

# FUSELABS

Future Social Experiences Labs

Lili Cheng

July 2012





Twigg

# Realtime Social Data

- August 2009: Twitter data: Public status messages
  - 12 million updates a day (20x our estimates)
  - Massive growth every month
- October 2010: Twitter data + Facebook public data
  - 500 million updates a day
  - 15,000 storage requests a second
- 2 guys + Windows Azure + Bing

# Link Expansion and Analysis

- Short Url expansion
- Content type
- Language
- Thumbnail
- Images
- Adult score
- Spam score
- (API aware)

Queen - Who Wants To Live Forever (High Quality) - YouTube

Who wants to live forever when love must die? Who waits forever anyway? Enjoy this single!

[www.youtube.com/watch?v=5L8-FTvSVxs](http://www.youtube.com/watch?v=5L8-FTvSVxs)



Queen - Who Wants To Live Forever (H...  
YouTube 4:55

**dannysullivan** Danny Sullivan  
interesting the auto-tune segment sparked more tweets than anything else at the #oscars -- including winners  
<http://tcm.ch/eRBkbl>



Link Expansion  
Metadata Extraction

Web

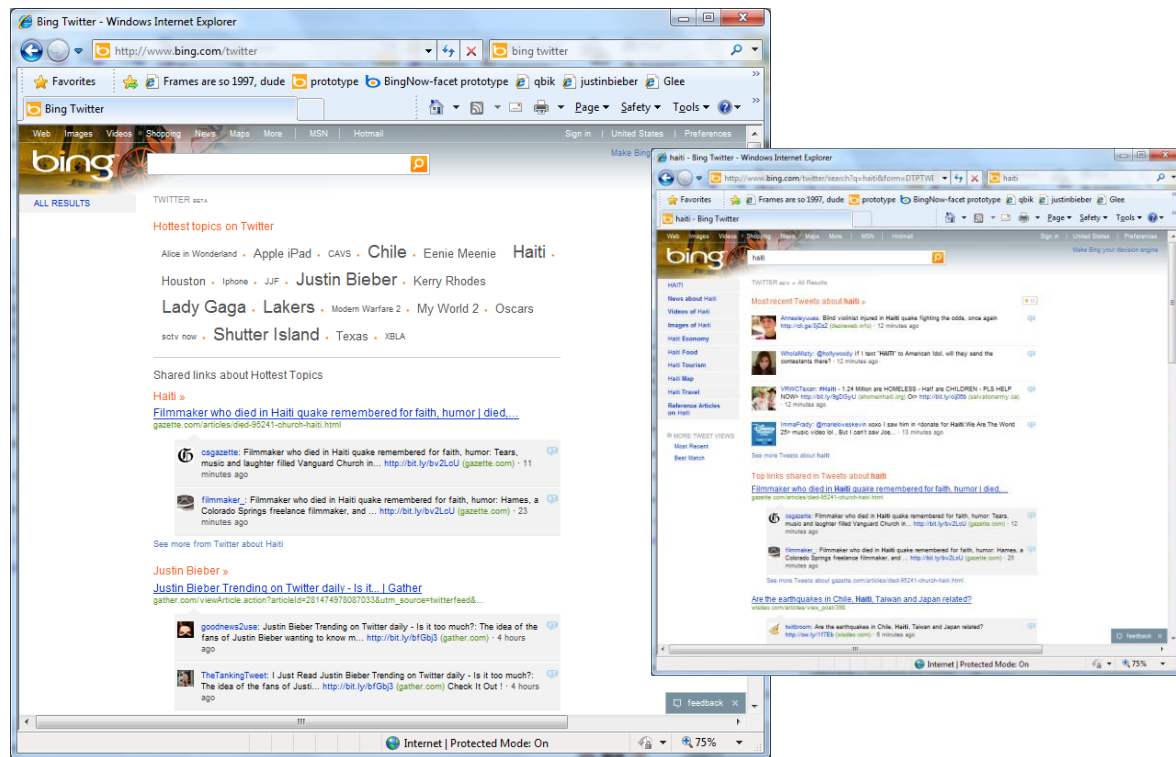
<http://trn.ch/eRBkbl>

← <http://someplace.com>



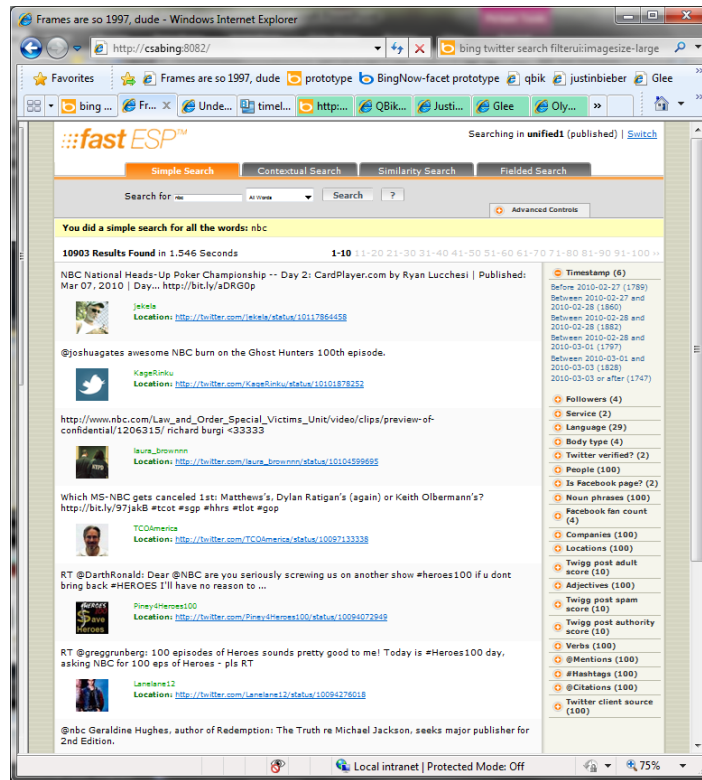
# Social Search

- Bing Now
- Twitter
- Facebook
- Bing Answers
- Indexing urls
- ....



# ...thanks FAST!

- Language detection
- Parts of speech
  - Nouns, verbs
  - Questions
- Entity extraction
  - People
  - Places
  - Topics
- Custom filters



# Social Apps and Components



USING THESE COMPONENTS

Personalize

Connect

Store

## Profile Service

- Multiple Logins
- Single Profile
- Customizable UI
- Simple REST access



## Recommendation Service



## jSearch

- Javascript based search library
- Normalizes all sources to a common schema
- Download only the sources you use
- Easily extended to support new sources
- Support RX.js async observable pattern
- Sources include RSS, OData, OpenGraph

## URL Analysis

- Classification:** adult, article audio bookmark, comment, event, file, folder, group, language, list, note, person, photo, album, place, playlist, product, review, service, status, video
- Extract meta-data from OpenGraph and MicroFormats
- Handles URL shorteners like bit.ly
- Generates web page thumbnails
- Extract media
- Many Sources: Facebook, YouTube, Bing, Google Maps, etc.

[danmyaufrin.com/2013/04/10/interesting-the-audio-tune-segment-sparked-more-interest-than-anything-else-at-the-#iwcars-including-winners](http://danmyaufrin.com/2013/04/10/interesting-the-audio-tune-segment-sparked-more-interest-than-anything-else-at-the-#iwcars-including-winners)  
<http://item.chicofeb08d>

## Distributed Search

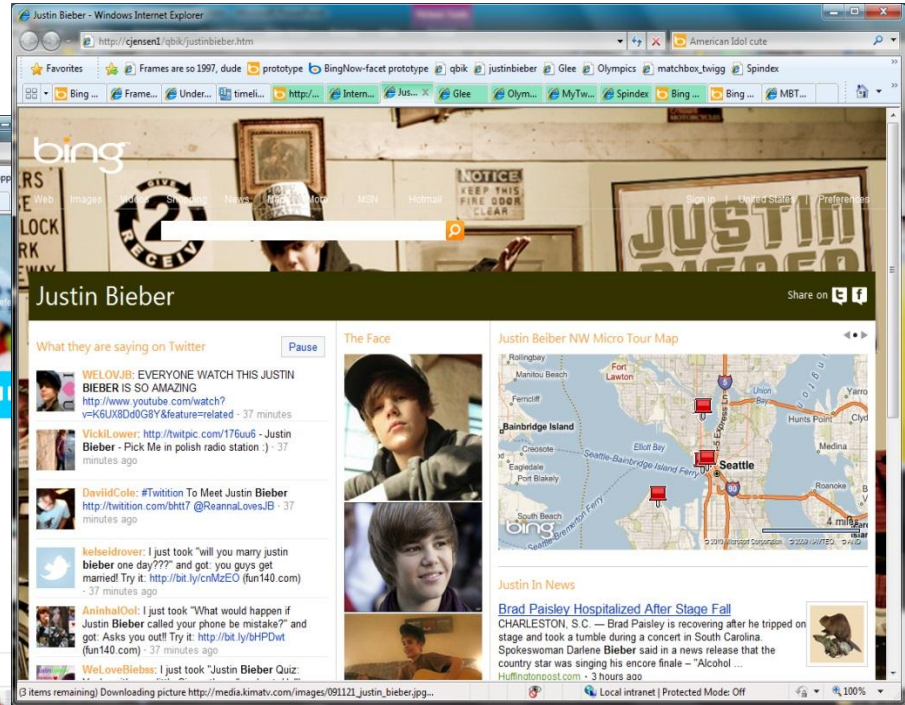
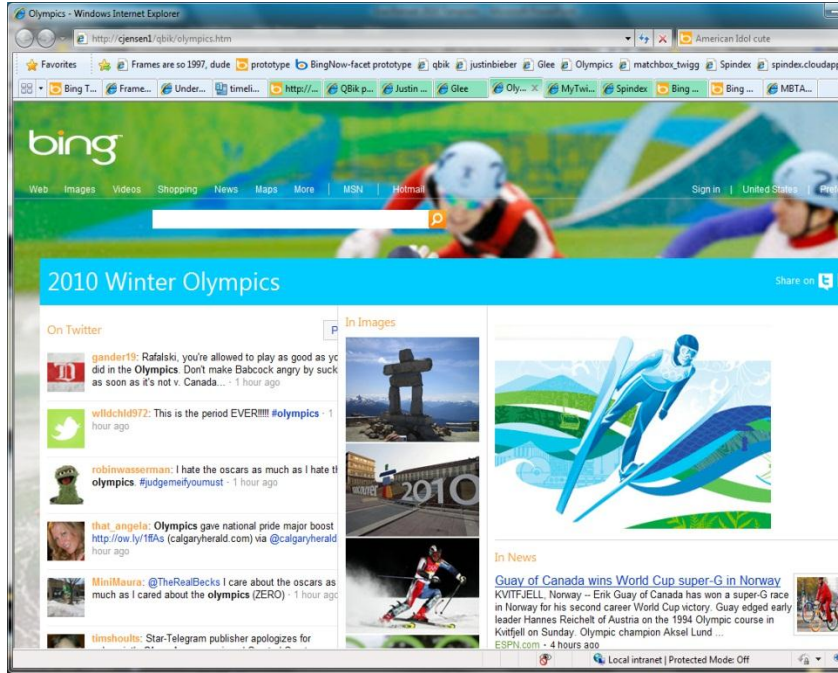
- REST API Store & Retrieve JSON
- Optional Schema with validation
- Full-text and relational search



## eSearch : Embedded Search & Storage Library

- Robust, Pluggable Item Storage
- 14 simple data types and vectors
- Relational:  $<$ ,  $=$ ,  $>$ ,  $\geq$ ,  $bits$
- Text: contains, near, phrase
- Ranking: explicit, hits and IR
- Multi-level sorting

# Authoring. Montage.

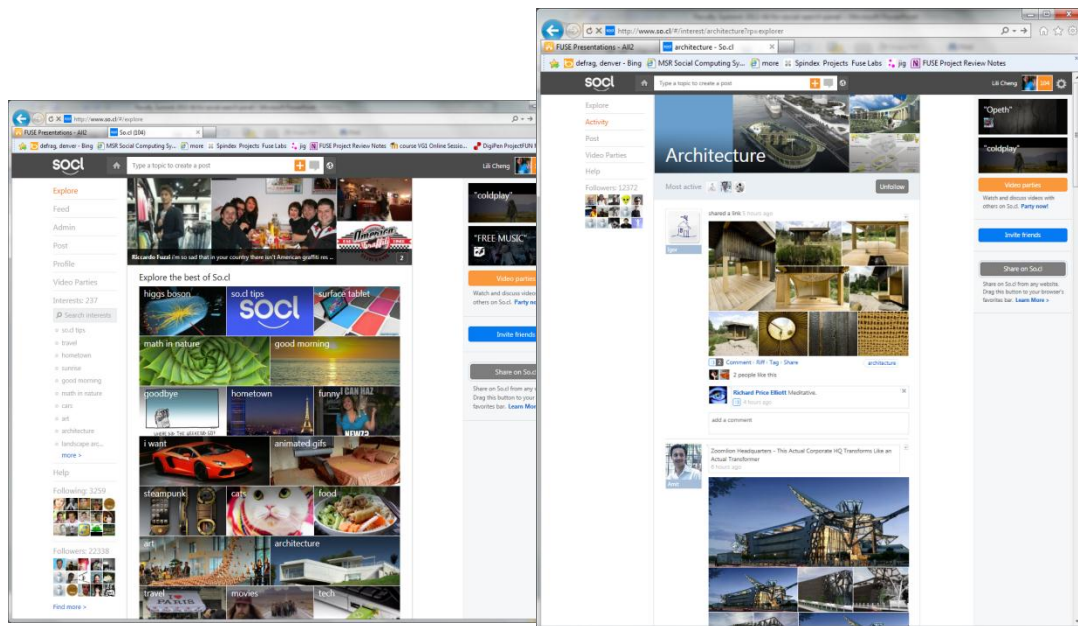
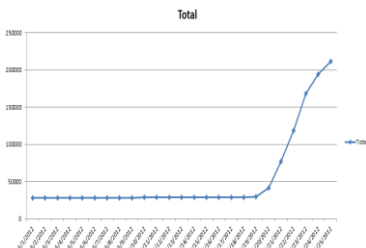




# So.cl

*Use search to express & share your ideas*

- So.cl
  - Integrates Search & Social
  - Data reuse
    - Bing API's
    - So.cl Public data & API's
- About 300,000 users



RECENT



## Springtime in the Southwest



**Dan Marshall** I love the springtime in the Southwest US. The blooming flowers speckle and turn it into an amazing place. 2 min ago

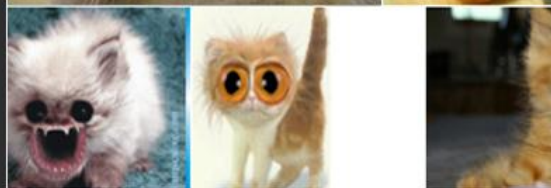


**Dan Marshall** You always post the most beautiful images. I love the southwest. 2 min ago



**Tom Jones** I love that area too 2 min ago

27 comments



## Crazy Cats



**Dan Marshall** Searched 2 min ago



## Carrots



**Dan Marshall** Nothing like the orange bits 2 min ago

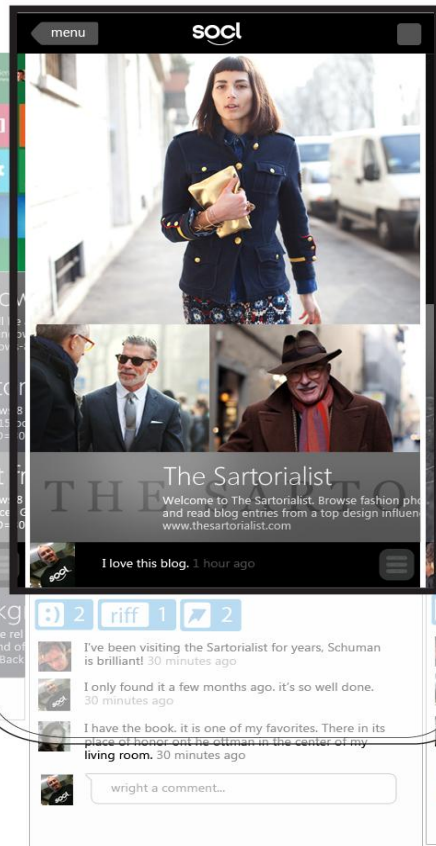
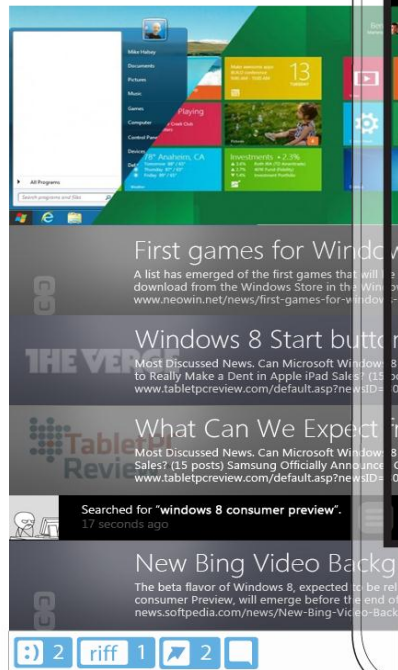
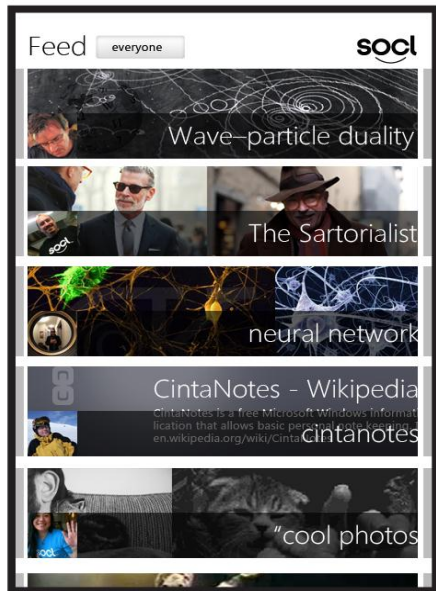


**Dan Marshall** You always post the most beautiful images. I love the southwest. 2 min ago



**Tom Jones** I love that area too 2 min ago

27 comments



# Usage: Creating, Socializing, Browsing

Factor Analysis for Associated Behaviors:

Three types of usage – creating, socializing, browsing

Component Matrix <sup>a</sup>			
Type: % Variance:	Creators 32%	Component Socialites 12%	Browsers 9%
Created post	.86	.17	.10
Invited	.01	-.16	.63
Followed	-.03	.10	.37
Added item to post	.83	.08	-.06
Searched	.81	.03	.17
Commented	.36	.64	.09
Liked post	.15	.58	.32
Liked comment	.13	.80	.06
Messaged	-.09	.50	-.08
Viewed person	.22	.47	.48
Navigated to All	.51	.37	.53
Joined party	.17	.09	.68

Factors about equally predict if user comes back

Regression Coefficients

	Beta	t	Sig
Creating	.14	5.28	.000
Socializing	.07	2.61	.000
Browsing	.19	7.20	.000

Browsing stronger predictor of overall activity level

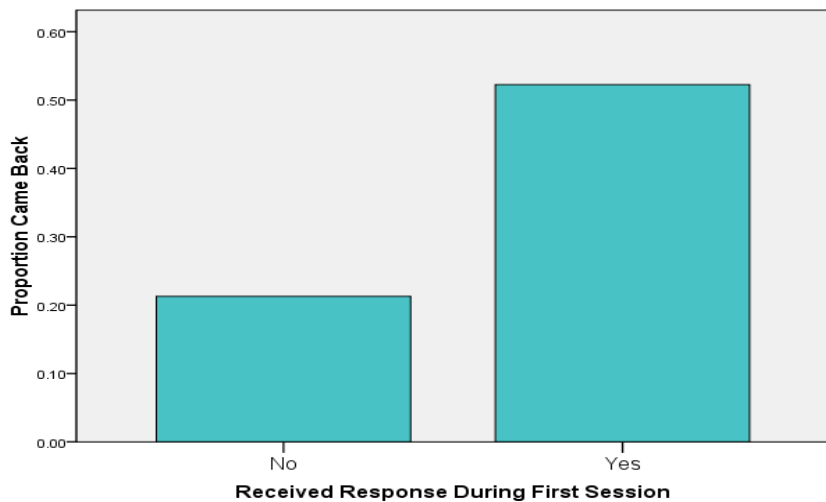
Regression Coefficients

	Beta	t	Sig
Creating	0.20	7.89	0.00
Socializing	0.17	6.58	0.00
Browsing	0.29	9.07	0.00



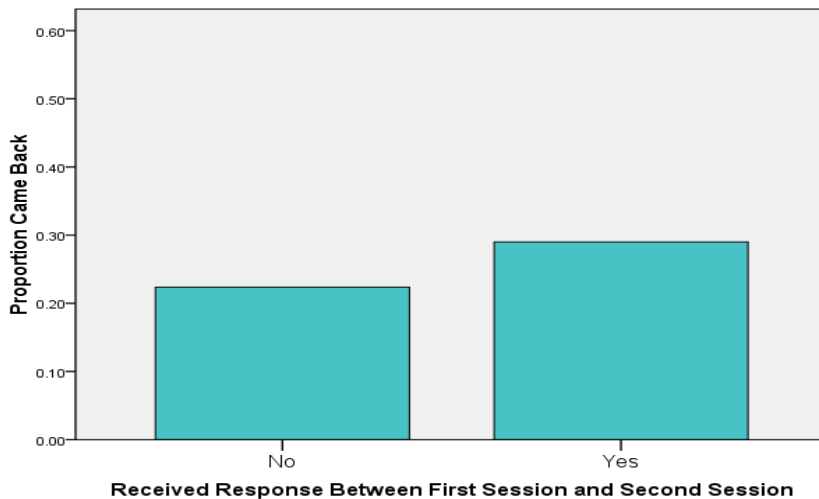
# Usage: Return Visits

- Social responses inspire people to return to site, especially if occurring *during first session*



N = 1268

N = 111



N = 1083

N = 293

Social responses to user: following, commenting, liking post, liking comment, riffing

# Lessons Learned

- Social networking feed:

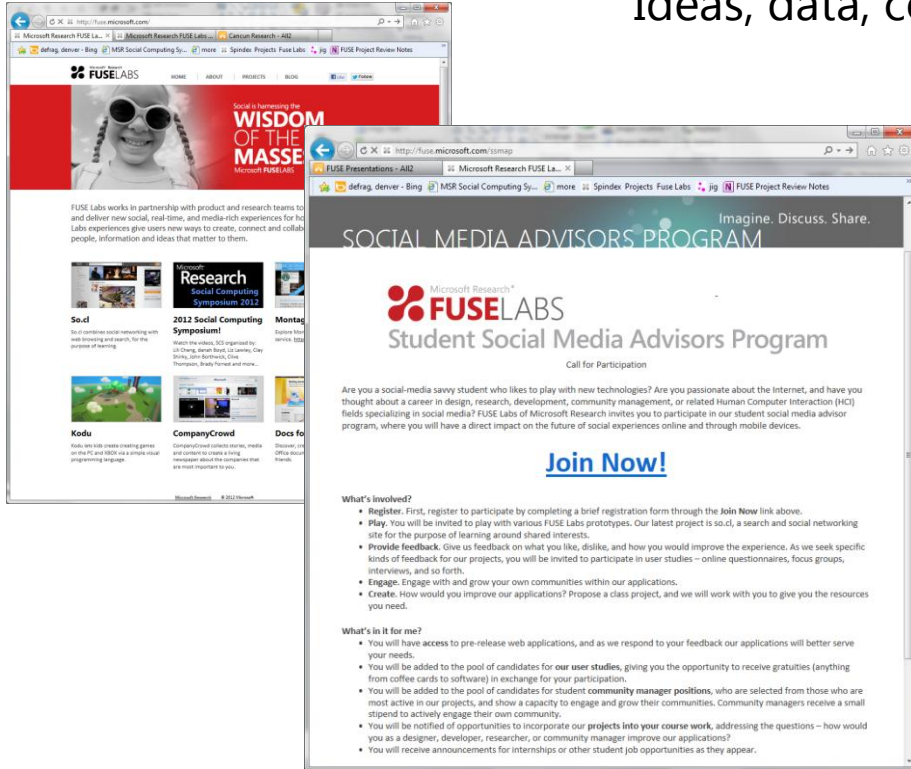
- Appropriate “flow of data” primary part of social user experience
- Critical mass of activity & “good users” define site norms and personality
- Discovery of new people and topics a person cares about key

- Social across “boundaries”:

- Across identities:
- Across time: interaction history creates rich profile over time
- Reuse of public data

# <http://fuse.microsoft.com/ssmap>

Ideas, data, community, apps



## Faculty Summit Day 3!

iSchool Social Media Research

Student Social Media Advisors Program

<http://fuse.microsoft.com/ssmap>

## Contacts:

Shelly Farnham [shellyf@microsoft.com](mailto:shellyf@microsoft.com)

Lee Dirks [ldirks@microsoft.com](mailto:ldirks@microsoft.com)

Lili Cheng [lilich@microsoft.com](mailto:lilich@microsoft.com)

[lilich@microsoft.com](mailto:lilich@microsoft.com)  
@lilich

<http://fuse.microsoft.com>  
<http://so.cl>