

Managing Online Business Communities

The ROBUST Project

Steffen Staab, Thomas Gottron
& ROBUST Project Team

- Information ecosystems
 - Employees
 - Business Partners, Customers
 - General Public



Valuable asset



Opportunities



Risks





Lotus Connections

Communities

- Employees
- Working groups
- Interest Groups
- Projects

Business value

- Task relevant information
- Collaboration
- Innovation

Volume

- 4,000 posts/day
- 386,000 employees
- 1.5GB content/day

Employees
Intranet



SAP Community Network (SCN)

Communities

- Customers
- Partners
- Suppliers
- Developers

Business value

- Products support
- Services
- Find business partners

Volume

- 6,000 posts/day
- 1,700,000 subscribers
- 16GB log/day

Business Partners
Extranet



MeaningMine

Communities

- Social media
- News
- Web fora
- Public communities

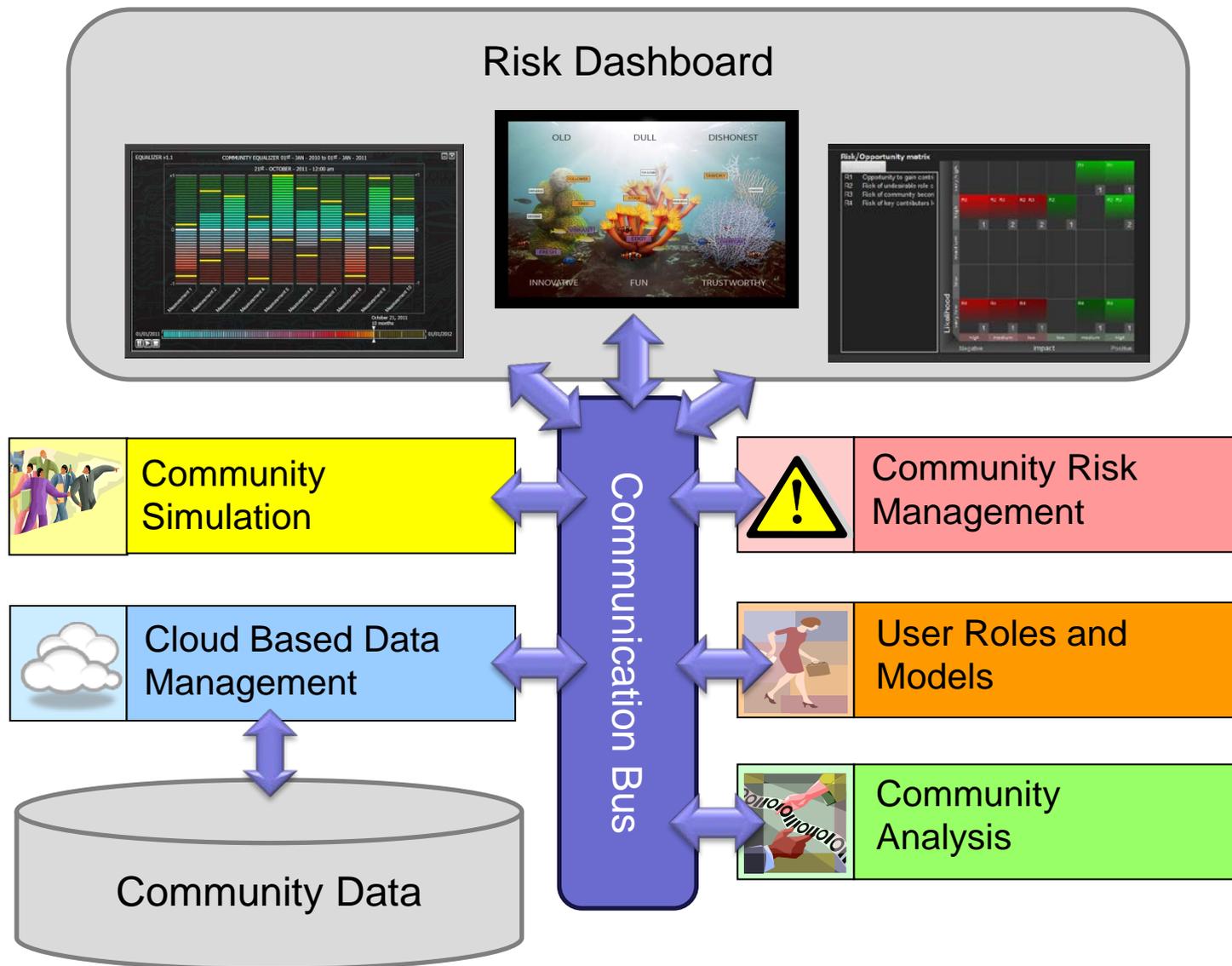
Business value

- Topics
- Opinions
- Service for partners

Volume

- 1,400,000 posts/day
- 708,000 web sources
- 45GB content/day

Public Domain
Internet



Community Analysis



- Feature Selection
- Structural Analysis
- Behavior Analysis
- Content Analysis
- Cross Community

Community Simulation



- Community Models
- Policy Models
- Influence
- Simulation
- Prediction

Community Risk Management



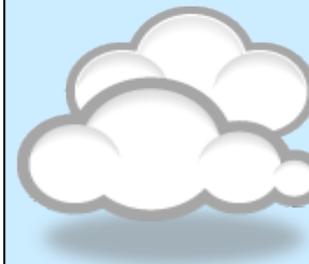
- Risk Formalization
- Risk Detection
- Risk Forecasting
- Risk Management
- Risk Visualization

User Roles and Models



- User Needs
- Motivation
- Roles
- Groups
- Community Value

Cloud Based Data Management



- Operations on Data
- Scalability
- Real Time
- Parallel Execution
- Stream Based

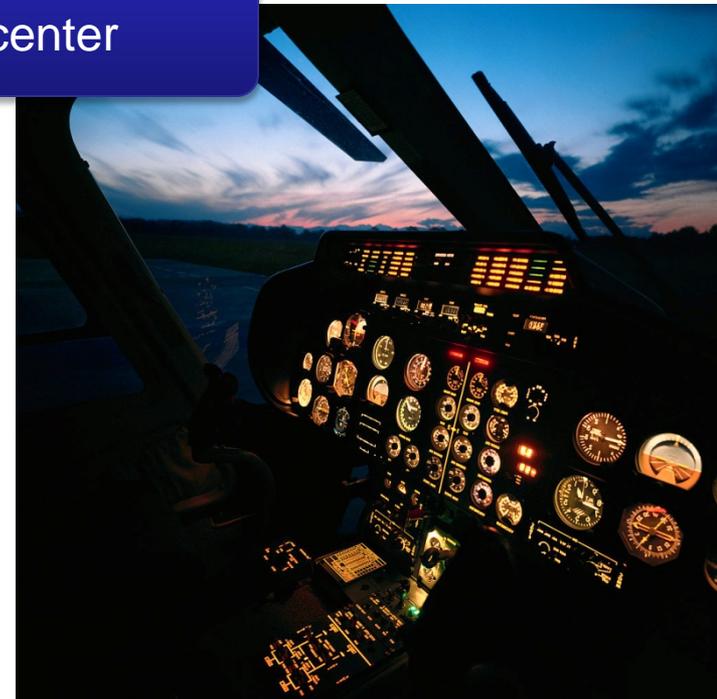
The ROBUST platform

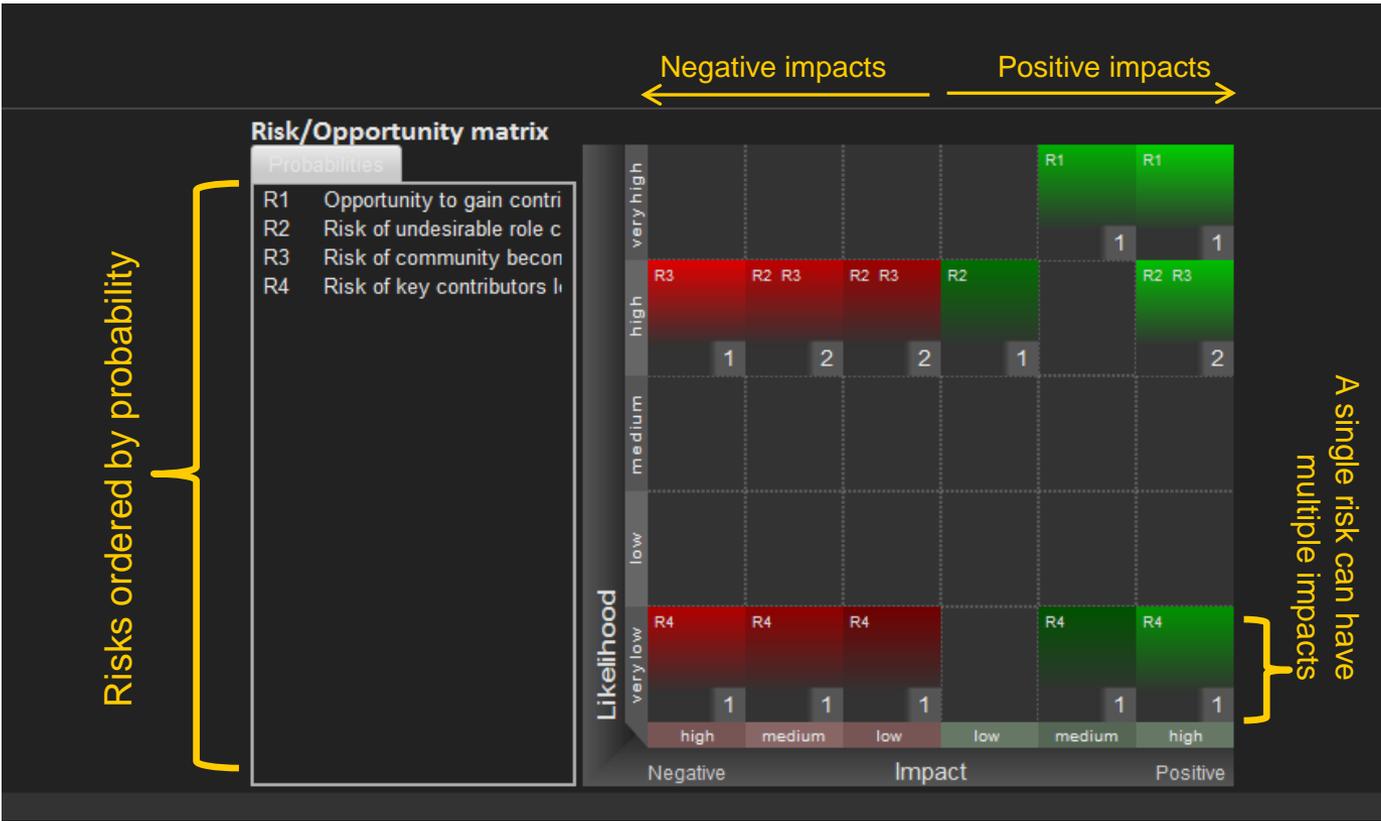
Risk Dashboard and Visualizations

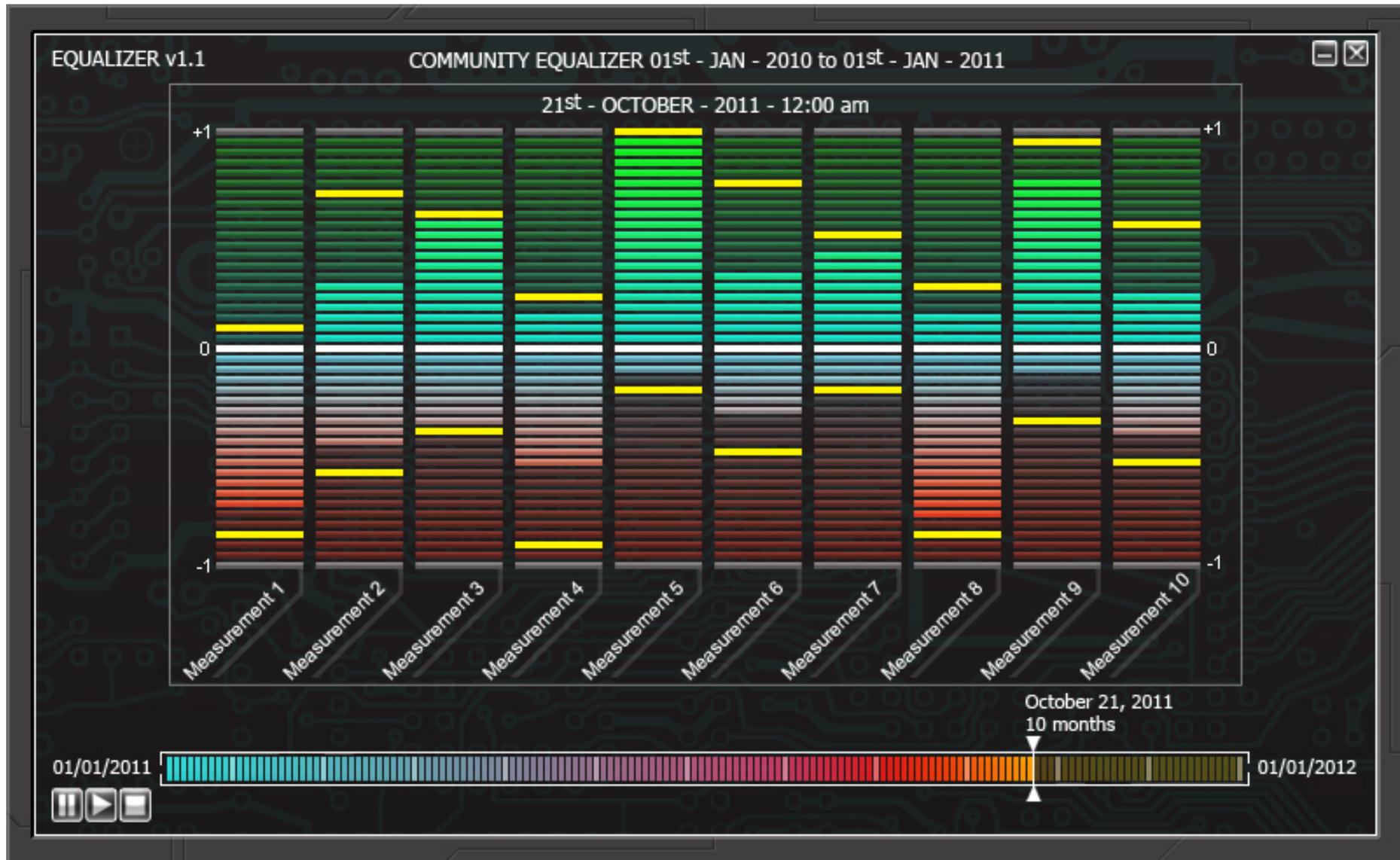


- Definition of risks/opportunities
- Monitoring
- Interaction with community
- Reaction and countermeasures

→ Need for a control center





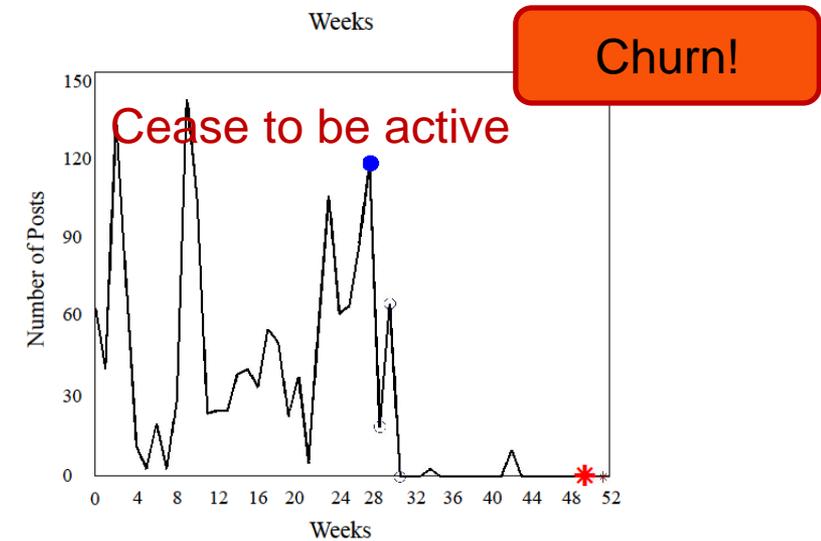
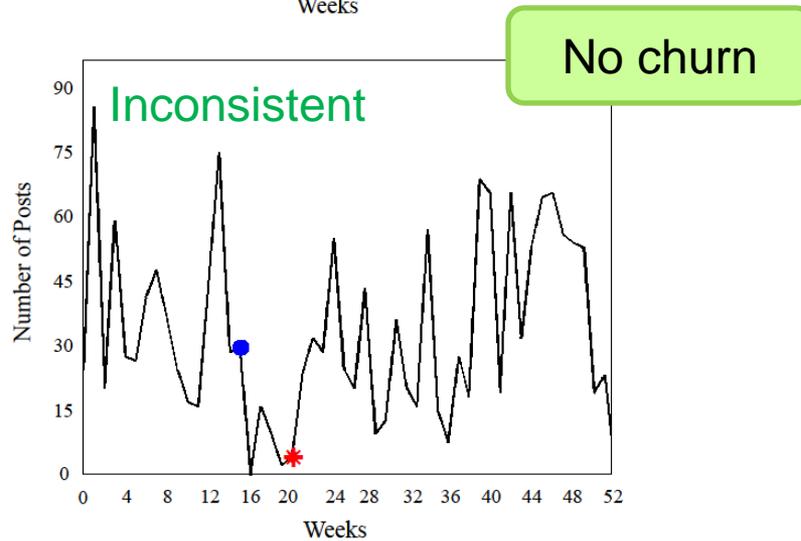
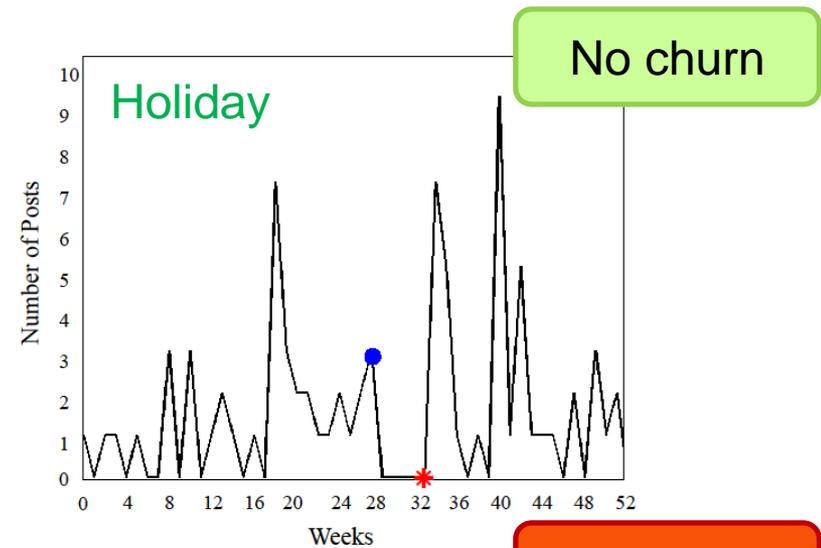
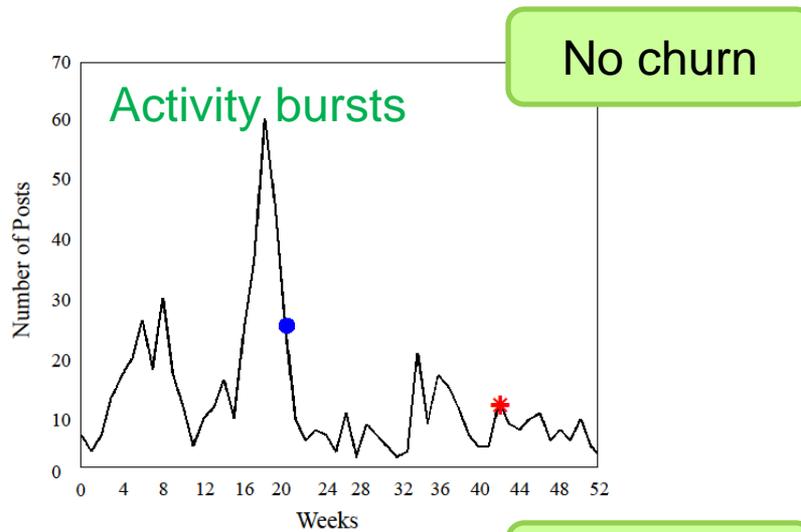


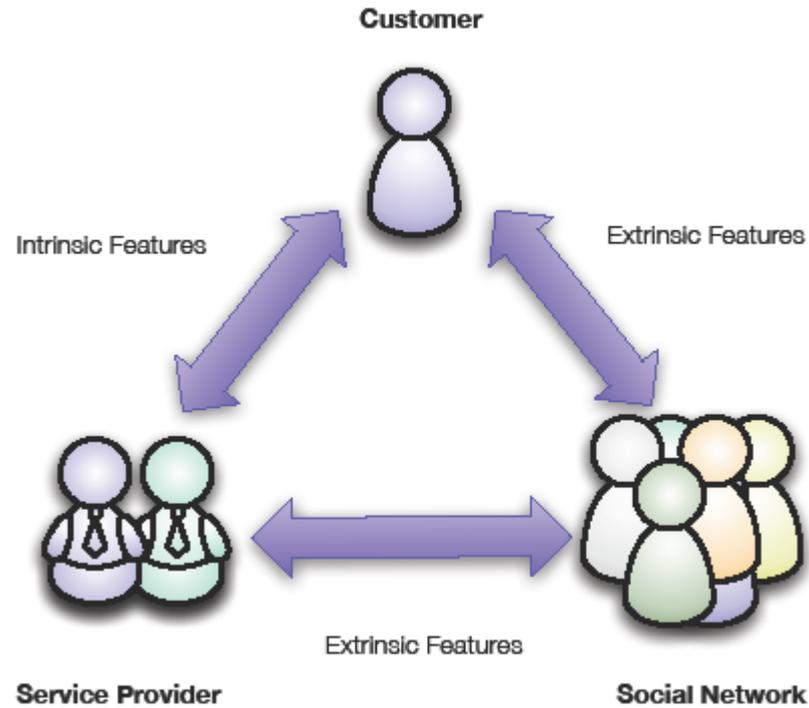
Community Analysis,
User Roles & Risk Management

Determine Risk:
Users Leaving the Community

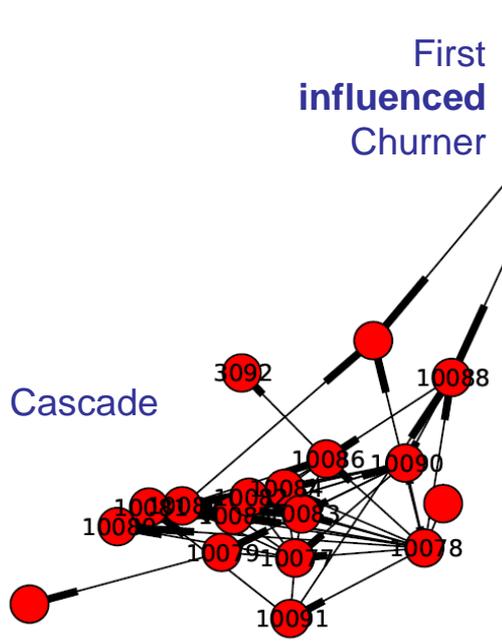
- Users churn = Users leave the community, become inactive
- Questions:
 - Why do users leave the community?
 - Who is leaving the community?
 - Impact?
- Tasks:
 - Detect “Churn”
 - Predict “Churn”
 - Evaluate “Churn”

What is Churn? Activity?

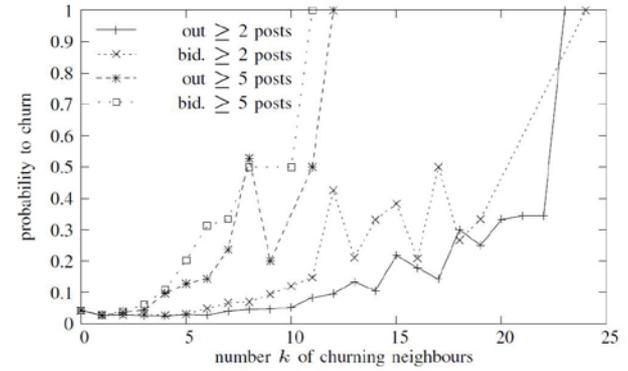
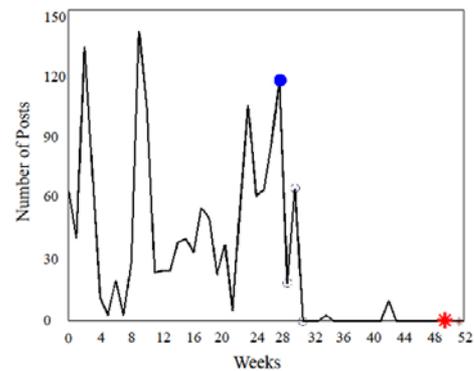




- Churn: churning users influence other users they communicate with



- Temporal ordering: Churners that churned subsequent to each other
- Frequently, after a slow start, resulting in a cascade of churning

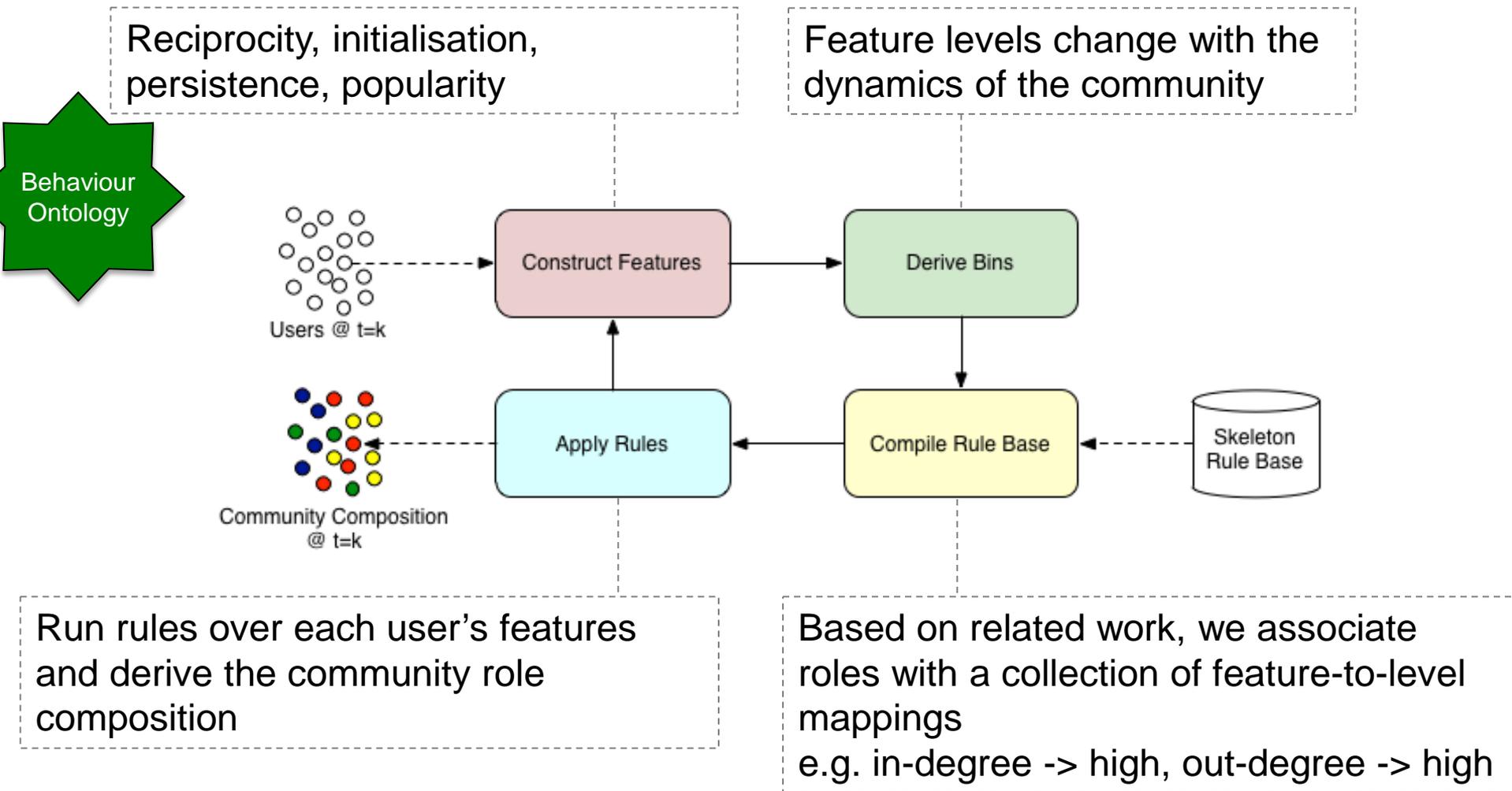


Churn of a single user → Influences his neighbours

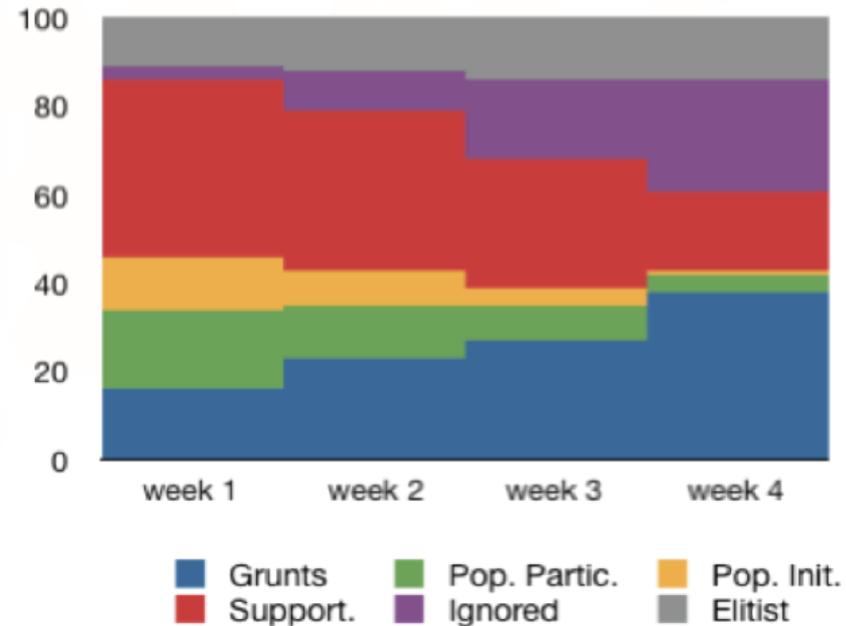
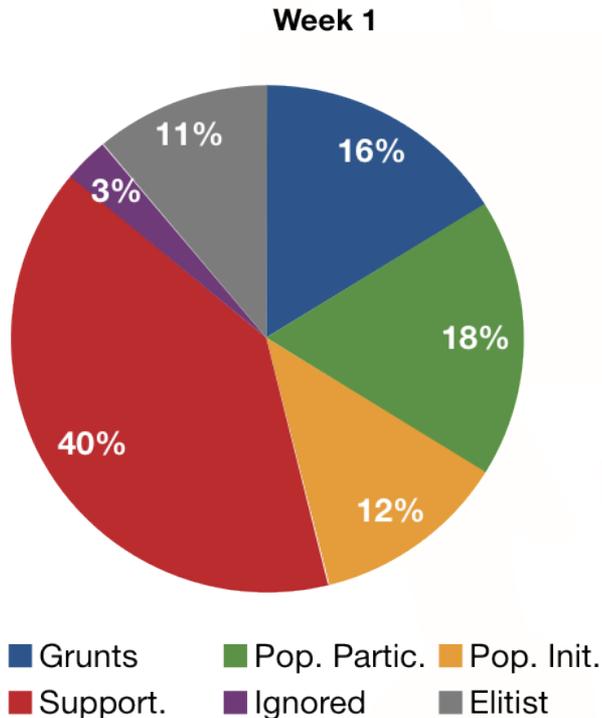
- A role represents the standing, or part, that a user has within a given community

Role \ Dimension	Reciprocity	Initialisation	Persistence	Popularity
Elitist	high: Threads low: Neighbours			low
Joining Conversationalist		low	high	
Popular Initiator		high		high
Supporter	medium	low		medium
Taciturn	low		low	
Ignored				low: No replies





- Role composition in community



- Development towards an unhealthy role composition!

Community & Content Analysis

Opportunity Detection: Discovering Interesting Contents

IBM Collaboration Solutions Marketing Community

Follow this Community Join this Community

IBM Collaboration Solutions Marketing Community



Overview

- Members
- Events (BETA Widget)
- Blog
- Wiki
- Bookmarks
- Files

Overview

This community is designed to foster collaboration and ignite innovation. This is your entry point to ICS Marketing, led by Kristen Lauria, VP of Marketing.

Tags: [collaboration_solutions](#), [ics_marketing](#), [kristen_lauria](#)

Events (BETA Widget)

W3C Social Business
08 - 10 Nov 2011
09:00 AM - 05:00 PM
Online event

View: [All current \(2\)](#) | [All](#)

[Feed for Events](#)

Blog

Drive 4Q IBM Lotus
KRISTEN LAURIA | Tuesday

Let's make it bigger
Sharnette Mitchell | 25 Oct

 [Jimmy Page](#) liked the file [HRLFRR - Social Stream Analysis.doc](#).
32 seconds ago

 [Yi Chun Wang](#) created a new bookmark named [blueDoodle](#), a good way to find a meeting time...
4 minutes ago

 [Chris Banes](#) edited the wiki page [Welcome](#) in the [Young Professional's Community and Social Events](#) wiki.
27 minutes ago

 [David Jones](#) commented on the [Jazz-Connections check-point Meeting Minutes](#) blog entry in the [Jazz-Connections Integration](#) blog.
2 hours ago

 [David Rose](#) created a new thread named [2015 Roadmap Q&A / Discussion](#) in the [Canada Services Community](#).
5 hours ago

 [Eitan Shapiro](#) wrote on the board of [Mark Smith](#).
Hi Mark, here is a feedback activity for the widgets...
yesterday

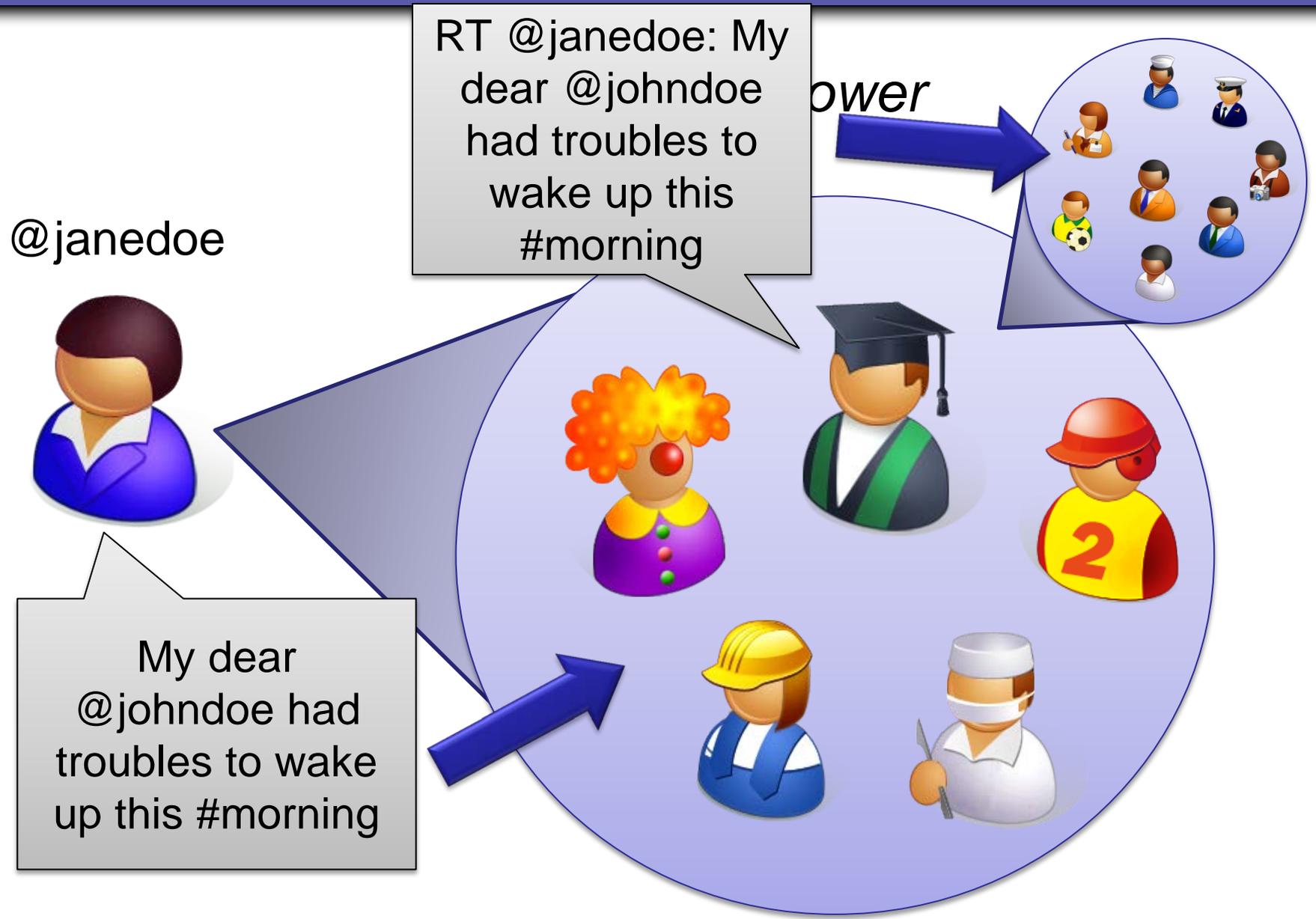
Analogies to Twitter!

Important Bookmarks

- [2011 ICS Activities](#)
- [Analyst Buzz](#)
- [Collaboration Soapbox](#)
- [Cross IBM Communications Contacts](#)
- [External Communication Vehicles](#)
- [IBM Social Business Update](#)
- [ICS Client References Wiki](#)
- [ICS Geo Contacts](#)
- [ICS Marketing org chart](#)
- [ICS Pipeline Community](#)
- [ICS ThumpBook Community](#)
- [ICS Tiger Team & Competitive Marketing](#)
- [Internal Communication Vehicles](#)
- [Lotus Messaging and Collaboration Blog](#)
- [Social Business Demand Program](#)
- [Town Hall charts / replays](#)

Members







- Retweet indicates quality
 - „of interest for others“
- Idea:
 - Learn to predict retweets!

Likelihood of retweet as metric for Interestingness

Aim: Prediction of probability of retweet

Logistic regression:

$$f(z) = \frac{1}{1 + e^{-z}}$$

$$z = w_0 + w_1 x_1 + w_2 x_2 + \dots + w_n x_n$$

Model parameters w_i learned on training data

Dataset	Users	Tweets	Retweets
Choudhury	118,506	9,998,756	7.89%
Choudhury (extended)	277,666	29,000,000	8.64%
Petrovic	4,050,944	21,477,484	8.46%

Feature	Dimensions	Weight
Constant	(intercept)	-5.45
Message feature	Direct message	-147.89
	Username	146.82
	Hashtag	42.27
	URL	249.09
Sentiment	Valence	-26.88
	Arousal	33.97
	Dominance	19.56
Emoticons	Positive	-21.8
	Negative	9.94
Exclamation	Positive	13.66
	Negative	8.72
Punctuation	!	-16.85
	?	23.67
Terms	Odds	19.79

Topic	Weight
social media market post site web tool traffic network	27.54
follow thank twitter welcome hello check nice cool people	16.08
credit money market business rate economy home	15.25
christmas shop tree xmas present today wrap finish	2.87
home work hour long wait airport week flight head	-14.43
twitter update facebook account page set squidoo check	-14.43
cold snow warm today degree weather winter morning	-26.56
night sleep work morning time bed feel tired home	-75.19

- Top-k interesting tweets for „beer“

Rang	Username	Tweet
1	BeeracrossTX	UK beer mag declares "the end of beer writing." @StanHieronymus says not so in the US. http://bit.ly/424HRQ #beer
2	narmmusic	beer summit @bspward @jhinderaker no one had billy beer? heehee #narm - beer summit @bspward @jhinde http://tinyurl.com/n29oxj
3	beeriety	Go green and turn those empty beer bottles into recycled beer glasses! http://bit.ly/2src7F #beer #recycle (via: @td333)
4	hblackmon	Great Divide beer dinner @ Porter Beer Bar on 8/19 - \$45 for 3 courses + beer pairings. http://trunc.it/172wt
5	nycraftbeer	Interesting Concept-Beer Petitions.com launches&hopes 2help craft beer drinkers enjoy beer they want @their fave pubs. http://bit.ly/11gJQN
6	carichardson	Beer Cheddar Soup: Dish number two in my famed beer dinner series is Beer Cheddar Soup. I hadn't had too.. http://bit.ly/1diDdF
7	BeerBrewing	New York City Beer Events - Beer Tasting - New York Beer Festivals - New York Craft Beer http://is.gd/39kXj #beer
8	delphiforums	Love beer? Our member is trying to build up a new beer drinker's forum. Grab a #beer and join us: http://tr.im/pD1n
9	Jamie_Mason	#Baltimore Beer Week continues w/ a beer brkfst, beer pioneers luncheon, drink & donate event, beer tastings & more. http://ping.fm/VyTwg
10	carichardson	Seattle and Beer: I went to Seattle last weekend. It was my friend's stag - he likes beer - we drank beer.. http://tinyurl.com/cpb4n9

[Home](#) | [Top 10](#) | [How does it work?](#) | [References](#)

LiveTweet



LiveTweet

Submit your tweet, and we tell you how interesting it is in the Twitter world right now:

Tell me!

Like 10
Tweet 15
+ reddit this!



EC Project 257859

<http://livetweet.west.uni-koblenz.de/>

Tomorrow I am going to ...

Interestingness:



... play tennis

Interestingness:



... play golf

Interestingness:



... do some gardening

Interestingness:



... spread wisdom

Interestingness:



... go to a Justin Bieber concert

Interestingness:

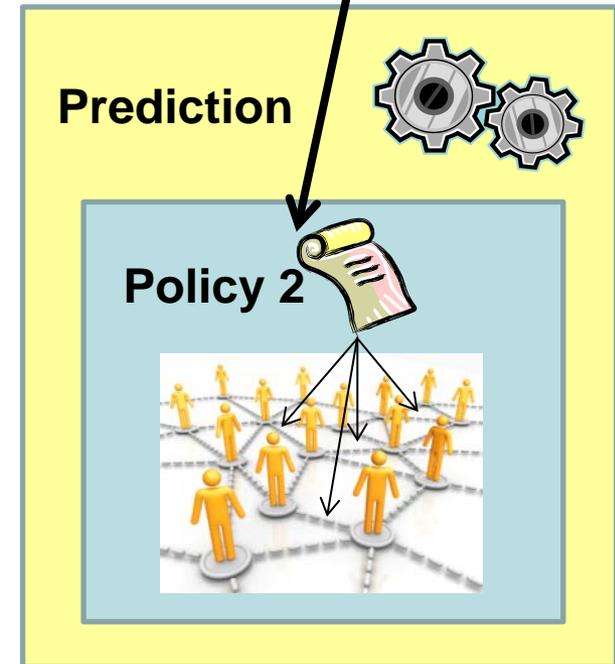
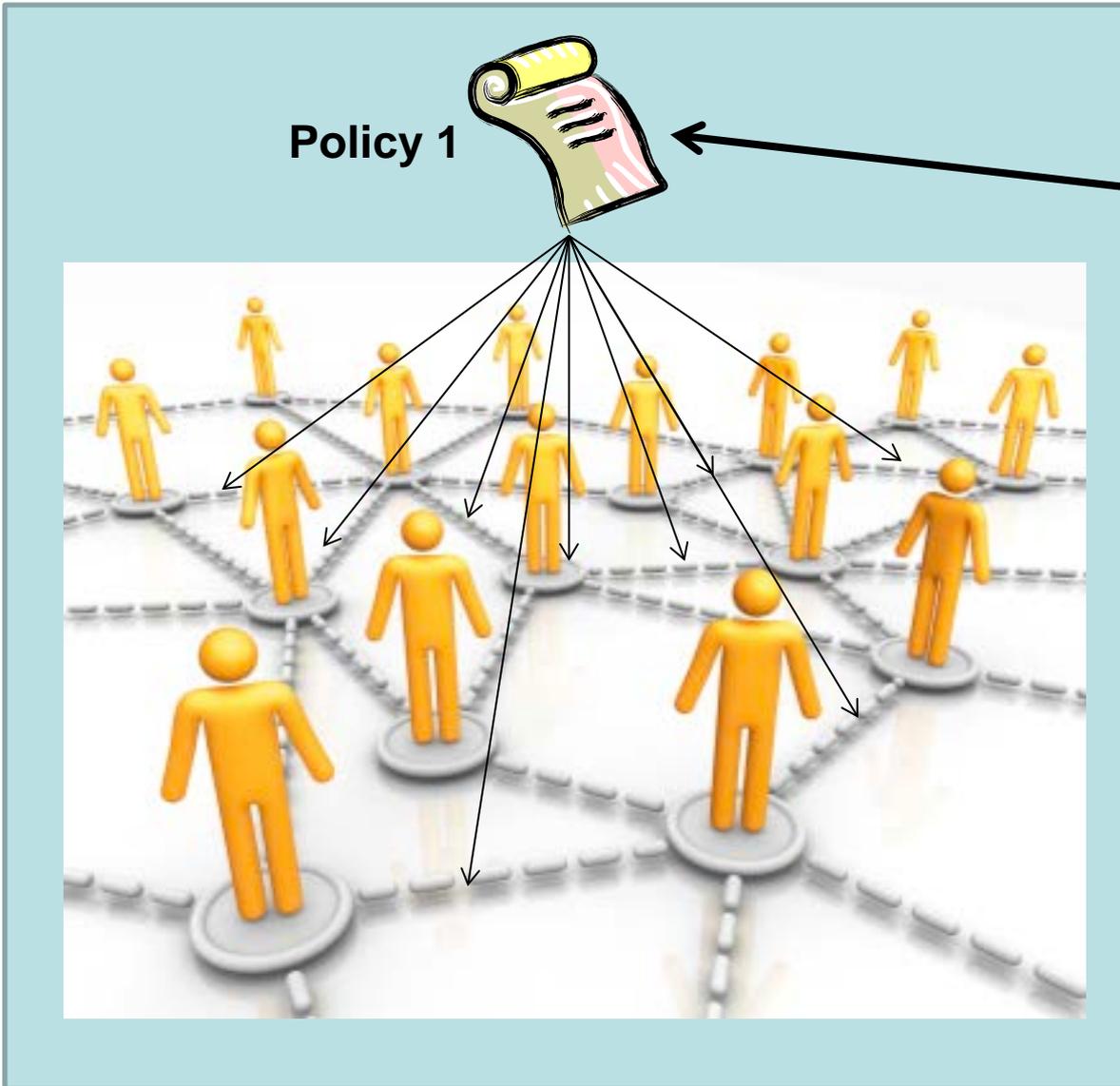


... win the lottery

Community Simulation

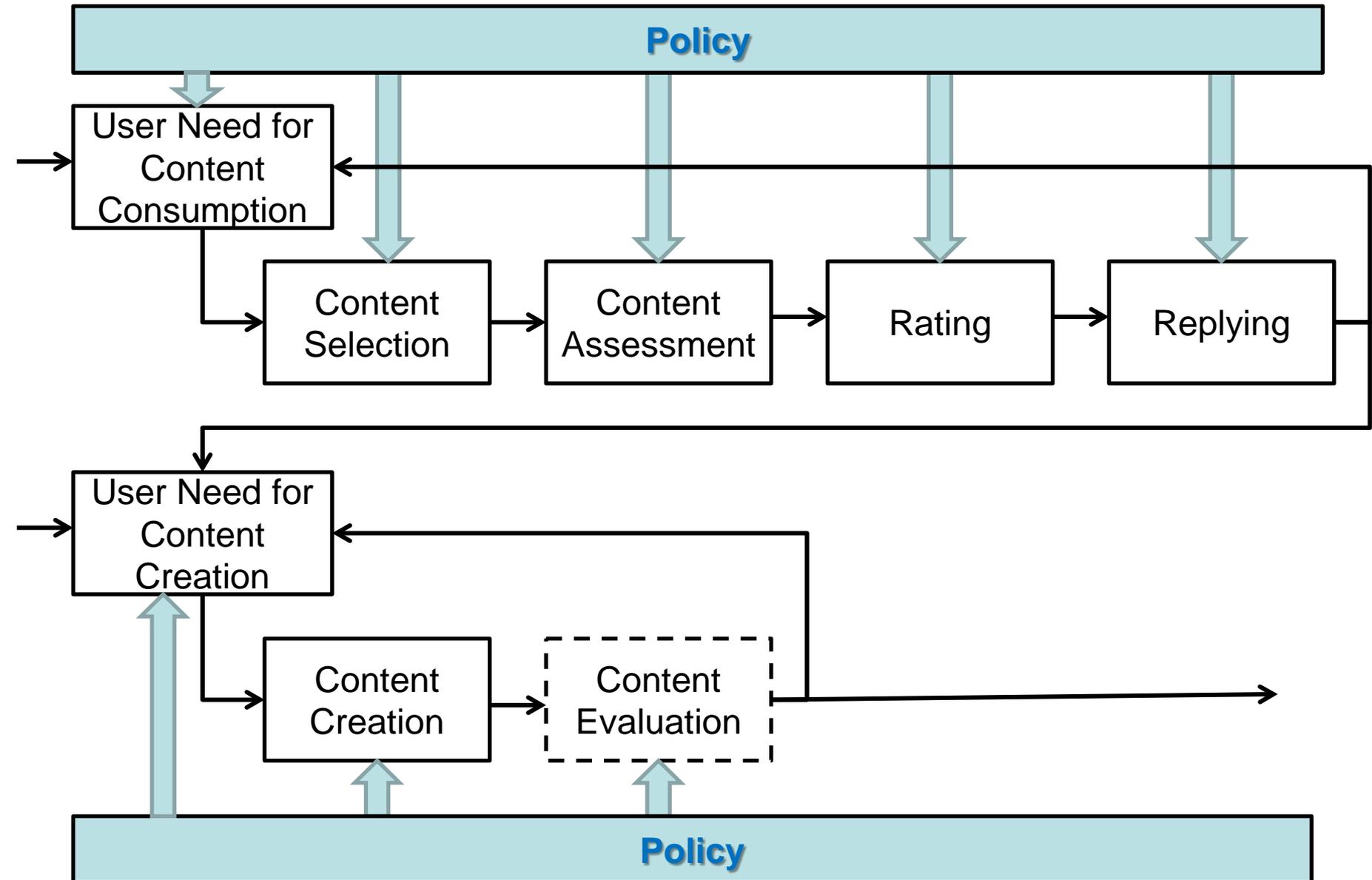
Risk Mitigation: Simulating Effects of Policy Changes

- **Governance of Communities**
 - Def.: Steering and coordinating actions of community members
- **Implementation:**
 - Direct intervention of community owner
 - Functionality of the community platform
- **Mitigation of risks:**
 - Change platform functionality
 - Impact?

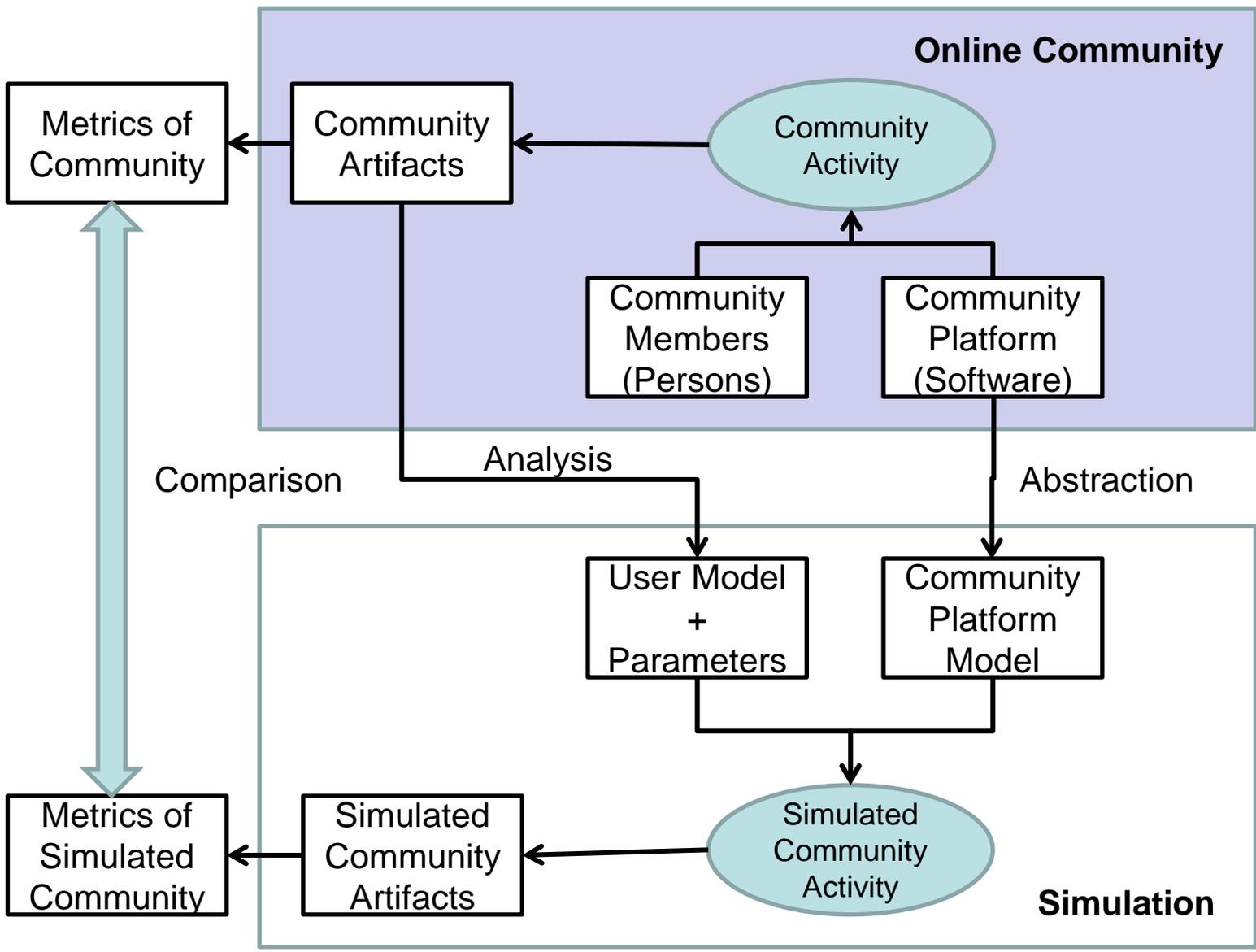


- Users generate content
- Users search content
- Users consume content
- Users interact with content
 - Rate content
 - Reply to content

- Where can a (platform) policy have effects?



- Part of Community: forum activity
- Goal: Steer user activity to specific forums
- User model parameters:
 - Activity rate for creating threads
 - Activity rate for creating replies
 - Preferences for activity in specific forums
- Community Model
 - **Varied** restrictions for thread creation
- Observed Metrics
 - Response time on threads



- Users search content
 - Presentation → Ranking threads in content views
 - Recency
 - Social Closeness
 - Topical closeness
 - Popularity
 - Observation: Influence which questions are answered
- Users generate content
 - Restrict number of questions asked per forum
 - Users turn to other questions
 - Observation: Response time in some fora reduced

Backend Technologies

Indexing Distributed Semantic Graphs

- Examples

- Male persons who have a public profile document
- Computing science papers authored by social scientists
- American actors who are also politicians and are married to a model.

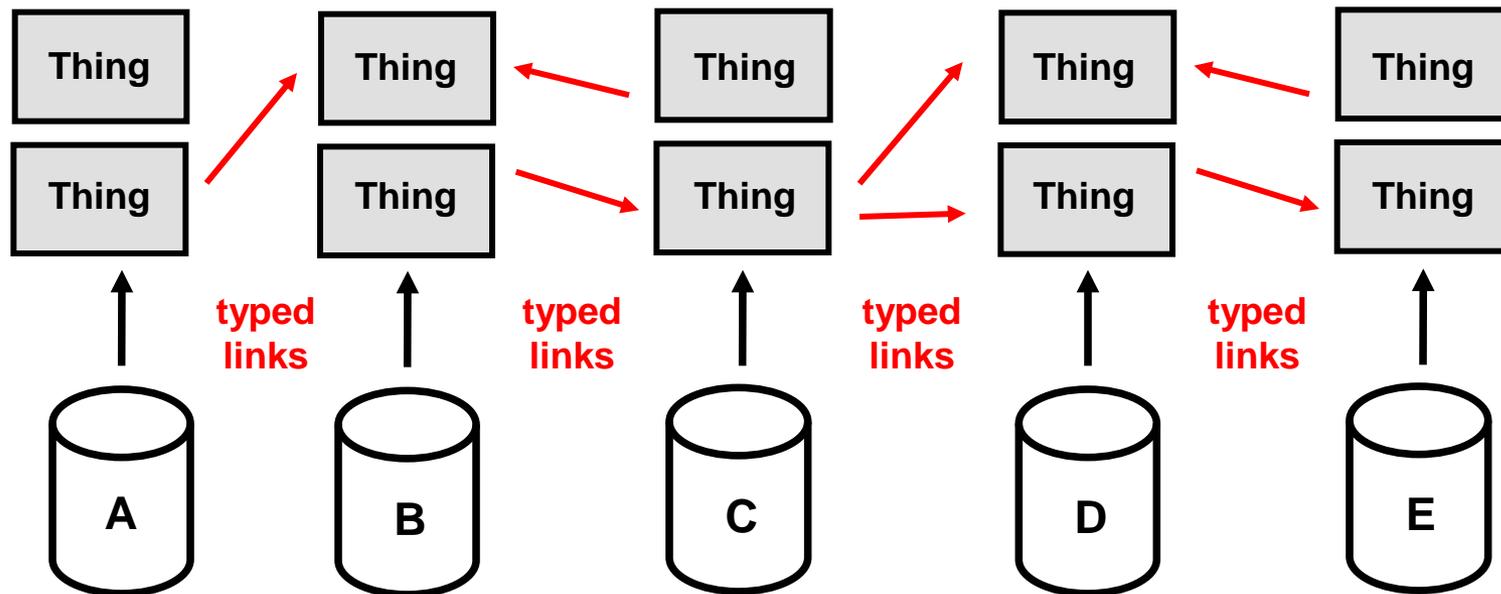
- Maybe specific databases available:

- Person search engines
- Bibliographic databases
- Movie database

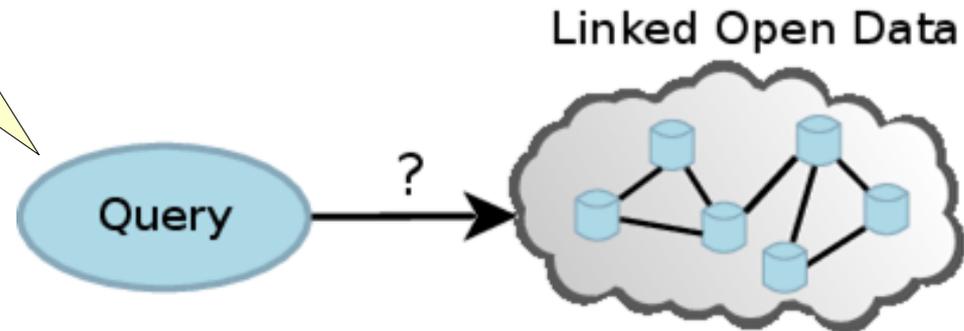
} How to integrate?

Semantic Web Technology to

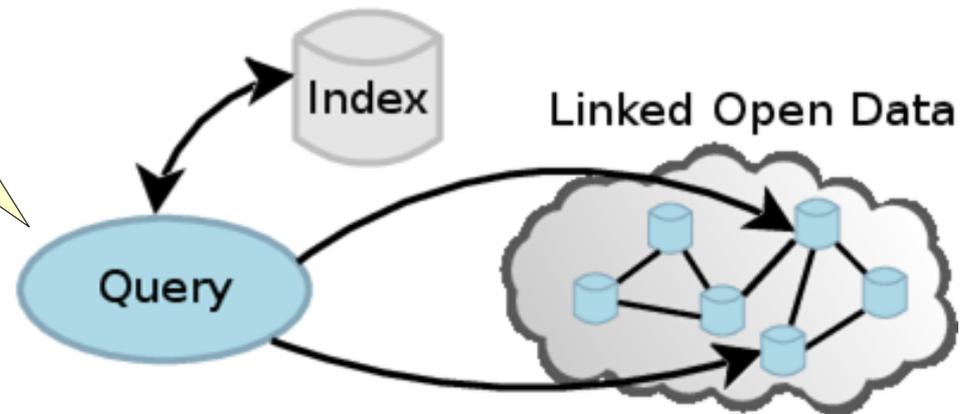
1. Provide structured data on the web
2. Link data across data sources

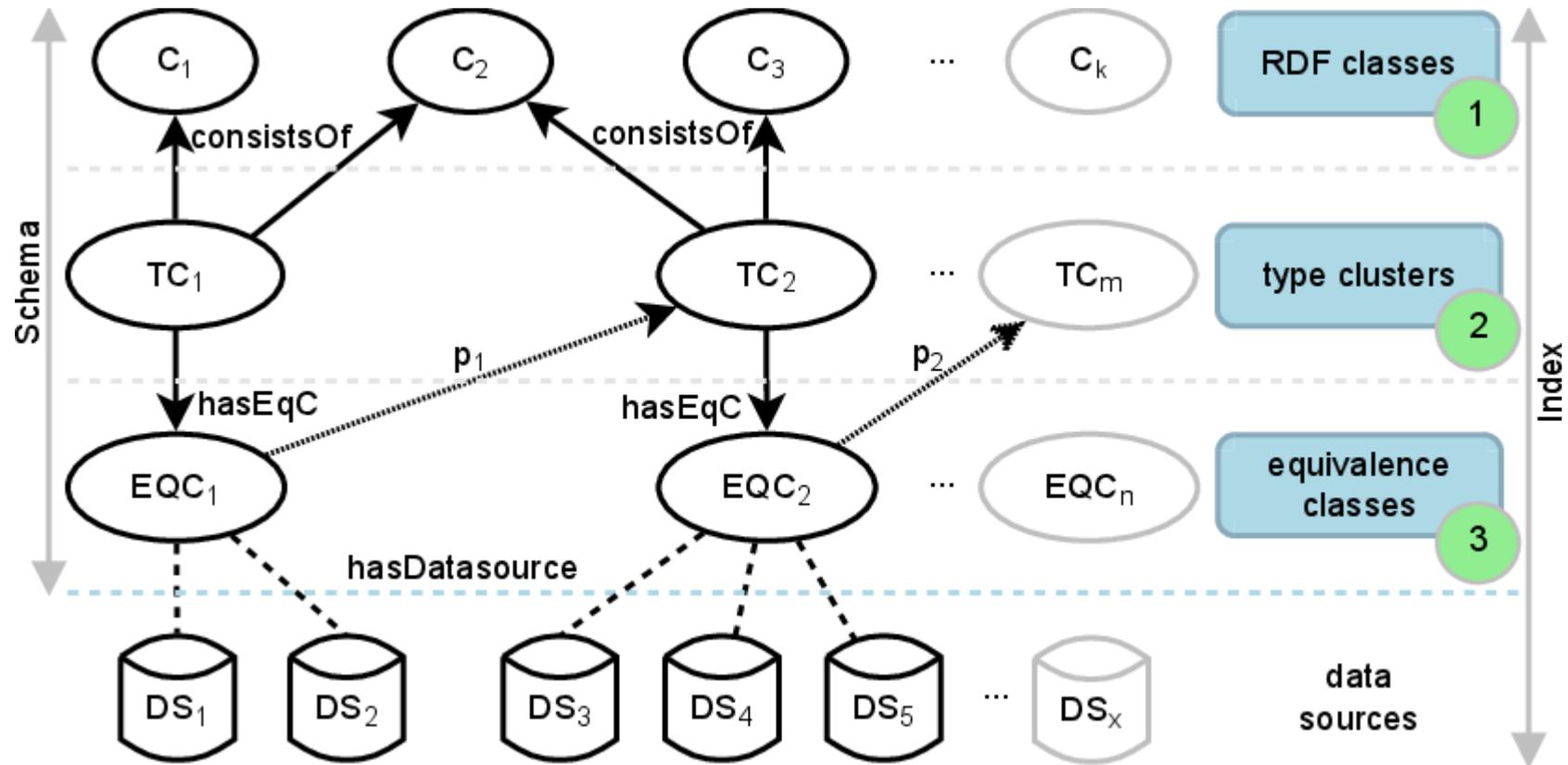



```
SELECT ?x
WHERE {
  ?x rdfs:type foaf:Person .
  ?x rdfs:type pim:Male .
  ?x foaf:maker ?y .
  ?y rdfs:type
    foaf:PersonalProfileDocument .
}
```

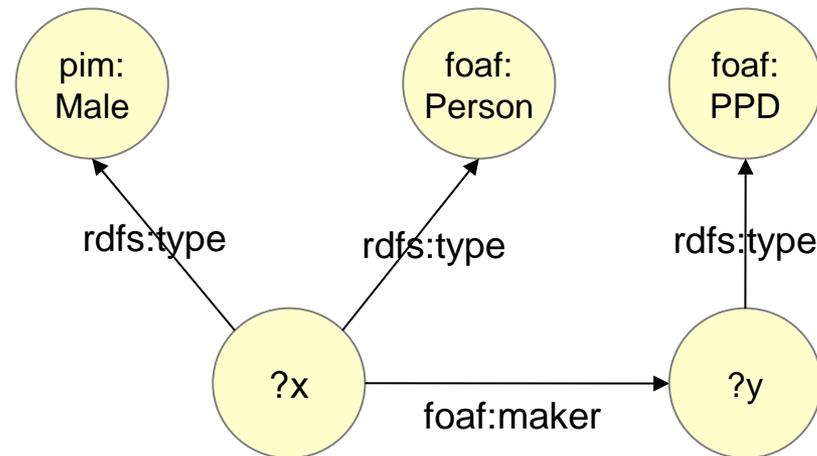
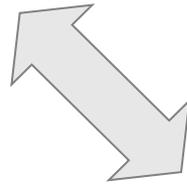


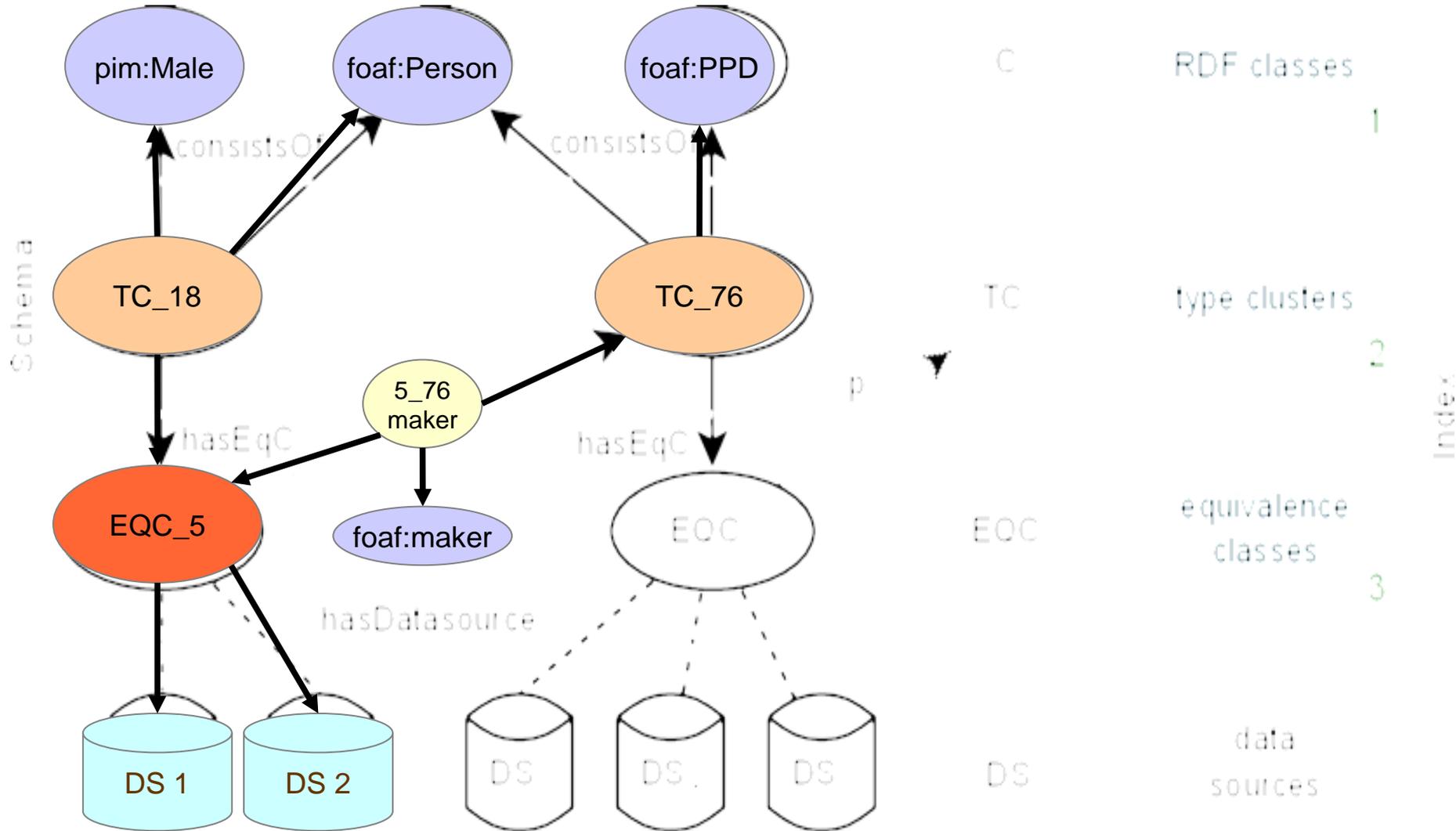
```
SELECT ?x
WHERE {
  ?x rdfs:type foaf:Person .
  ?x rdfs:type pim:Male .
  ?x foaf:maker ?y .
  ?y rdfs:type
    foaf:PersonalProfileDocument .
}
```





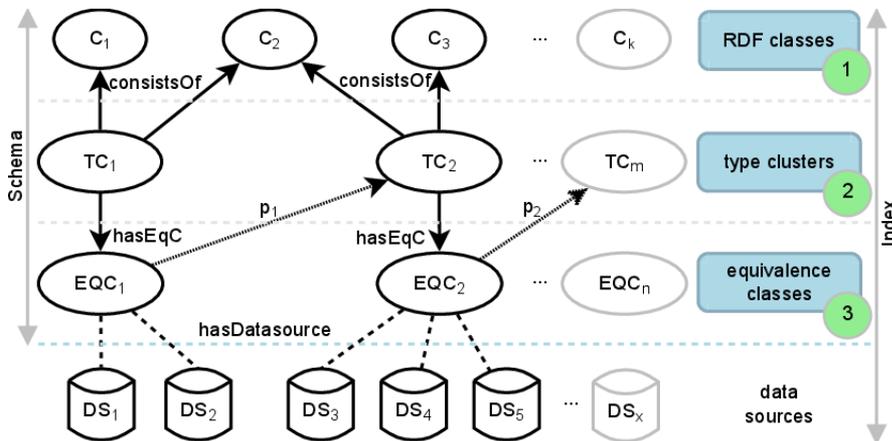
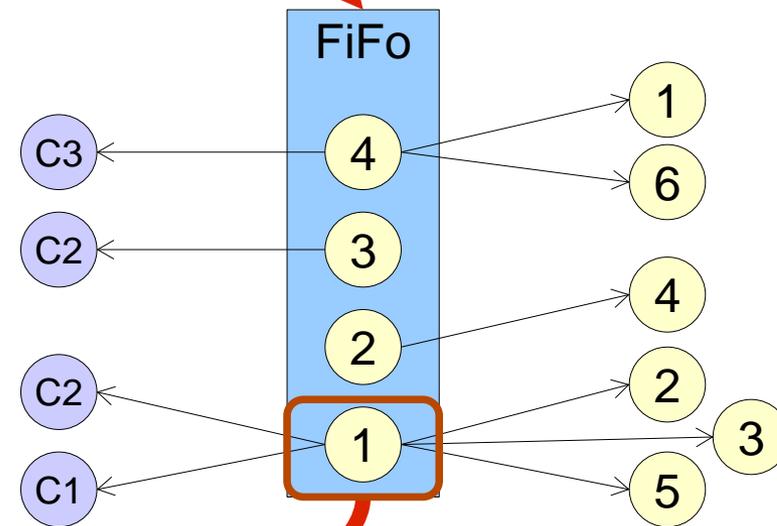
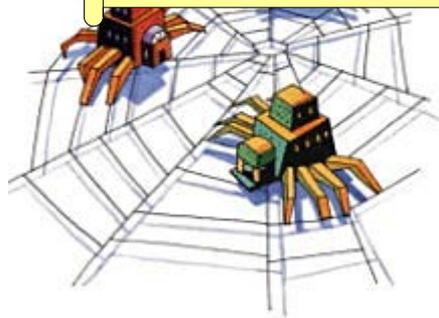
```
SELECT ?x
FROM ...
WHERE {
  ?x rdfs:type foaf:Person .
  ?x rdfs:type pim:Male .
  ?x foaf:maker ?y .
  ?y rdfs:type
    foaf:PersonalProfileDocument .
}
```





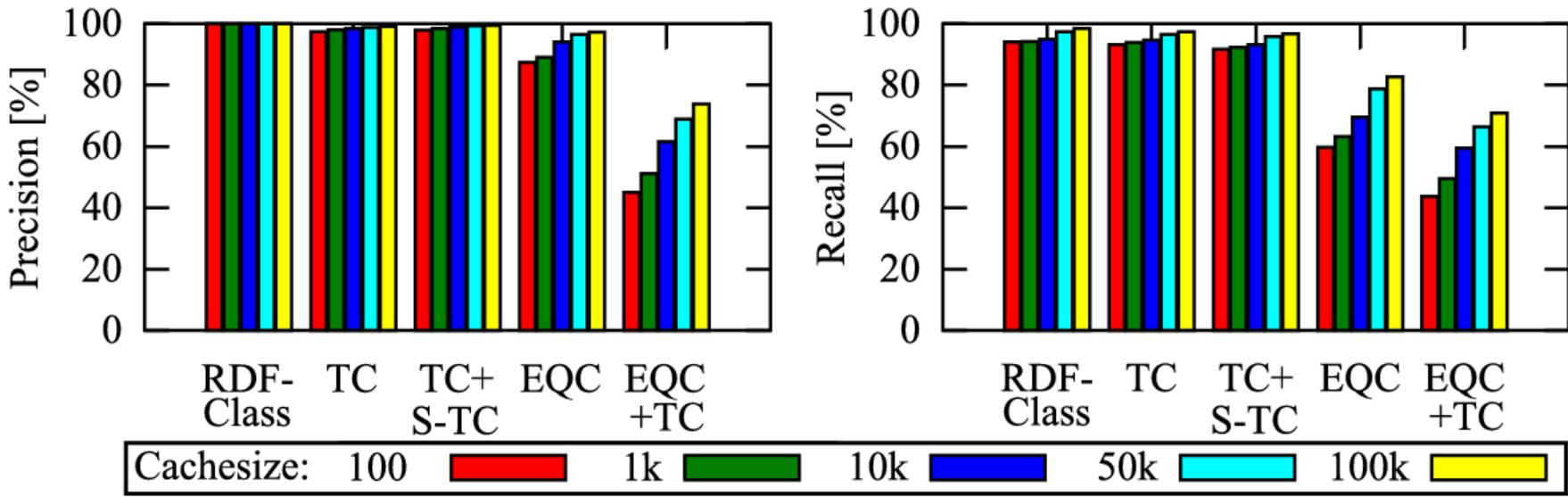
- Stream of data (coming from a LD crawler)

... D16, D15, D14, D13, D12, D11, D10, D9, D8, D7, D6, D5, D4, D3, D2, D1



Does it work good?

Comparison of stream based vs. Gold standard Schema on 11 M triple data set



Managing Business Communities

Lessons learned

- Business communities vary with regard to
 - Interaction
 - Interests
 - Type of conversation
- Novel analysis techniques needed:
 - Integration of different data sources
 - Simulation of policy changes
 - Value of users
- Concrete needs confirmed by project external companies looking for such technology

- *S. Angeletou, M. Rowe, H. Alani. **Modelling and analysis of user behaviour in online communities.*** In Proceedings of the 10th international conference on The semantic web - Volume Part I,
- *A. Che Alhadi, T. Gottron, J. Kunegis, N. Naveed. **Livetweet: Microblog retrieval based on interestingness.*** In TREC'11: Proceedings of the Text Retrieval Conference 2011, 2011.
- *A. Che Alhadi, T. Gottron, J. Kunegis, N. Naveed. **Livetweet: Monitoring and predicting interesting microblog posts.*** In ECIR'12: Proceedings of the 34th European Conference on Information Retrieval, 2012.
- *M. Karnstedt, M. Rowe, J. Chan, H. Alani, C. Hayes. **The effect of user features on churn in social networks.*** In WebSci '11: Proceedings of the 3rd International Conference on Web Science, 2011.
- *M. Konrath, T. Gottron, A. Scherp. **Schemex – web-scale indexed schema extraction of linked open data.*** In Semantic Web Challenge, Submission to the Billion Triple Track, 2011.
- *M. Konrath, T. Gottron, S. Staab, A. Scherp. **Schemex—efficient construction of a data catalogue by stream-based indexing of linked data.*** Journal of Web Semantics, 2012. to appear.
- *N. Naveed, T. Gottron, J. Kunegis, A. Che Alhadi. **Bad news travel fast: A content-based analysis of interestingness on twitter.*** In WebSci '11: Proceedings of the 3rd International Conference on Web Science, 2011.
- *N. Naveed, T. Gottron, J. Kunegis, A. Che Alhadi. **Searching microblogs: Coping with sparsity and document quality.*** In CIKM'11: Proceedings of 20th ACM Conference on Information and Knowledge Management, 2011.
- *F. Schwagereit, A. Scherp, S. Staab. **Survey on governance of user-generated content in web communities.*** In WebSci '11: Proceedings of the 3rd International Conference on Web Science, 2011.

Thank You!