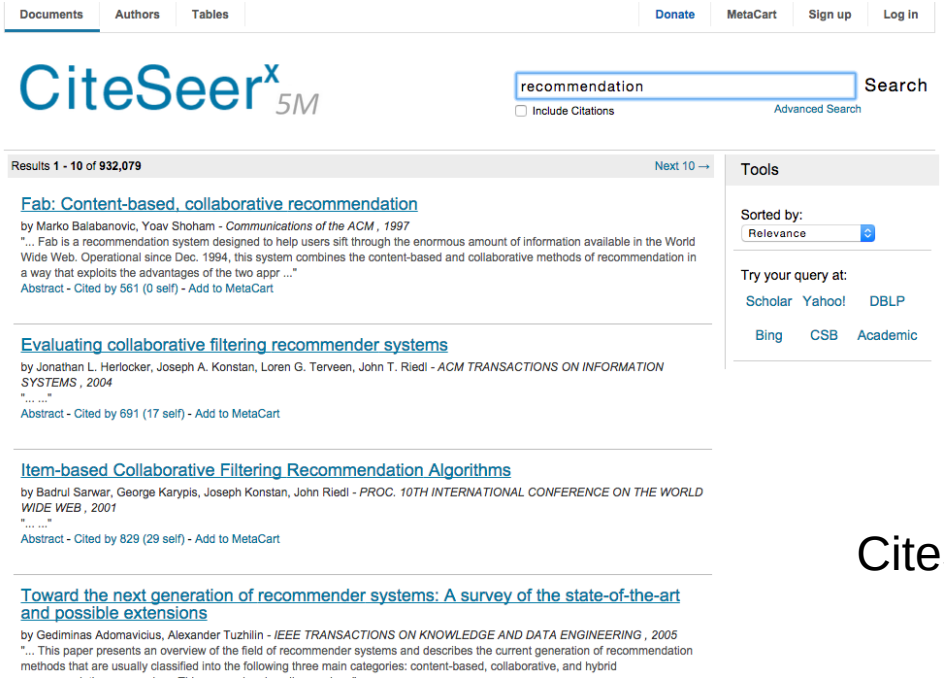
- Unlike traditional search systems (e.g. Google Scholar, Microsoft Academic Search, Citeseer), as the **source system** we use a recent search system having an **open user model**



Google Scholar

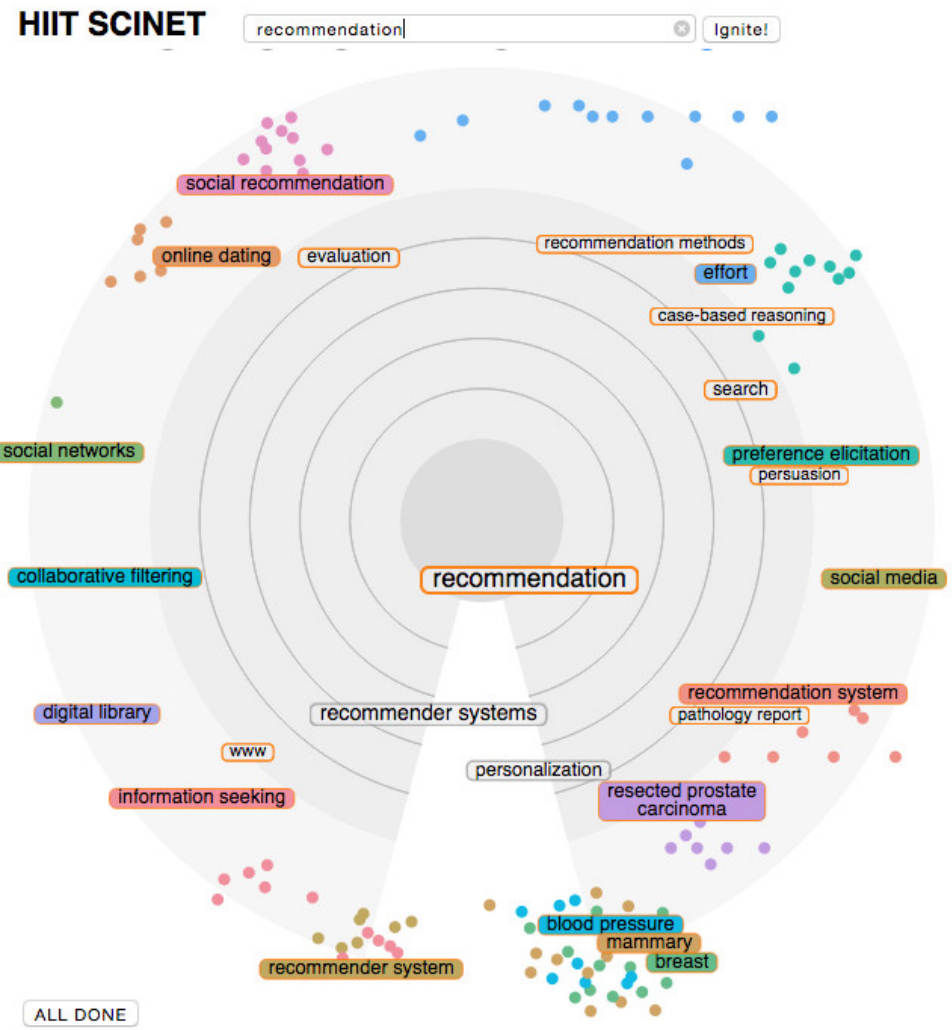


Microsoft Academic Search



CiteSeer

Source system: **SciNet** system for interactive exploratory search of scientific documents



Articles [show bookmarked (0)]

new **Recommendations on recommendations**
Rolf Molich, Kasper Hornbaek, Josephine Scott (Conference on Human Factors in Computing Systems, 2007-01-01)
evaluation | usability comment | usability recommendation | usability test problems | usability testing | user interface | recommendation
`This interactive session discusses the q...`

new **The Universal Recommender**
Jérôme Kunegis, Alan Said, Winfried Umbrath (arXiv.org, 2009-01-01)
recommender systems | recommendation
`We describe the Universal Recommender, a...`

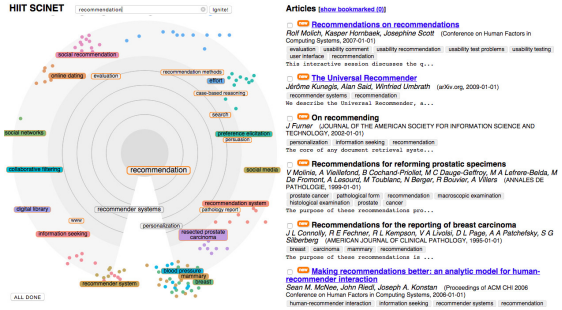
new **On recommending**
J Furner (JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 2002-01-01)
personalization | information seeking | recommendation
`The core of any document retrieval syste...`

new **Recommendations for reforming prostatic specimens**
V Molinie, A Vieillefond, B Cochand-Priollet, M C Dauge-Geffroy, M A Lefrere-Belda, M De Fromont, A Lesourd, M Toublanc, N Berger, R Bouvier, A Villers (ANNALES DE PATHOLOGIE, 1999-01-01)
prostate cancer | pathological form | recommendation | macroscopic examination | histological examination | prostate | cancer
`The purpose of these recommendations pro...`

new **Recommendations for the reporting of breast carcinoma**
J L Connolly, R E Fechner, R L Kempson, V A Livolsi, D L Page, A A Patchefsky, S G Silberberg (AMERICAN JOURNAL OF CLINICAL PATHOLOGY, 1995-01-01)
breast | carcinoma | mammary | recommendation
`The purpose of these recommendations is ...`

new **Making recommendations better: an analytic model for human-recommender interaction**
Sean M. McNee, John Riedl, Joseph A. Konstan (Proceedings of ACM CHI 2006 Conference on Human Factors in Computing Systems, 2006-01-01)
human-recommender interaction | information seeking | recommender systems | recommendation

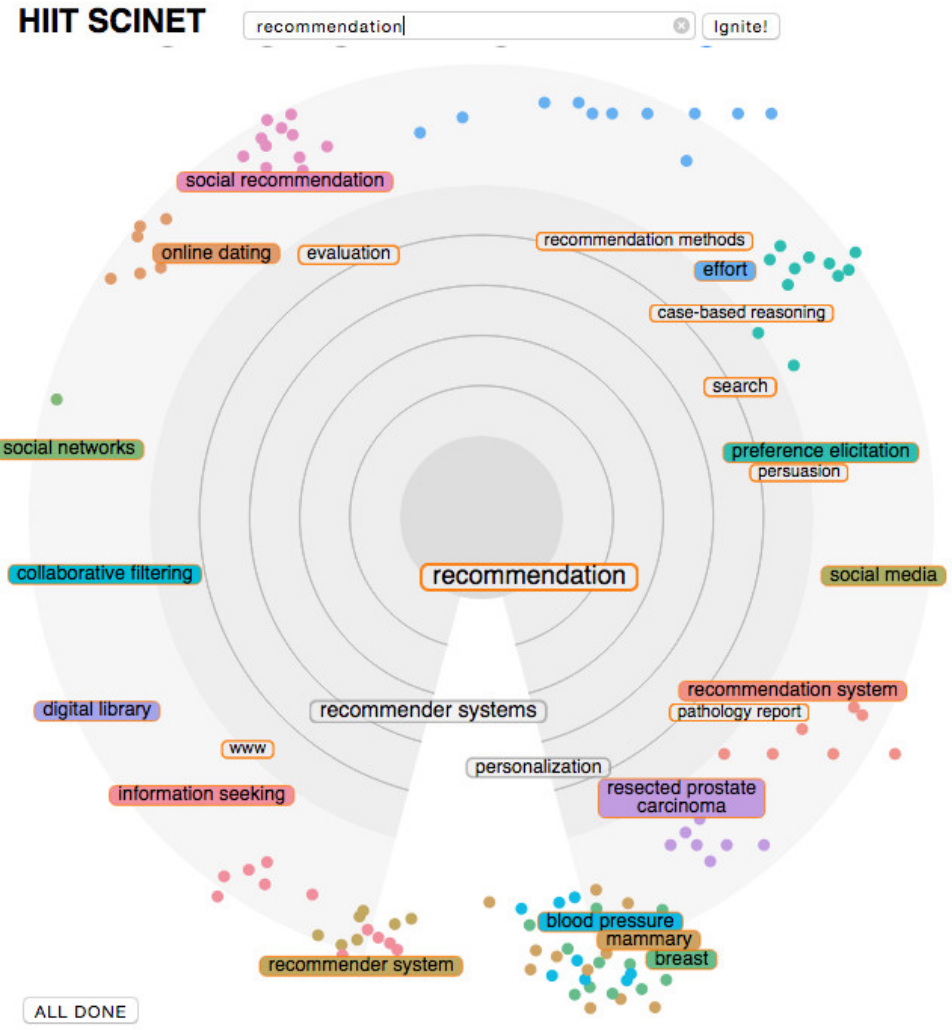
Source system: **SciNet** system for interactive exploratory search of scientific documents



Exploratory search system. Indexes over 50m scientific documents from Thomson Reuters, ACM, IEEE & Springer

- Goes beyond text-based queries.
- User can direct exploratory search by interacting with an open user model.
- Significantly improves information seeking task performance and quality of retrieved information.
- Open user models are promising for cross-system transfer.

SciNet **opens its user model:** users can interact with a **visualization** of the model, and curate the model by feedback.



Articles [show bookmarked (0)]

new **Recommendations on recommendations**
Rolf Molich, Kasper Hornbaek, Josephine Scott (Conference on Human Factors in Computing Systems, 2007-01-01)
evaluation | usability comment | usability recommendation | usability test problems | usability testing | user interface | recommendation
`This interactive session discusses the q...`

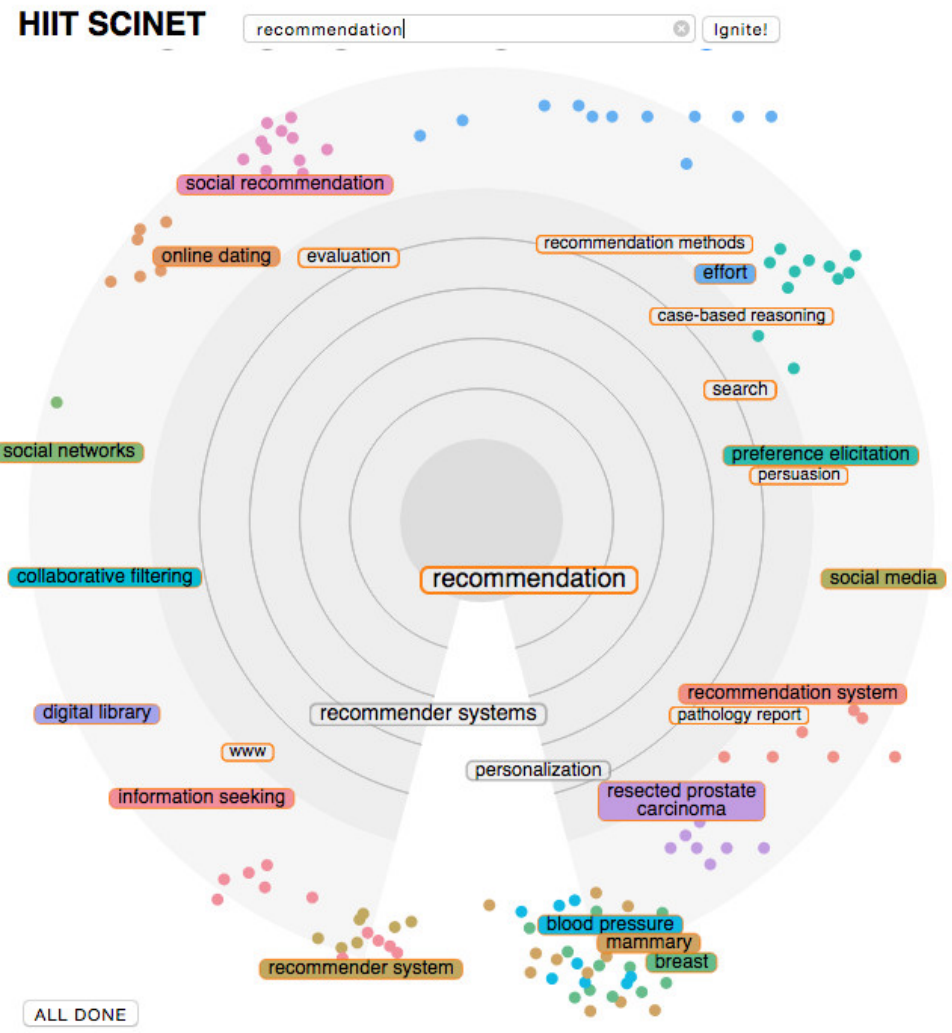
new **The Universal Recommender**
Jérôme Kunegis, Alan Said, Winfried Umbrath (arXiv.org, 2009-01-01)
recommender systems | recommendation
`We describe the Universal Recommender, a...`

new **On recommending**
J Furner (JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 2002-01-01)
personalization | information seeking | recommendation
`The core of any document retrieval syste...`

new **Recommendations for reforming prostatic specimens**
V Molinie, A Vieillefond, B Cochand-Priollet, M C Dauge-Geffroy, M A Lefrere-Belda, M De Fromont, A Lesourd, M Toublanc, N Berger, R Bouvier, A Villers (ANNALES DE PATHOLOGIE, 1999-01-01)
prostate cancer | pathological form | recommendation | macroscopic examination | histological examination | prostate | cancer
`The purpose of these recommendations pro...`

new **Recommendations for the reporting of breast carcinoma**
J L Connolly, R E Fechner, R L Kempson, V A Livolsi, D L Page, A A Patchefsky, S G Silberberg (AMERICAN JOURNAL OF CLINICAL PATHOLOGY, 1995-01-01)
breast | carcinoma | mammary | recommendation
`The purpose of these recommendations is ...`

new **Making recommendations better: an analytic model for human-recommender interaction**
Sean M. McNee, John Riedl, Joseph A. Konstan (Proceedings of ACM CHI 2006 Conference on Human Factors in Computing Systems, 2006-01-01)
human-recommender interaction | information seeking | recommender systems | recommendation

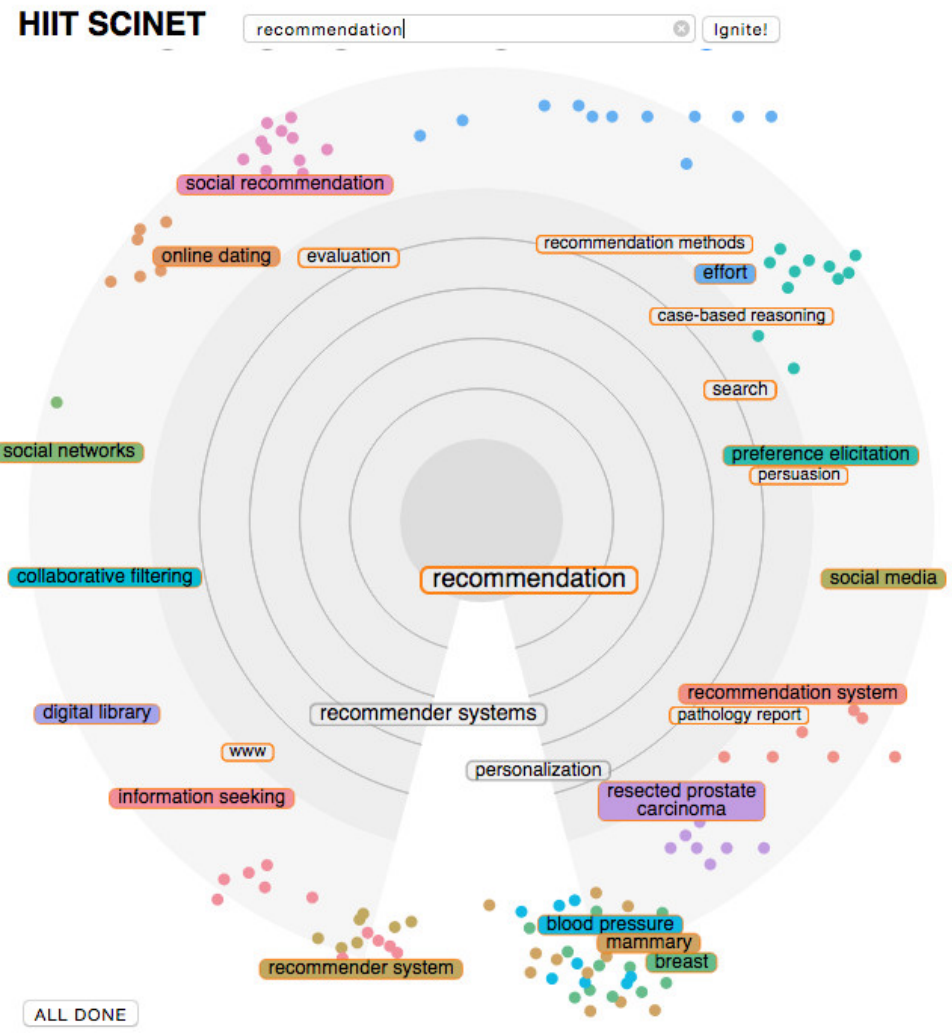


User intent modeled as vector of interests over keywords.

Predicted interests visualized, organized by correlation of interest wrt. future feedback.

Users can drag keywords to change their importance.

A new model is inferred from user feedback. Documents and keyword display change.



Technically:

- model is inferred by Bayesian inference of a linear predictor of user interest over keywords, fitted to feedback.
- Exploration-exploitation is modeled by taking upper confidence bounds of keyword relevances.
- Keywords visualized by dimensionality reduction of their relevances across future feedback.

Ways of Transferring a User Model

Our interest is to use

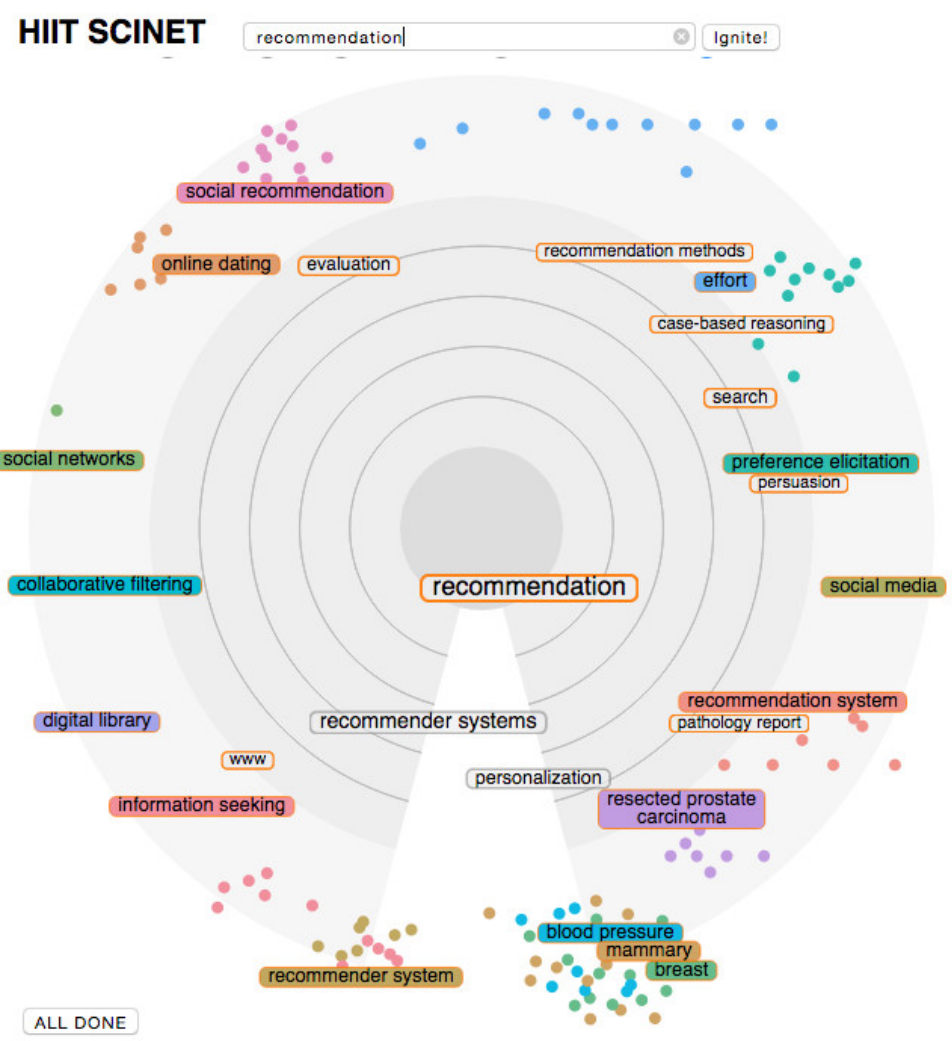
1. the **whole content of the open user model**

2. its **curated subset** (the keywords the user moved in the process of curation).

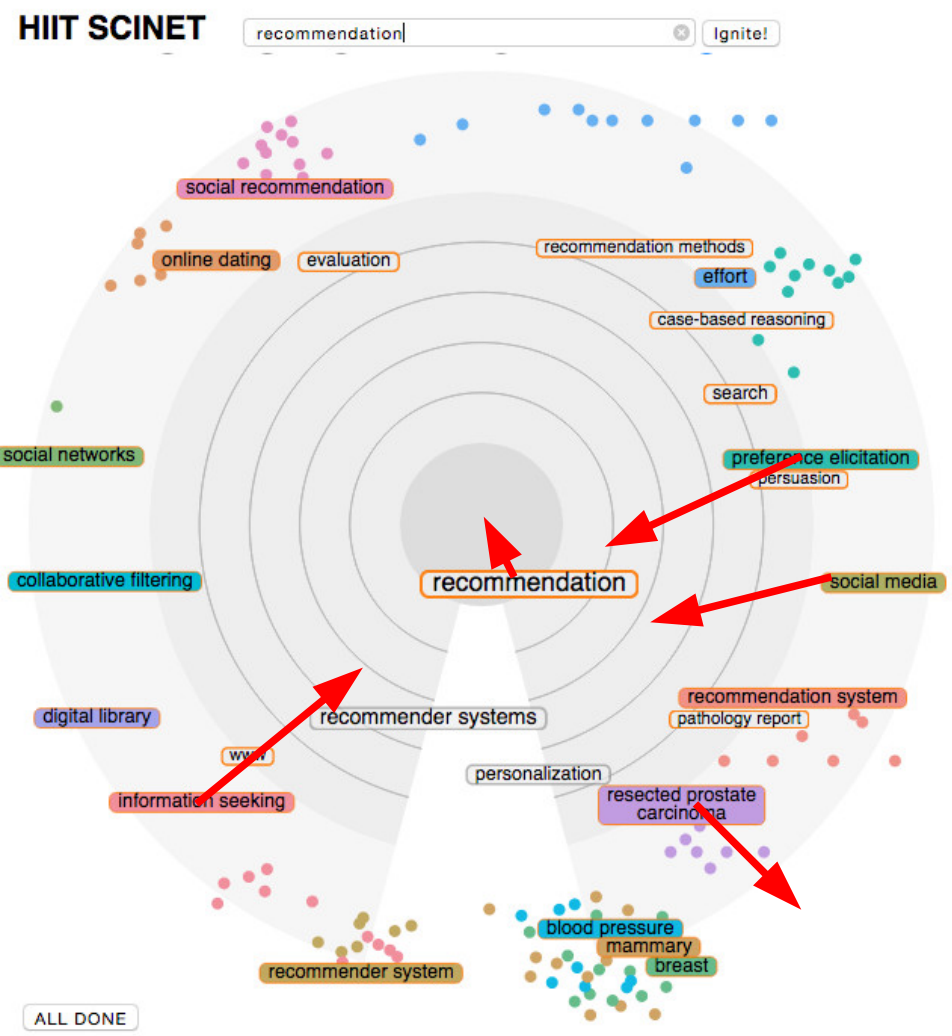
As a baseline, we also explore transfer of:

3. the **relevant documents** selected by the user during search (could be considered a hidden, implicit user model)

4. a broader set of all **documents retrieved** in response to user queries (weaker reflection of user interests)


HIIT SCINET
recommendation
Ignite!
social recommendation
online dating
evaluation
recommendation methods
effort
case-based reasoning
search
social networks
preference elicitation
persuasion
collaborative filtering
recommendation
social media
digital library
recommender systems
recommendation system
pathology report
www
personalization
information seeking
resected prostate carcinoma
blood pressure
mammary
breast
recommender system
ALL DONE

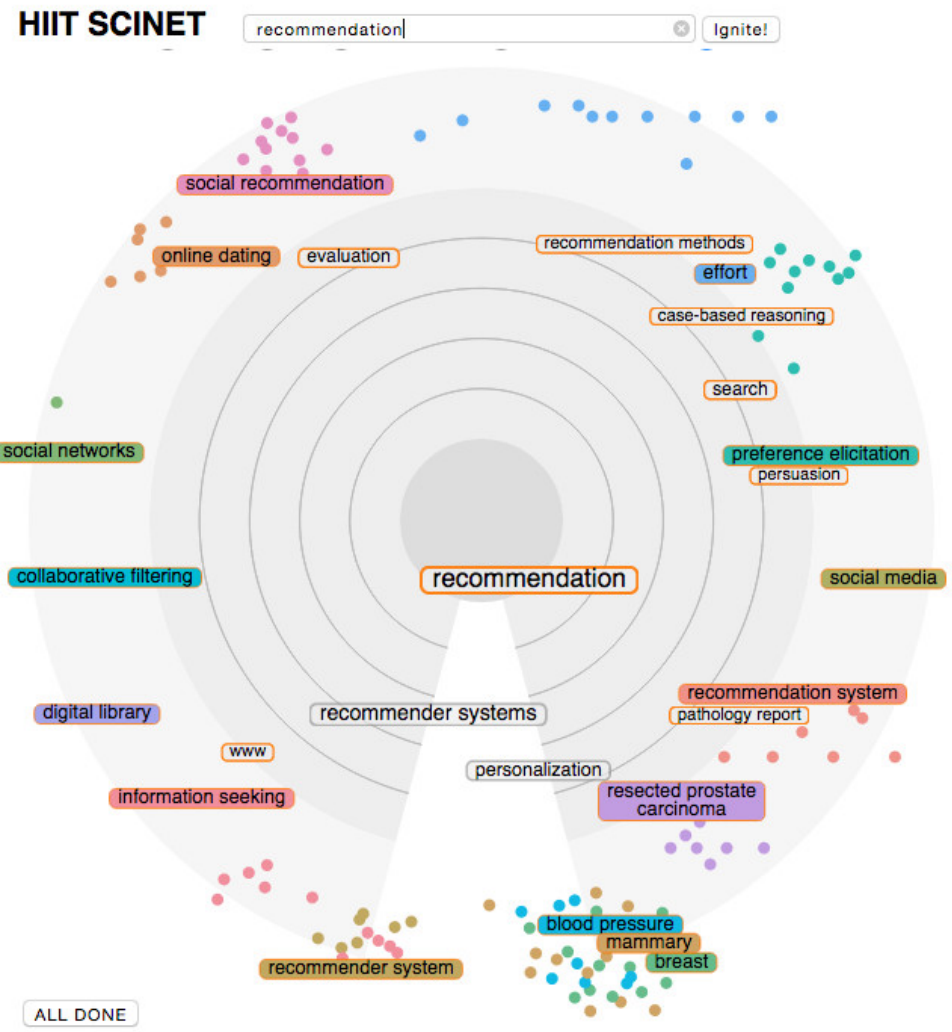
Explicit model 1: manipulated keywords



Take keywords **manipulated** by users, with their **assigned** relevances, convert to unigrams, and form a pseudo-document (bookmarked talk abstract) from them.

(discard unigrams not occurring in target system)

Explicit model 2: shown keywords



At each iteration,
Take keywords
seen by users,
with their **predicted** relevances, convert to unigrams, and form a pseudo-document (bookmarked talk abstract) from them.

(discard unigrams not occurring in target system)

Articles [show bookmarked (0)]

new **Recommendations on recommendations**
Rolf Molich, Kasper Hornbaek, Josephine Scott (Conference on Human Factors in Computing Systems, 2007-01-01)
evaluation | usability comment | usability recommendation | usability test problems | usability testing | user interface | recommendation
`This interactive session discusses the q...`

new **The Universal Recommender**
Jérôme Kunegis, Alan Said, Winfried Umbrath (arXiv.org, 2009-01-01)
recommender systems | recommendation
`We describe the Universal Recommender, a...`

new **On recommending**
J Furner (JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 2002-01-01)
personalization | information seeking | recommendation
`The core of any document retrieval syste...`

new **Recommendations for reforming prostatic specimens**
V Molinie, A Vieillefond, B Cochand-Priollet, M C Dauge-Geffroy, M A Lefrere-Belda, M De Fromont, A Lesourd, M Toublanc, N Berger, R Bouvier, A Villers (ANNALES DE PATHOLOGIE, 1999-01-01)
prostate cancer | pathological form | recommendation | macroscopic examination | histological examination | prostate | cancer
`The purpose of these recommendations pro...`

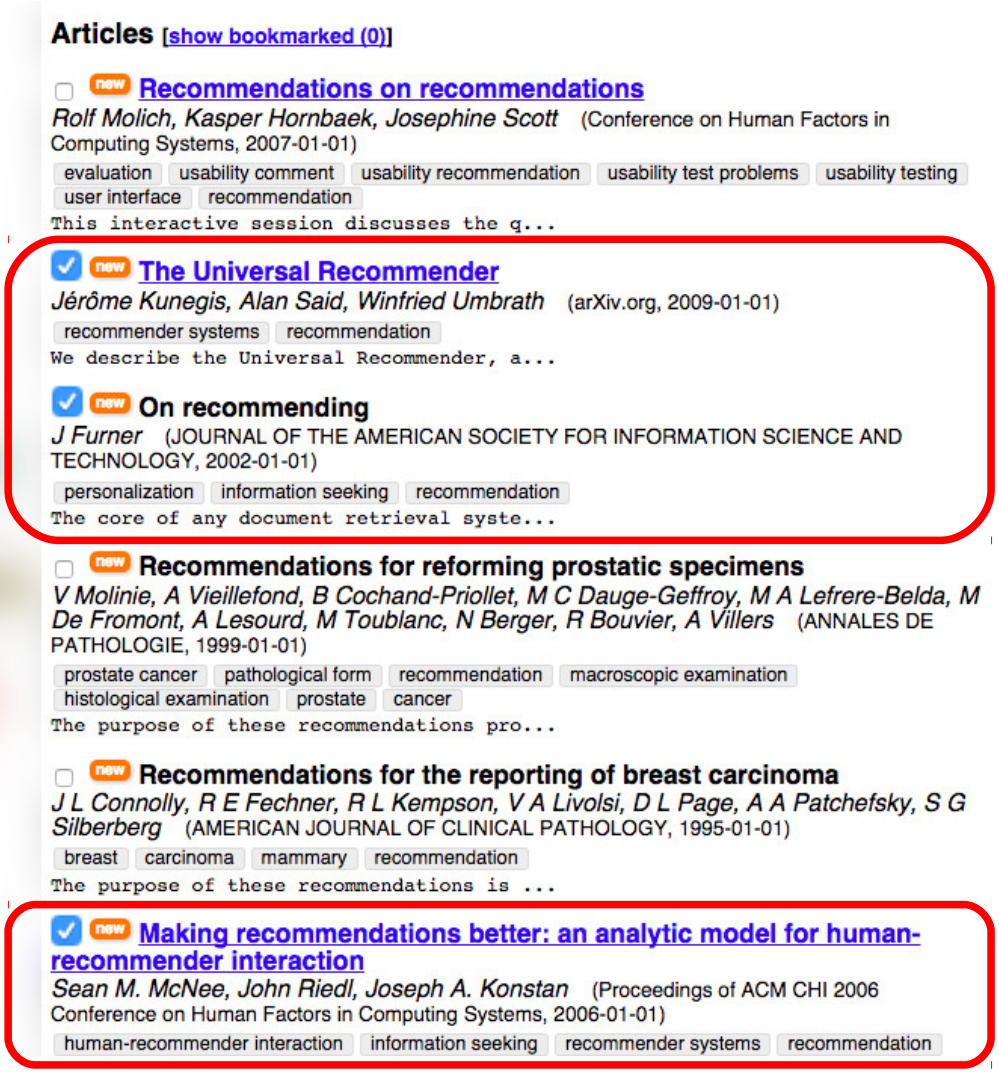
new **Recommendations for the reporting of breast carcinoma**
J L Connolly, R E Fechner, R L Kempson, V A Livolsi, D L Page, A A Patchefsky, S G Silberberg (AMERICAN JOURNAL OF CLINICAL PATHOLOGY, 1995-01-01)
breast | carcinoma | mammary | recommendation
`The purpose of these recommendations is ...`

new **Making recommendations better: an analytic model for human-recommender interaction**
Sean M. McNee, John Riedl, Joseph A. Konstan (Proceedings of ACM CHI 2006 Conference on Human Factors in Computing Systems, 2006-01-01)
human-recommender interaction | information seeking | recommender systems | recommendation

Implicit model 1: bookmarked documents

Scientific documents **bookmarked** by the user during the search session are implicit information about user interests.

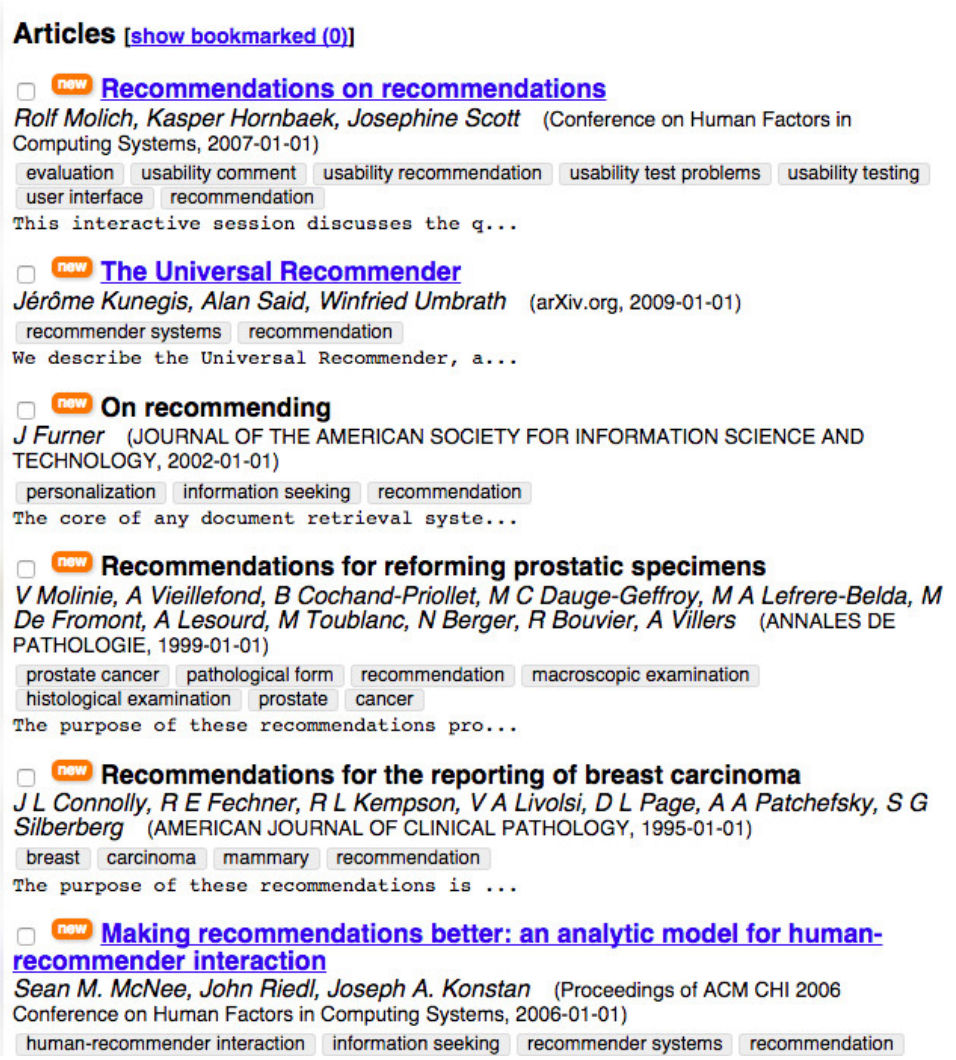
Convert each into unigrams, add into CoMeT as a bookmarked talk.



Implicit model 2: seen documents

Scientific documents **seen** by the user during the search session are implicit information about user interests (momentary responses to user search).

Convert each into unigrams, add into CoMeT as a bookmarked talk.



Articles [show bookmarked (0)]

new **Recommendations on recommendations**
Rolf Molich, Kasper Hornbaek, Josephine Scott (Conference on Human Factors in Computing Systems, 2007-01-01)
evaluation | usability comment | usability recommendation | usability test problems | usability testing | user interface | recommendation
`This interactive session discusses the q...`

new **The Universal Recommender**
Jérôme Kunegis, Alan Said, Winfried Umbrath (arXiv.org, 2009-01-01)
recommender systems | recommendation
`We describe the Universal Recommender, a...`

new **On recommending**
J Furner (JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 2002-01-01)
personalization | information seeking | recommendation
`The core of any document retrieval syste...`

new **Recommendations for reforming prostatic specimens**
V Molinie, A Vieillefond, B Cochand-Priollet, M C Dauge-Geffroy, M A Lefrere-Belda, M De Fromont, A Lesourd, M Toublanc, N Berger, R Bouvier, A Villers (ANNALES DE PATHOLOGIE, 1999-01-01)
prostate cancer | pathological form | recommendation | macroscopic examination | histological examination | prostate | cancer
`The purpose of these recommendations pro...`

new **Recommendations for the reporting of breast carcinoma**
J L Connolly, R E Fechner, R L Kempson, V A Livolsi, D L Page, A A Patchefsky, S G Silberberg (AMERICAN JOURNAL OF CLINICAL PATHOLOGY, 1995-01-01)
breast | carcinoma | mammary | recommendation
`The purpose of these recommendations is ...`

new **Making recommendations better: an analytic model for human-recommender interaction**
Sean M. McNee, John Riedl, Joseph A. Konstan (Proceedings of ACM CHI 2006 Conference on Human Factors in Computing Systems, 2006-01-01)
human-recommender interaction | information seeking | recommender systems | recommendation

Experiments

- 20 researchers from University of Helsinki:
  14 male, 6 female; 10 PhD researchers and 10 research assistants
- SciNet: Search relevant papers to their interest
  - "Write down 3 areas of your research interests. Imagine you are preparing for a course/seminar for each interest. Search scientific documents you find useful for preparing for the courses/seminars."
  - Bookmark at least 5 documents for each interest.
  - 7min demonstration, 30min for task
  - Complex enough: users must interact with the system to get needed information. Broad enough to reveal research interests.
- CoMeT: Rate 500 talks (Jan 1 to May 17, 2013)
  - Consider attending (Yes/No)? If yes, rate willingness 1 – 5
  - 7min demonstration, 75min for task
- All interactions logged (shown/manipulated keywords shown/bookmarked documents, queries, read abstracts...)

Non-cold-start Setting

- We first evaluated a traditional non-cold-start learning setting
- 10-fold cross-validation setup, in each fold rank the held-out CoMeT talks by 3 predictors
- **Centroid:** rank test talks by cosine similarity to centroid of bookmarked talks
- **k-Nearest-Neighbor:** find nearest training neighbors for each test talk, rank by s_{pos}-s_{neg} (sum of cosine similarities to positive nearest neighbors - sum of cosine similarities to negative neighbors)
- **positive-only kNN:** find nearest positive-rated talks, rank by sum of cosine similarity to them

Non-cold-start Setting

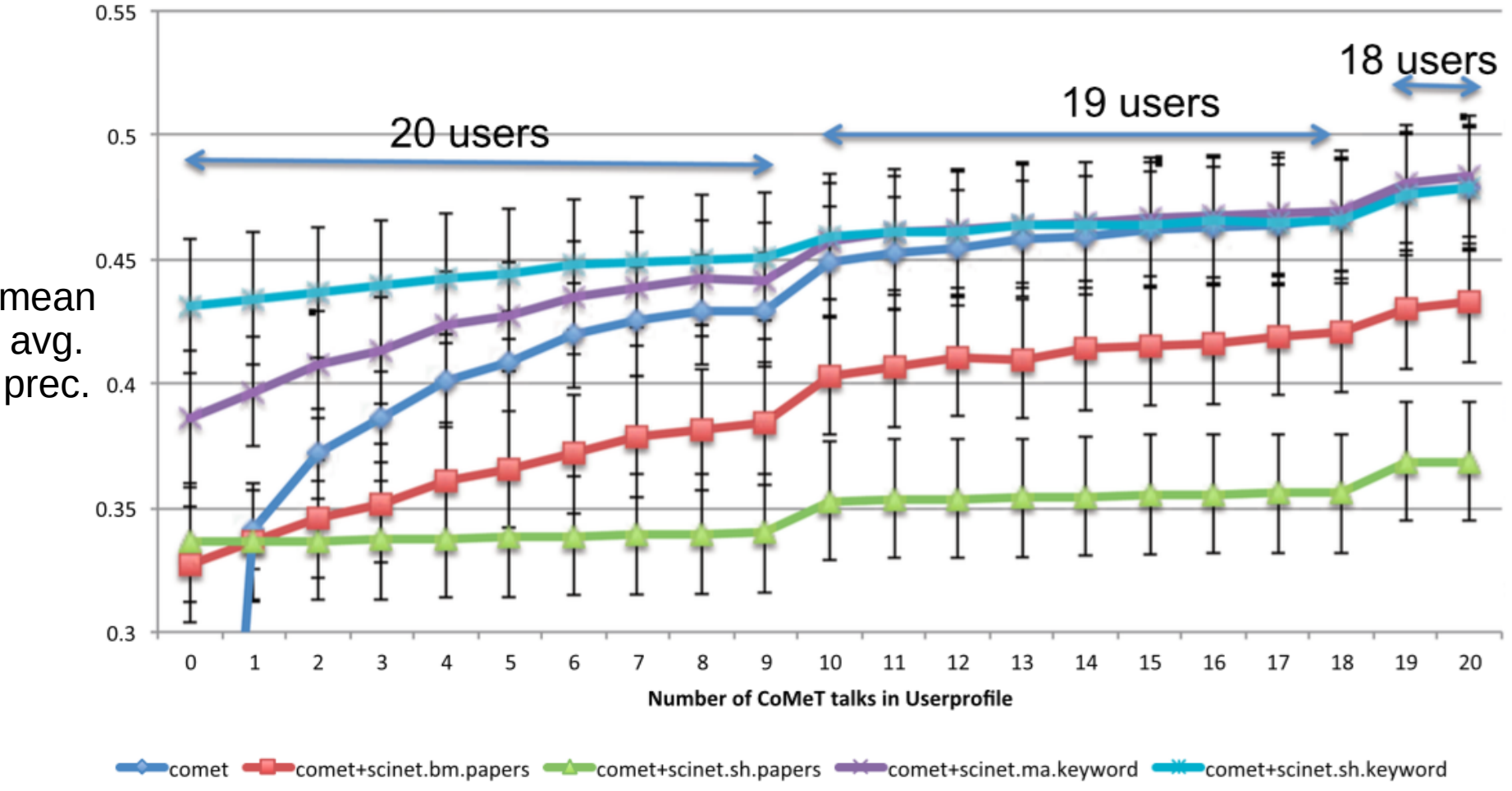
- Results evaluated by Mean Average Precision of ranked test talks
  (mean of precision values at positive test talks in the ranking, averaged over users and folds)
- no significant improvement from transfer compared to baseline in non-cold-start setting from traditional or open-user-model approach
- User profiles in CoMeT had enough data to work well on their own

Mean Average Precision		Centroid	k-NN				k-NN.PO			
			5nn	10nn	20nn	30nn	5nn.po	10nn.po	20nn.po	30nn.po
baseline		0.47	0.45	0.47	0.48	0.46	0.48	0.50	0.50	0.50
Implicit User Model	ex.papers	0.42	0.44	0.45	0.45	0.44	0.43	0.44	0.44	0.44
	im.papers	0.36	0.36	0.36	0.35	0.35	0.36	0.37	0.36	0.36
Explicit Open User Model	ex.keywords	0.48	0.46	0.48	0.48	0.47	0.49	0.51	0.51	0.50
	im.keywords	0.47	0.46	0.48	0.49	0.49	0.48	0.49	0.49	0.48

Cold-start Setting

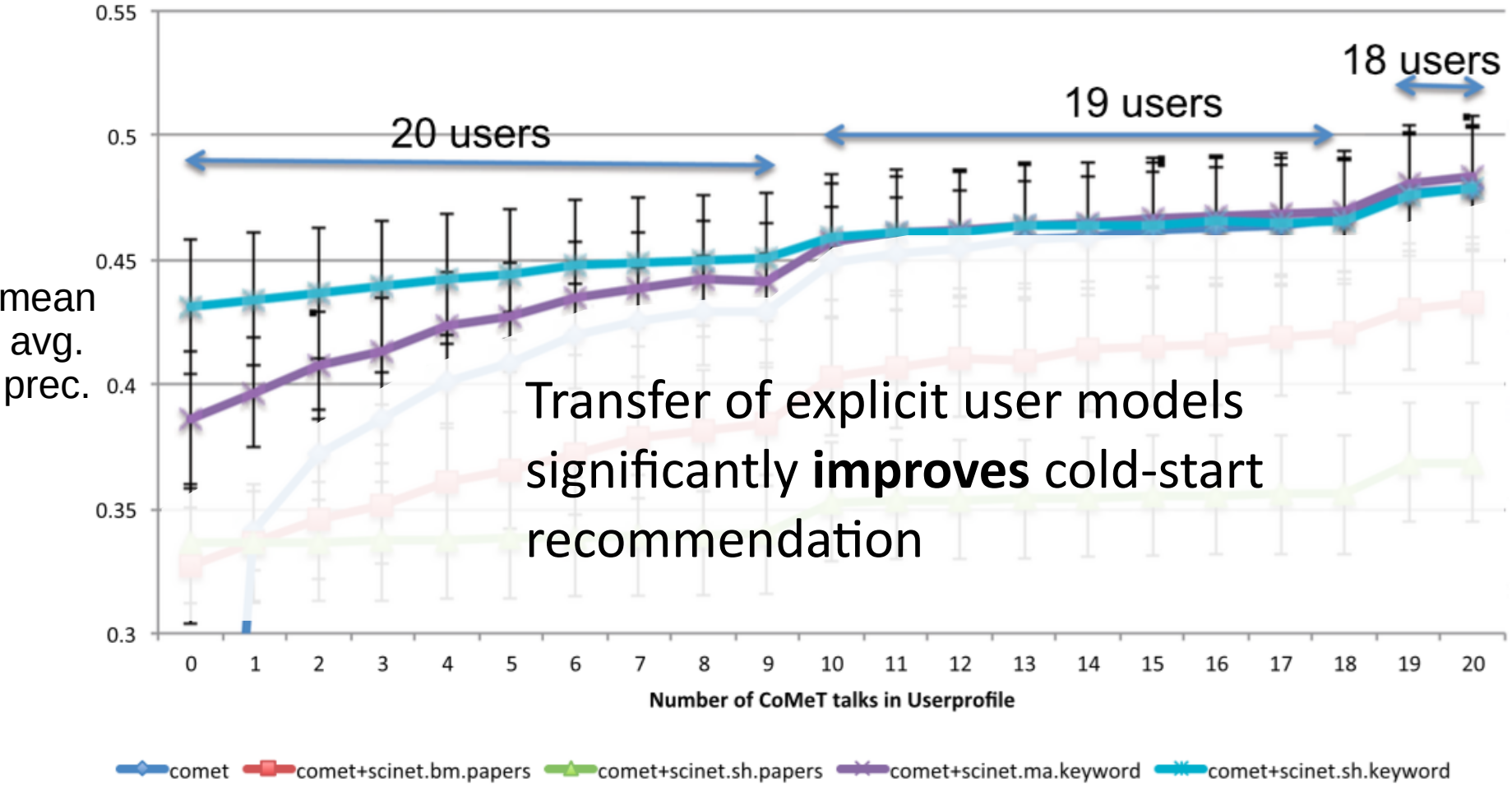
- In each cross-validation fold we subsample a small pool of cold-start talks (0-20 positive talks, proportionally same amont of negative talks)
- Cold-start talks used to predict test talk ranking, evaluate by mean average precision
- We report average results over 10 subsamplings
- Same predictors as before (Centroid, k-Nearest-Neighbor, positive-only kNN)

Cold-Start Impact



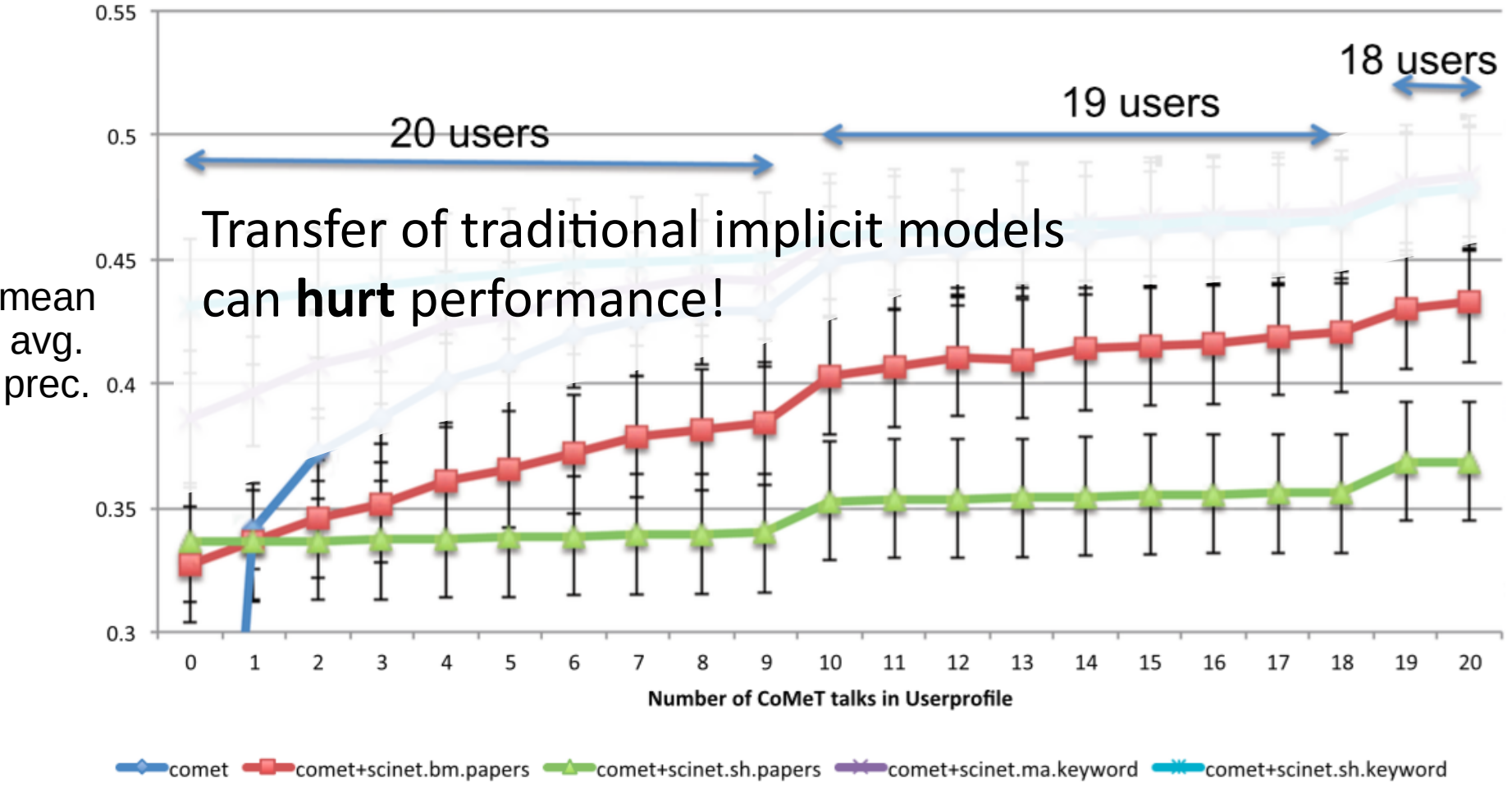
Results with centroid predictor shown: positive-only kNN performs essentially the same, and outperforms kNN

Cold-Start Impact



Results with centroid predictor shown: positive-only kNN performs essentially the same, and outperforms kNN

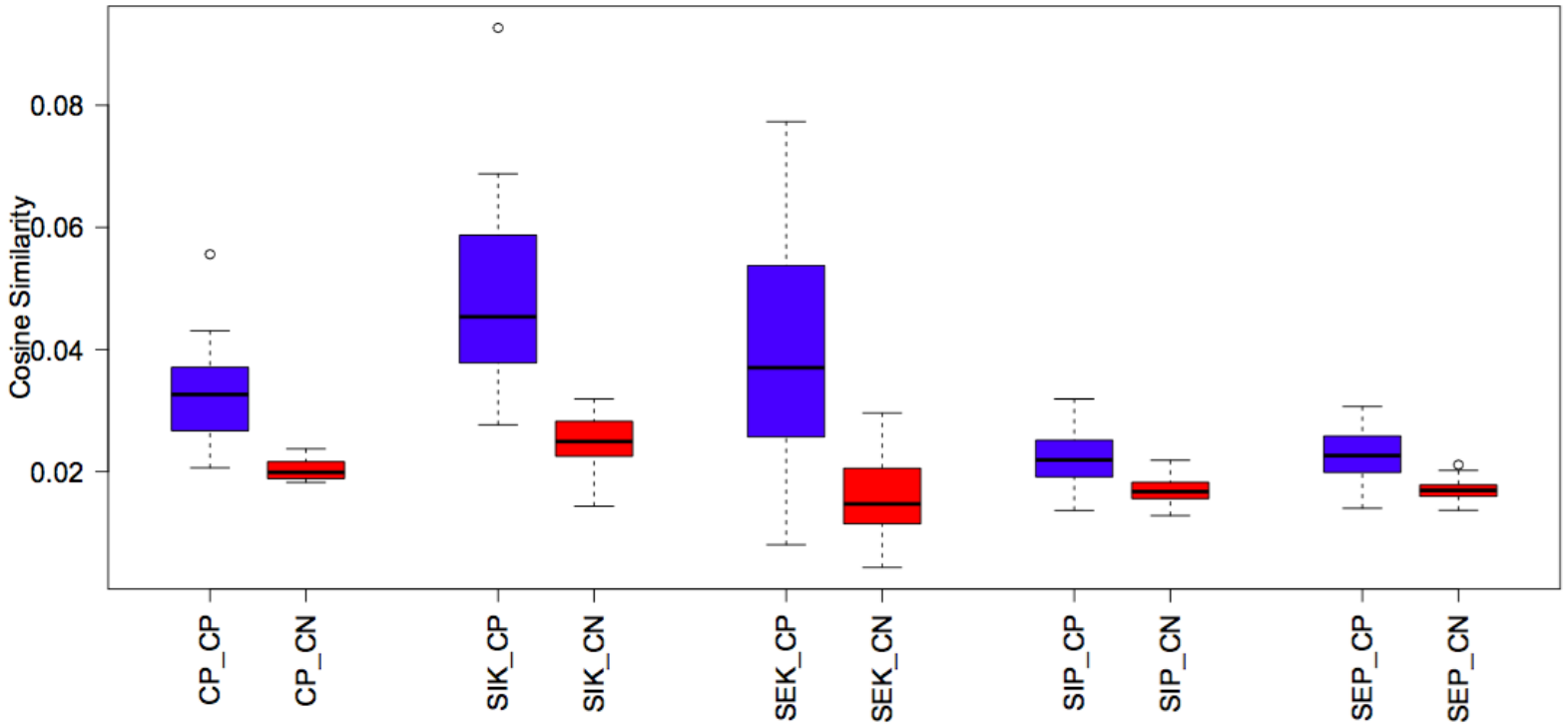
Cold-Start Impact



Results with centroid predictor shown: positive-only kNN performs essentially the same, and outperforms kNN

Analysis

- Cosine similarities between different information types



Explicit open user models have good similarity to positive-rated talks, well separated from uninteresting talks

Implicit models from papers are far from bookmarked talks. They do not separate positive-rated from uninteresting talks.
→ add more noise than value

Summary

- Cross-system personalization by transferring an **explicit, open, and editable user model.**
- Transfer from a literature search system to a talk recommendation system.
- Cross-system model transfer is challenging: no impact in general case
- However, **significant impact in cold-start case!**
- **Use of open, explicitly curated user models is critical** for the success of user model transfer
- Transferring implicit models (here through shown or bookmarked documents) can damage performance