



Information-seeking on the Web with Trusted Social Networks - from Theory to Systems

Tom Heath

Knowledge Media Institute, The Open University
/ Platform Division, Talis



1. Problem Statement and Research Questions
2. Source Selection in Social Networks
3. Technical Approach
4. Data Sources: Revyu.com and Beyond
5. Hoonoh Trust Algorithms and Hoonoh.com
6. Conclusions and Future Work



Problem Statement and Research Questions



Problem

How do you find information that's relevant to you personally?



- Questions that are easy to answer, but hard to get 'right':
 - “hotel in paris”
 - “plumber in milton keynes”
 - “back pain specialist”
 - ...etc...



Bill Gates and I need different search results







Google™

[Web](#) [Images](#) [Groups](#) [News](#) [Froogle](#) [more »](#)

hotel paris luxurious discreet "excellent service"

Search: ☒ the web ☐ pages from the UK





Google™

[Web](#) [Images](#) [Groups](#) [News](#) [Froogle](#) [more »](#)

hotel paris quirky alternative cheap

Search

Search: ☒ the web ☐ pages from the UK



- Information Overload!
 - Bill and I want different information from the same query
- Keywords aren't expressive enough
 - It's hard to convey tastes or preferences to personalise a search query



Wanted!

- Some means to:
 - constrain the search space
 - prioritise results
 - identify the right information **for you**



Social networks and word of mouth recommendation are the answer!





"Many information gathering tasks are better handled by finding a referral to a human expert rather than by simply interacting with online information sources"

(Kautz, Selman and Shah, 1997)



My Approach

- Better exploit existing social processes to support information-seeking on the Web
- Provide personalised relevance in information-seeking through your trusted social network



- Characteristics
 - Source-centricity
 - Task-adaptivity
- Benefits
 - Increased personal relevance
 - Spam resistance
 - More complex trust judgements
 - Openness to additional (offline) information



Research Questions

1. How do people choose information and recommendation sources from among members of their social network?
2. Which factors influence judgements about the relevance and trustworthiness of these information and recommendation sources?
3. How do the characteristics of the task being performed affect these judgements?



Research Questions

4. To what extent can general principles derived from answers to the previous questions be operationalised as computational algorithms that replicate the process of seeking information and recommendations through social networks? (*can we operationalise these principles algorithmically?*)
5. How feasible is the implementation of user-oriented systems that exploit such algorithms? (*can we implement systems based on these algorithms?*)
6. If such systems can be implemented, how do they perform relative to human performance of equivalent tasks?



Source Selection in Social Networks



my friends know loads of stuff...





...but who knows what, exactly?





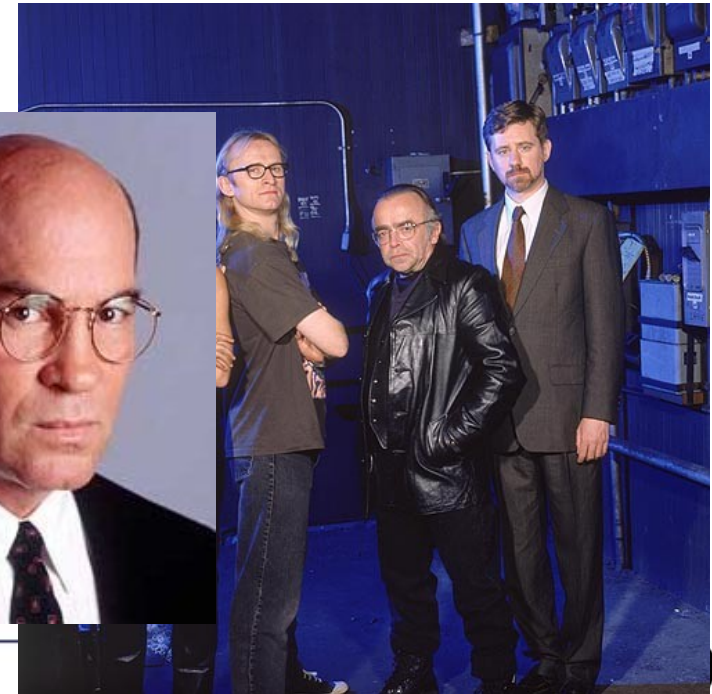
...and who is the best person to ask?





Who should Fox Mulder trust for restaurant recommendations?

trust no one





- Study of Source Selection in Word-of-Mouth Information-Seeking
 - Exploratory, qualitative study of how people choose information and recommendation sources
 - Questions
 - Who do people seek recommendations from in *different scenarios*?
 - How do they *decide whether or not to trust* this information?



- In depth interviews with 12 participants
- 4 recommendation seeking scenarios
 - plumber, hotel, back pain, holiday activities
 - variation by task modality and criticality
 - “who would you ask for recommendations, and why”
- Qualitative analysis to identify key themes



5 Trust Factors in Word of Mouth Recommendation



The source has relevant expertise, which may be formally validated through qualifications or acquired over time (35)

“I would probably go and ask my friend who is a plumber or my friend who is a gas fitter, working on the principle that their domain expertise, their knowledge, is in a similar area”



The source has experience of solving similar scenarios, but without extensive expertise (41)

“People i know in the area, it’s good to have word of mouth, you know they’ve got experience good or bad”



The source does not have vested interests in a particular resolution to the scenario (9)

“With travel agents you’d have to question what they were promoting to you - is it because they get commission?”



The source has characteristics in common with the recommendation seeker such as shared tastes, standards, viewpoints, interests, or expectations (24)

“[I] may not ask people who I don’t feel comfortable with, who haven’t got the same values as me, or have a completely different lifestyle that I don’t relate to”



The source has previously provided successful recommendations to the recommendation seeker (3)

“I looked on the internet yesterday about going to see a masseur, but they were too expensive so I’ll go back to [ask] my sister as I had a good experience with [recommendations from] her before”



How the Factors are Used



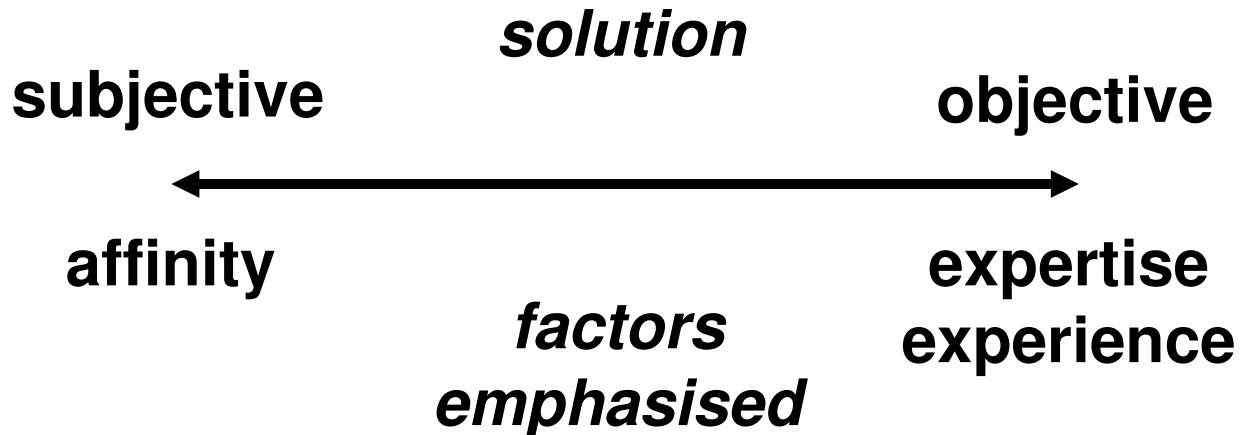
expertise and **experience**
cited most frequently



characteristics of the task influenced
the choice of trusted sources



Trust and Task Characteristics

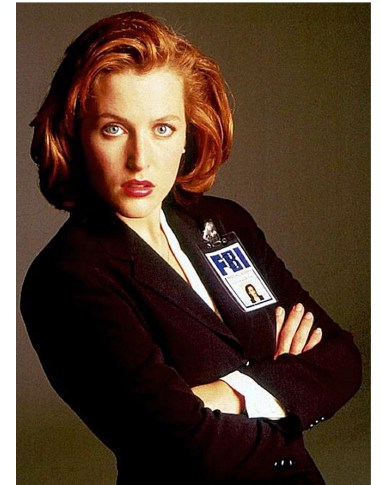


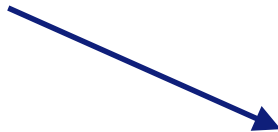
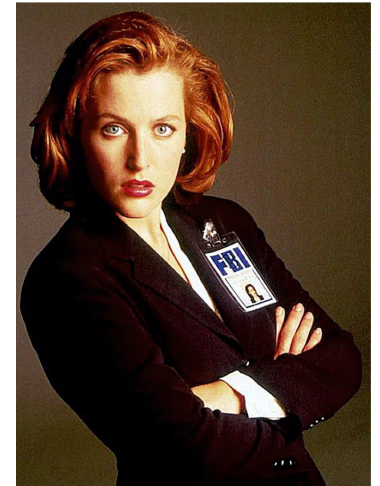


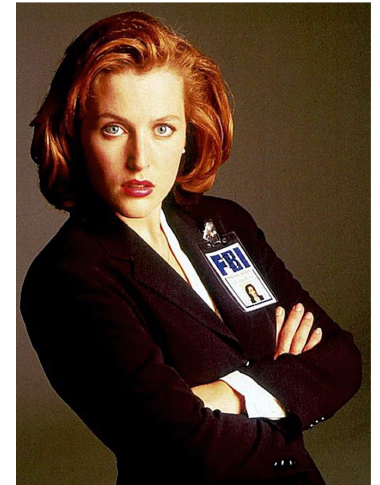
Who should Fox Mulder trust for restaurant recommendations?













Technical Approach



Hoonoh: a Source-centric Search Engine

- Results:
 - people first, items second
- Ranking of results according to:
 - who you know
 - who knows what
 - trust metrics
 - task characteristics



- Social network information (who knows who)
 - FOAF
- Topics (who knows what)
 - tagging
- Trust metrics (who to trust)
 - mine from (Semantic) Web data, based on trust algorithms
- Task profiles (task context)
 - see later



Data Sources: Revyu.com and Beyond



- Semantic Web reviewing and rating site
- Developed to address shortcomings in existing data sources
- Enables integration of review data with social networks
- Allows easy reuse of review data in computing trust metrics

What is Revyu.com?

Revyu.com is a web site where you can review and rate things. Unlike many other reviewing sites on the web, Revyu.com lets you review and rate absolutely anything you can name.

Start a New Review

Recent Reviews [RSS WEB FEED](#) [\(All Reviews\)](#)

1. [The Turf Tavern, Oxford](#) by [fjane37](#)
2. [Riviera Hotel, Haeundae, Busan, Kor...](#) by [tom](#)
3. [Ops Bakery, Haeundae, Busan, Korea](#) by [tom](#)
4. [Motorcare Service Centres](#) by [hockeyshooter](#)
5. [The Fine Burger Company](#) by [Paddy](#)
6. [Park Farm Parking - Bristol Airport](#) by [tigger](#)
7. [The Chap live at the Cube, Bristol,...](#) by [martinp](#)
8. [CheapoAir.com, Cheapo Air Review](#) by [BennyBob](#)
9. [www.travelodge.co.uk](#) by [tom](#)
10. [best restaurants in Paris](#) by [davis9](#)

Popular Tags [\(All Tags\)](#)

[accommodation](#) [alberta](#) [article](#) [aswc2007](#) [banff](#) [bar](#)
[beer](#) [book](#) [bristol](#) [cafe](#) [central-milton-keynes](#) [coffee](#)
[comedy](#) [curry](#) [devon](#) [eating](#) [eating-out](#) [fenny-stratford](#) [film](#)
[food](#) [france](#) [hotel](#) [indian](#) [iswc2007](#) [london](#)
[milton-keynes](#) [movie](#) [music](#) [organic](#) [paper](#) [paris](#)
[pub](#) [real-ale](#) [research](#) [restaurant](#)
[semantic-web](#) [shop](#) [shopping](#) [stony-stratford](#)
[takeaway](#) [travel](#) [web](#) [wolverton](#)

Top Reviewers [\(All People\)](#)

[AdamRae](#) [AdrianStevenson](#) [al](#) [AlanJ](#) [AlexLittle](#) [Aneta](#) [bouquet](#)
[castagna](#) [Crash](#) [DnyaneshRajpath](#) [drewp](#) [esfandia](#) [Fin](#) [Fouad](#)
[glittrgirl](#) [hockeyshooter](#) [jccq](#) [Kake](#) [magicrebirth](#) [Mark](#)
[Marta](#) [martinp](#) [mgaved](#) [Paddy](#) [Paul](#) [pip](#) [redwards](#) [revking](#)
[Rui](#) [Sanyukta](#) [smonroe](#) [sofia](#) [Stefania](#) [stoermer](#) [teddypolar](#)
[tom](#) [trickybastard](#) [vladtn](#) [wcrosbie](#) [xcv](#)

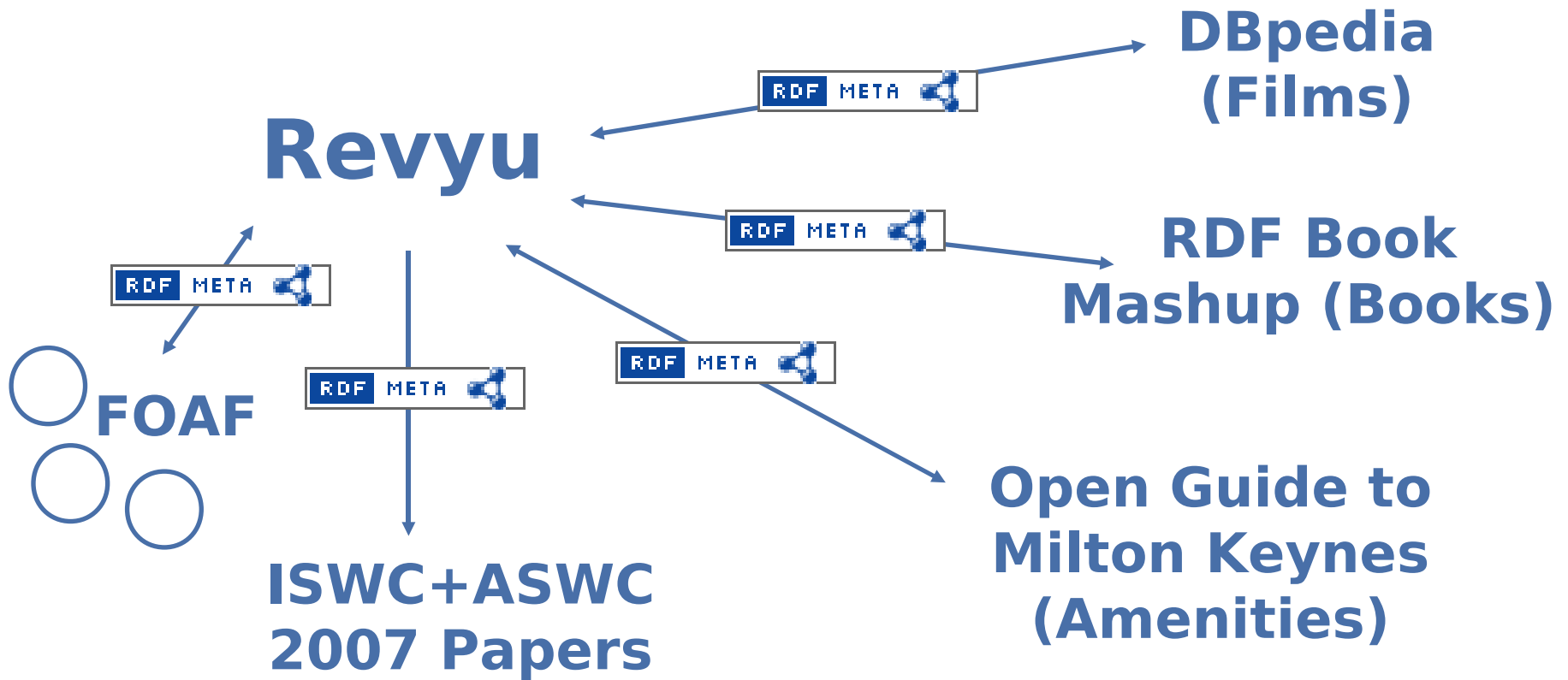
Blog.Revyu.com

- Read [Blog.Revyu.com](#) for news, announcements, fixed bugs, and features delivered on Revyu.com.

Revyu SPARQL Endpoint

- Do You Run a Web Site? You can use reviews from Revyu.com on your site by querying our [SPARQL Endpoint](#).
- Want to learn more about SPARQL? Read the [SPARQL entry on Wikipedia](#) for an overview and links to resources.

- HTML interface for regular users to read and create reviews
 - no SW knowledge required
 - live, public, used daily by real people
- Reviews published transparently in RDF
 - Uses Review vocab, FOAF, Tag and SKOS data also in RDF
- HTML and RDF both crawlable
- SPARQL endpoint to query Revyu data
- Linked with other data sources, according to Linked Data principles





Publishing Linked Data

```
<owl:Thing
```

```
  rdf:about="http://revyu.com/things/iswc-aswc-2007-in-use-paper-715-dbpedia-a">
```

```
  <rdfs:label>DBpedia: A Nucleus for a Web of Open Data (Article)</rdfs:label>
```

```
  <owl:sameAs
```

```
    rdf:resource="http://data.semanticweb.org/conference/iswc-aswc/2007/tracks/in-use/papers/715">
```

```
  <rdf:type rdf:resource="http://swrc.ontoware.org/ontology#InProceedings"/>
```

```
</owl:Thing>
```




- Enhances the user experience without duplicating data
- Revyu currently consumes FOAF, DBpedia, OpenGuides and RDF Book Mashup data

Reviews by *tom* (169)

[Riviera Hotel, Haeundae, Busan, Korea](#)

★★★★☆ on 11 Nov 2007

Pretty average is the best description for this hotel. The concrete jungle feel of the forecourt and entry through the bargain-laden department store don't help the image. It's not so bad inside, and the reception staff are really friendly.

The rooms are clean and comfy enough, if a little bit tired. The mini-bar is good for when you get the munchies and the fridge makes a good place to store some snacks (bought for a fraction of the price at the convenience store down the road!).

In an ideal world there'd be more hangers (isn't that always the case?), an iron, some decent lighting, more power sockets and some drawers. Perfectly bearable for a week, but not exactly home from home.

[Ops Bakery, Haeundae, Busan, Korea](#)

★★★★★ on 11 Nov 2007

Situated just round the corner from the Riviera Hotel, this European-style bakery is a bit of a saviour if you can't manage Korean food for breakfast, lunch *and* dinner. They sell a range of very tasty Italian breads, pastries (French, Danish), and savoury snacks like baguette pizzas, and it wouldn't be at all out of place in a fashionable area of a major European city.

Don't be fooled by the things that have labels only in Korean and look like oval jam doughnuts (you wish!), even if they're in amongst other sweet things; the one I tried was actually filled with a mild curry filling that may or may not have been vegetarian (think British-style curry pasty, in a doughnut rather than pastry). It was not at all what I expected but actually very tasty.

The system seems to be (at least it worked for me) to help yourself by taking a tray and set of tongs from by the door, putting the things you want on your tray and taking them to the till for wrapping. Receipts are possible if you need them; pointing at the part of the till where they are printed and smiling seems to do the job.

It's not particularly cheap, but decent enough value for money and compares favourably to paying for a hotel breakfast. There's a tea shop inside as well off to one side. I didn't try it but it looks decent enough.

www.travelodge.co.uk

★★★★☆ on 29 Oct 2007

This is the main national Web site for Travelodge, the hotel group. It's rubbish. The navigation is totally unpredictable and often fails. Hitting the Back button often takes you several pages back, and I kept experiencing errors saying "Your search has timed out. Please search for offer rooms again". Very tedious so I

Do you know **tom**? [Login](#) or [Register](#) to add **tom** to your Network

[Web Feed of tom's Latest Reviews](#)

RSS WEB FEED

About **tom** (Tom Heath)



[tom's Home Page](#)


tom's location:
[Borough of Milton Keynes](#)

tom's Interests

[Semantic Web](#)
[ESWC2006 Semantic Web](#)
[Technologies](#)
[Beer](#)

[RDF Metadata About tom](#)

RDF META

 [Add to del.icio.us](#)

Reviews by *tom* (169)

Riviera Hotel, Haeundae, Busan, Korea

★★★★☆ on 11 Nov 2007

Pretty average is the best description for this hotel. The concrete jungle feel of the forecourt and entry to the bargain-laden department store don't help the image. It's not so bad inside, and the reception staff are really friendly.

The rooms are clean and comfy enough, if a little bit tired. The TV and the fridge makes a good place to store some snacks (bought from the store down the road!).

In an ideal world there'd be more hangers (isn't that always the case?), an iron, some decent lighting, more power sockets and some drawers. Perfectly bearable for a week, but not exactly home from home.

Ops Bakery, Haeundae, Busan, Korea

★★★★★ on 11 Nov 2007

Situated just round the corner from the Riviera Hotel, this European-style bakery is a bit of a saviour if you can't manage Korean food for breakfast, lunch *and* dinner. They sell a range of very tasty Italian breads, pastries (French, Danish), and savoury snacks like baguette pizzas, and it wouldn't be at all out of place in a fashionable area of a major European city.

Don't be fooled by the things that have labels only in Korean and look like oval jam doughnuts (you wish!), if they're in amongst other sweet things; the one I tried was actually filled with a mild curry filling that may not have been vegetarian (think British-style curry pasty, in a doughnut rather than pastry). It was not all what I expected but actually very tasty.

The system seems to be (at least it worked for me) to help yourself by taking a tray and set of tongs from the door, putting the things you want on your tray and taking them to the till for wrapping. Receipts are possible if you need them; pointing at the part of the till where they are printed and smiling seems to do the job.

It's not particularly cheap, but decent enough value for money and compares favourably to paying for a hot breakfast. There's a tea shop inside as well off to one side. I didn't try it but it looks decent enough.

www.travelodge.co.uk

★★★★☆ on 29 Oct 2007

This is the main national Web site for Travelodge, the hotel group. It's rubbish. The navigation is totally unpredictable and often fails. Hitting the Back button often takes you several pages back, and I kept experiencing errors saying "Your search has timed out. Please search for offer rooms again". Very tedious s



About tom (Tom Heath)



[tom's Home Page](#)

[tom's location:
Borough of Milton Keynes](#)

[tom's Interests](#)

[Semantic Web
ESWC2006 Semantic
Web Technologies
Beer](#)

[RDF Metadata About
tom](#)





Broken Flowers

Links

Homepage: <http://www.brokenflowersmovie.com/>
See Also: http://en.wikipedia.org/wiki/Broken_flowers

Tags

[bill-murray](#) [film](#) [jessica-lange](#) [jim-jarmusch](#) [julie-delpy](#) [movie](#) [sharon-stone](#)

Reviews (1)

★★★★★ [by tom on 30 Jan 2007](#)

Broken Flowers provides a fantastic vehicle for a classic deadpan Bill Murray performance. The film centers around his character Don, who one day receives a letter from an ex-girlfriend, telling him he has a teenage son. The letter is unsigned, so (with encouragement from his neighbour) he sets off round the country, visiting each the exes who could be the mother of his son. Predictably they're all different in personality and life situation, giving plenty of raw material for awkward silences and dubious encounters. This is great viewing for any Bill Murray fans, or anyone who likes their humour intelligent and a little bit quirky. The soundtrack is also excellent, and deserves a separate review.

What do you think of **Broken Flowers**? [Write Your Own Review...](#)

Broken Flowers



directed by [Jim Jarmusch](#)



Broken Flowers



directed by **Jim Jarmusch**

RDF Metadata About Broken Flowers



Broken Flowers

Links

Homepage: <http://www.brokenflowersmovie.com/>
See Also: http://en.wikipedia.org/wiki/Broken_flowers

Tags

[bill-murray](#) [film](#) [jessica-lange](#) [jim-jarmusch](#) [julie-delpy](#) [movie](#) [sharon-stone](#)

Reviews (1)

★★★★★ **by tom on 30 Jan 2007**

Broken Flowers provides a fantastic vehicle for a classic deadpan Bill Murray performance. The film centers around his character Don, who one day receives a letter from an ex-girlfriend, telling him he has a teenage son. The letter is unsigned, so (with encouragement from his neighbour) he sets off round the country, visiting each the exes who could be the mother of his son. Predictably they're all different in personality and life situation, giving plenty of raw material for awkward silences and dubious encounters. This is great viewing for any Bill Murray fans, or anyone who likes their humour intelligent and a little bit quirky. The soundtrack is also excellent, and deserves a separate review.

What do you think of **Broken Flowers**? [Write Your Own Review...](#)



- Conference attendance data
 - from ESWC2006
- Tagging data
 - from delicious.com



Hoonoh Trust Algorithms and Hoonoh.com



- Algorithms for Computing Knowledge and Trust Relationships
 - Based on findings of earlier empirical study
 - Use reviews and background data sources as input
 - Generate data for Hoonoh source-centric search engine



- Based on the 3 most significant factors
 - Expertise
 - Experience
 - Affinity
- Some proxy metrics required
 - Credibility (Expertise)
 - Usage (Experience)



Credibility Algorithm (Expertise)

- Person -> Topic Relationship
1. For a particular tag, get all items tagged with that tag
 2. For each item find its mean rating
 3. For each review of the item calculate how far the rating varies from the mean rating for that item
 4. Low rating distance = high credibility score for that review
 5. Sum each reviewer's credibility scores for reviews of items tagged with the tag
 6. Find each reviewer's mean credibility score for that tag



Usage Algorithm (Experience)

- Person -> Topic Relationship
1. For a particular tag
 2. Count how many times each reviewer has reviewed an item tagged with that tag (by anyone)
 3. This gives a reviewer's tag count
 4. Find the highest of these tag counts across all users of the tag
 5. Each reviewer's usage score for a tag =
their tag count / highest tag count + constant



Affinity Algorithm

- Person -> Person Relationship
- 1. Get all reviews by a user A and another user B
- 2. Count the number of items reviewed by both
- 3. Calculate *item overlap ratio*:
number of items reviewed by both / total reviews by A
- 4. For each overlapping item
 - 1. Find the distance between the two ratings
 - 2. Low rating distance = High rating overlap
 - 3. Find *mean rating overlap* between Users A and B
- 5. combine the *item overlap ratio* and *mean rating overlap* to produce a measure of *affinity*(A,B)



Usage (Experience) Scores from delicious.com Data

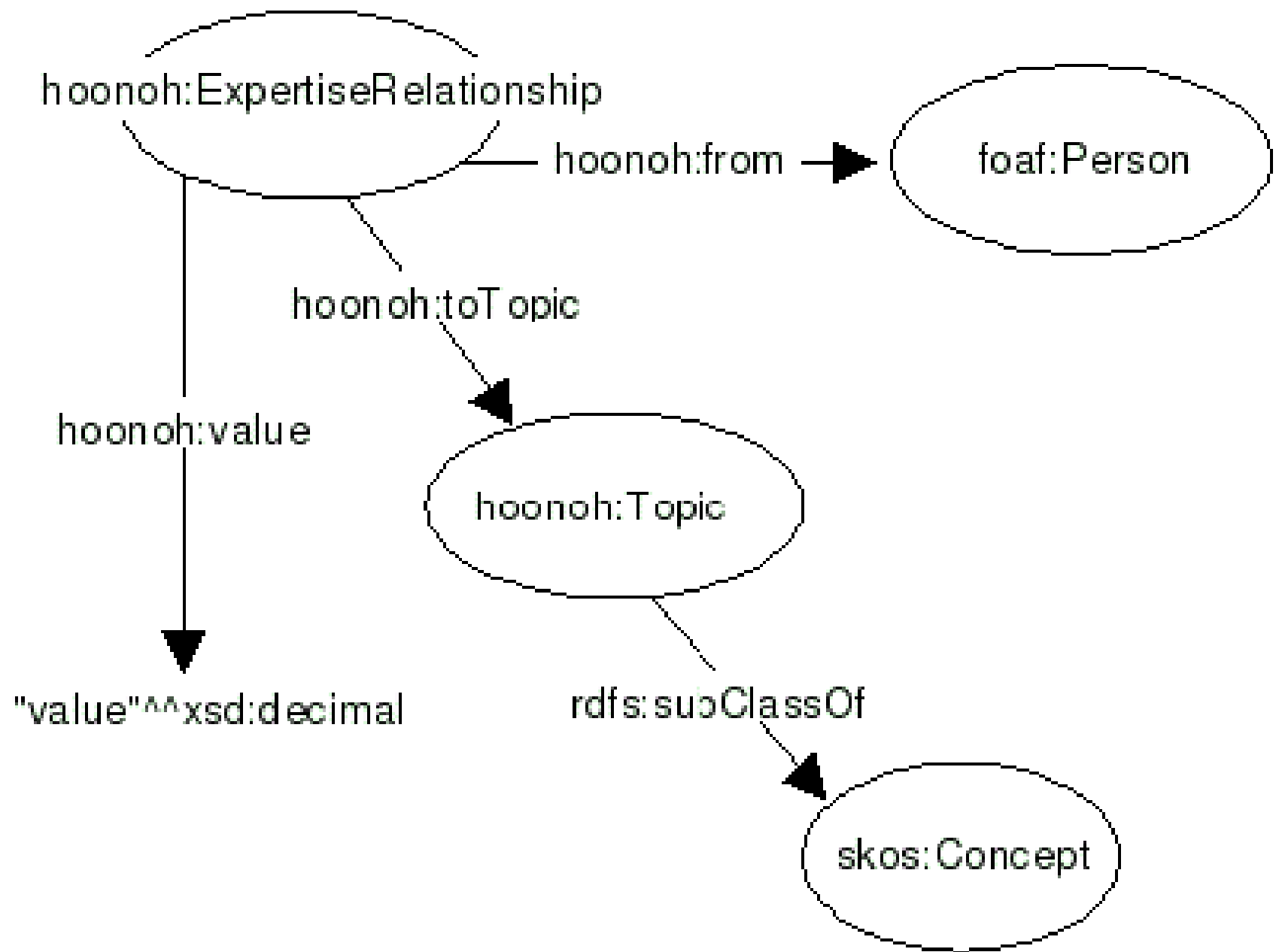
- Person -> Topic Relationship
1. Get a user's most used tags from delicious.com
 2. Compare against existing (Revyu-derived) usage scores
 3. If no previous relationship exists then add new usage topics and nominal scores
 4. Or raise existing scores to a nominal level
 5. Could easily be extended to Flickr, Digg, etc.

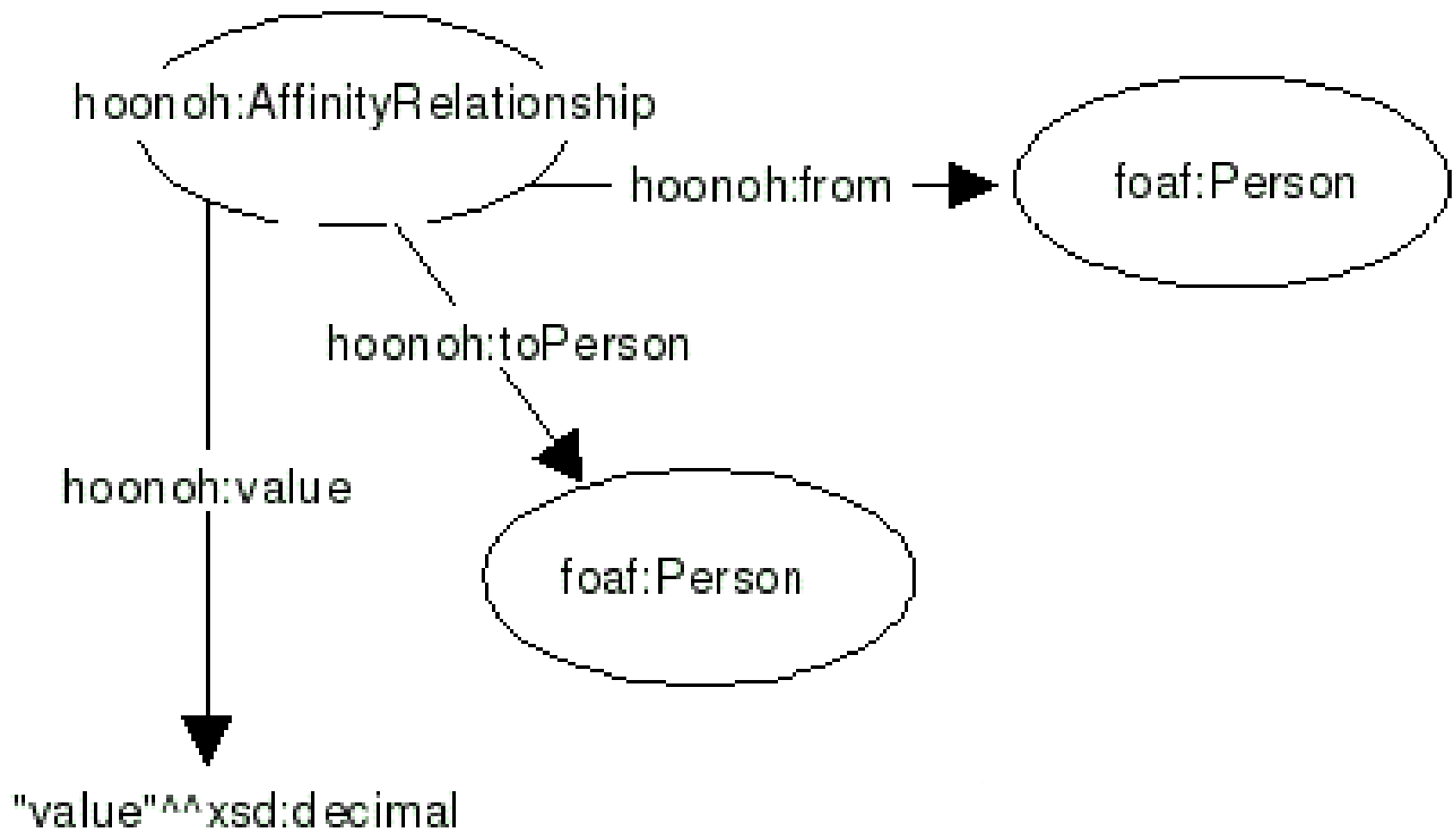


- For representing derived trust metrics
- Why a new ontology?
 - insufficient expressivity in existing ontologies
 - no evidence that trust is a binary relationship
 - binary values less meaningful when combining evidence from multiple sources
 - more amenable to SPARQL querying (and post-processing)



- Trust Relationships are classes
 - TopicalRelationship(s)
 - ExpertiseRelationship
 - ExperienceRelationship
 - ImpartialityRelationship
 - InterpersonalRelationship(s)
 - AffinityRelationship
 - TrackRecordRelationship







- Trust metrics generated using Hoonoh algorithms
- Stored in dedicated triplestore
- Combined with other data sources (e.g. FOAF)
- Used to power Hoonoh.com source-centric search engine



- “Tells you who you know who knows”
- Search for information sources within your social network
 - Public, Web-based system supporting information-seeking through social networks
 - Source-centric approach
 - Enables ranking of sources according to the generated trust metrics

Find Out Who Knows About a Topic

Find Out Who Knows About

Submit

example queries: [film](#) [restaurant milton-keynes](#)

Who Knows About **film**?

Weight Results by: Experience | Expertise | Affinity |

[1] 1.733 Tom Heath - what do they know about film?

[2] 1.146 Martin Poulter - what do they know about film?

[3] 1.042 slowman - what do they know about film?

[4] 1.042 AdamRae - what do they know about film?

[5] 1.042 mqaved - what do they know about film?

Who Knows About **film**?

Limit Results to: Friends (55) + [Friends of Friends](#) + [Friends of Friends of F](#)

Weight Results by: [Experience](#) | [Expertise](#) | Affinity |

[1] 0.683 [Crash](#) - [what do they know about film?](#)

[2] 0.675 [drewp](#) - [what do they know about film?](#)

[3] 0.592 [hockeyshooter](#) - [what do they know about film?](#)

[4] 0.242 [Fin](#) - [what do they know about film?](#)

[5] 0.142 [cancer](#) - [what do they know about film?](#)

What Does *drewp* Know About *film*?

film Reviews

Review of: Idiocracy

5/5 on 29 March 2007

Very silly, lots of great satire. ([More of Idiocracy... at Revyu.com](#))

Review of: You, Me and Dupree

4/5 on 24 November 2006

Pleasant formula comedy. I imagine I'd watch it again (if I ever watched movies c the theater) ([More of You, Me and Dupr... at Revyu.com](#))

Review of: casino royale

3/5 on 28 November 2006



Evaluation



- Methodology
 - Two scenarios
 - Restaurants in Milton Keynes, Professional Camera Equipment
 - Reviews written on Revyu.com
 - Trust metrics computed using Hoonoh algorithms
- Question
 - To what extent do Hoonoh rankings of information sources correlate with reported rankings from participants?
- Some significant results, particularly for *experience*
- Room for further improvements, probably



Conclusions and Future Work



Conclusions

- People make sophisticated source selection decisions when information-seeking
- Perceived trustworthiness is source, relationship and task dependent
- These factors can be operationalised as algorithms
- Possible to mine Web data to generate trust metrics that approximate reported judgements



- More source data required!
- Development of non-proxy metrics
- Greater use of semantics
 - Concepts rather than tags
 - Semantic propagation of trust relationships
 - Trust decay
- Combining trust metrics derived from multiple sources



- Thesis
 - <http://tomheath.com/thesis>
- Contact Details
 - <http://tomheath.com/id/me>
 - tom.heath@talys.com
- Acknowledgements:
 - Enrico Motta
 - Marian Petre
 - Martin Dzbor

