Structured data from microblogs

- Semantic Web
  - typically published in large chunks
  - relatively static

- microblogs
  - published in tiny chunks
  - 600+ per second
  - is any of this data useful?
  - what can we capture?
Data about people

- users and accounts
- use FOAF and SIOC
- e.g. SemanticTweet (Flinter)
Data about tweets

- blogs and posts
- use SIOC and SIOC Types
- e.g. SMOB (Passant et al.)
What about data in a tweet?

- **nanotations**: structured information in microblog posts
TwitLogic

- capture microblog data...
- ...including nanotations
- translate it into RDF
- produce an RDF stream
- ... and expose it as Linked Data
- search and query in real time
For example...

#sioclog (see http://bit.ly/2uAWo2) makes Linked Data from IRC logs.

Who would have guessed such a funny movie as #ZombieLand (3/4) could be made around zombies?

#websci2010 (= #websci10) is all atwitter with presentations containing the word "tweet".
A demo

Semantic Conference Tweets
A demo by Zhenning Shangguan, Joshua Shinavier and Jin Guang Zheng

Tweets about http://data.semanticweb.org/conference/iswc2009

- rdf:type
  - owl:Thing
  - swc:ConferenceEvent
  - swc:ConferenceEvent

- owl:sameAs

- foaf:based_near

- swc:completeGraph

fuzzy.com
RT @tommyh: @ldodds' US #linkeddata overlays from the #iswc2009 #linkeddatatathon gets a mention in #www2010 consuming #linkeddata tutorial
da day ago

Joshua Antishinavier
Bogus #iswc2009 tweet, to test @shangz and Jin's new TwitLogic demo interface.
3 days ago

Jie Bao
Twitter brain storming #3: use cases (find all #iswc2009 tweets; twits about don jones drop)
5 months ago
More details

- semantic nanoformats
- publication and data ownership
- implementation
- real-time, real-place semantic search
- future work:
  - continuous queries
  - **Twitter Annotations**
Questions...
Links

• TwitLogic
  • http://twitlogic.fortytwo.net

• Tetherless World
  • http://tw.rpi.edu

• Franz Inc.
  • http://www.franz.com

• Contact
  • josh@fortytwo.net
  • @joshsh