Self-organized Semantic NeLI

Through Self-organized Approaches Using Semantic Web, Social Networking and Web 2.0 Frameworks

M. O. Faruque Sarker

PhD Candidate
Cognitive Robotics Research Centre
Newport Business School
University of Wales, Newport

2 November 2010

Outline

Background: Web 2.0, Semantic web and Self-organization

Why self-organization is relevant?

- Web 2.0 is based on bottom-up processes
  $\longrightarrow$ users generate contents through explicit and implicit social interactions.
- Complex patterns can be resulted from simple behavioural rules
  $\longrightarrow$ complexity is not necessarily outcome of sophisticated agent cognition

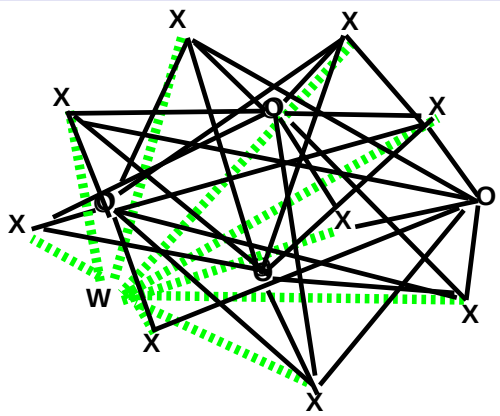
A Self-organized approach

- Pattern formation through the local interactions internal to the system [a]
- Needs four elements:[b]
  i) continuous flow of information, ii) concurrence, iii) learning and iv) forgetting of agents
- Scalable, adaptable and robust solutions
  $\longrightarrow$ Depends on less communication/computation, almost no user/environment modelling

[a] Camazine *et al.*, Self-organization in Biological Systems, 2001.
[b] Sarker & Dahl. *LNCS* 6234, 2010.

Figure: A termite nest from bottom-up approach

Self-organized Semantic Search through Attractive Field Model



Source nodes (o)	items to be searched
Agent nodes (x)	user search agents e.g., mozilla browser
Black solid edges	attractive fields that correspond to an agent's stimuli for related items
Green edges	attractive fields of unrelated items shown as (w)

Figure: The attractive filed model (AFM)

Strength of an attractive field: $S_j^i \propto \frac{k_j^i}{d_{ij}} \phi_j$

Parameters

k_j^i i agent's sensitization for item j
$\longrightarrow$ How much i is sensitive to j item?

d_{ij}: i agent's location preference
$\longrightarrow$ How far j is located from i?

ϕ_j: relative urgencies of item j.
$\longrightarrow$ What is the level of urgency of j?

Learning search patterns

E.g.: Increasing item sensitization of agents by a rate k_{INC}

$$If\ item\ is\ rated: \ k_j^i \rightarrow k_j^i + k_{INC}$$

$\longrightarrow$ A form of positive feedback

A Frameworks for Semantic Search and Navigation

Key aspects

- Easy and interactive, compatible, support for different users' personalization
- Reasoning new information from existing information

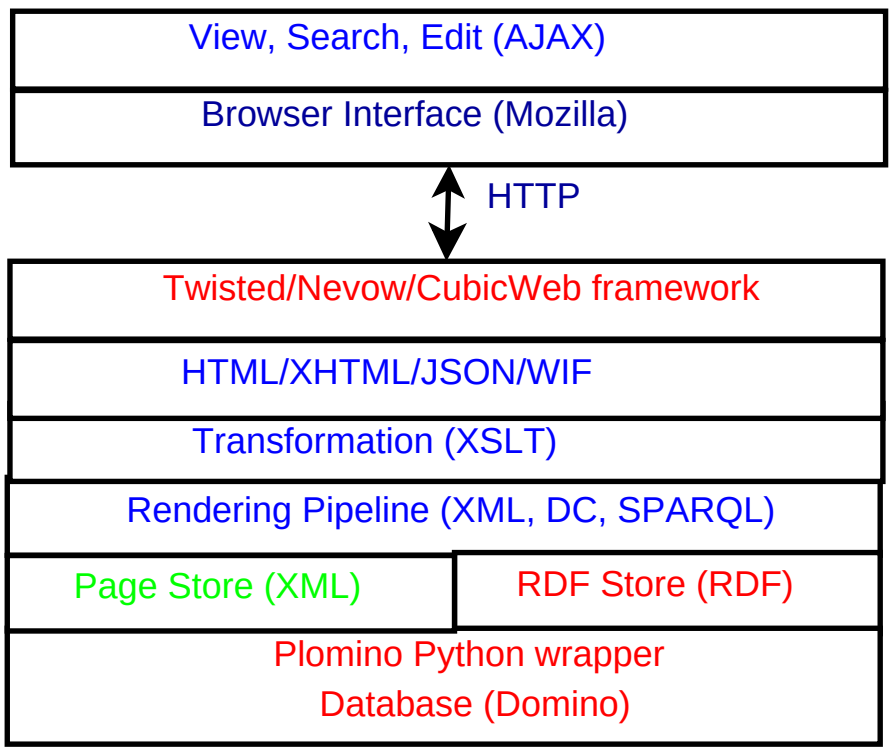


Figure: Python based semantic wiki framework

Model	Tools/Frameworks
Database storage	Domino, Plomino.dominoimport
Webpage /Knowl-edge base	XML, Resource Description Framework, RDF/XML
Data rendering	XML, Dublin Core, SPARQL Protocol and RDF Query Language
Page transfor-mation	XML Stylesheet XSLT
Web publishing	HTML/XHTML, JavaScript Object Notation, Wiki Ex-change Format
Backend	Python web frameworks
Frontend	JavaScript/AJAX

Web 2.0 in NeLI: Wiki, Blogs, RSS, Podcast and more

Adopting web 2.0 in NeLI

- Blogs: One-to-many interactive timely discussions → e.g. Dimov's *Clinical Cases & Images*[a]
- Wikis: Bottom-up content creation, editing and discussions → e.g. Flu Wiki[b]
- RSS Feeds: Up-to-date feeds on mobile devices *anytime anywhere*
- Social Bookmarking: Collaborative tagging, link sharing → e.g. IBM's Dogear
- Podcast or Vodcast: Audio-visual material for increasing awareness

[a] http://casesblog.blogspot.com/
[b] http://www.fluwikie.com/

Constraints

- Lack of authoritative control over content
- Lack of accuracy
- Information overload
- Anonymity
- Requires moderating → e.g. Ganfyd[a]
- Roll-back is costly
- Requires enforcing policies for ensuring privacy and copyright issues

[a] http://www.ganfyd.org/

A System for Exploiting Social Networks (SN)

Mining SN

- Relevant page identification, pre-processing and extraction
  $\longrightarrow$ Facebook/Twitter API
- Integration
  $\longrightarrow$ Probabilistic approach for name disambiguation, classification
- Database storage, indexing and access
  $\longrightarrow$ Domino, Python wrapper
- Social network modelling
  $\longrightarrow$ semantic modelling, clustering using statistical and socio-biological models
- Search services
  $\longrightarrow$ Hot topic, disease outbreak, expert search

Disseminating public health information over SN

- Automated delivery based on SN clusters/ subscription
- Relevant information (recall self-organized search)
- Low overhead of information based on user feedback

A Framework for Infectious Disease Monitoring through SN

HealthMap: Global Infectious Disease Monitoring Site

- Popular news-based
  ⟶ alerts from multi-sources, last 30 day stats
- web 1.0 framework
  ⟶ no metadata storage, no knowledge-base, needs human intervention

Going Beyond HealthMap

- Social-network based:
  ⟶ biological & statistical models
- Semantic web technology:
  ⟶ inferred knowledge
- Personalized alerts
  ⟶ over Facebook, Twitter, mobile devices
- Self-organized design and operation
  ⟶ scalable, minimum site maintenance overhead, fault-tolerant and mostly automated

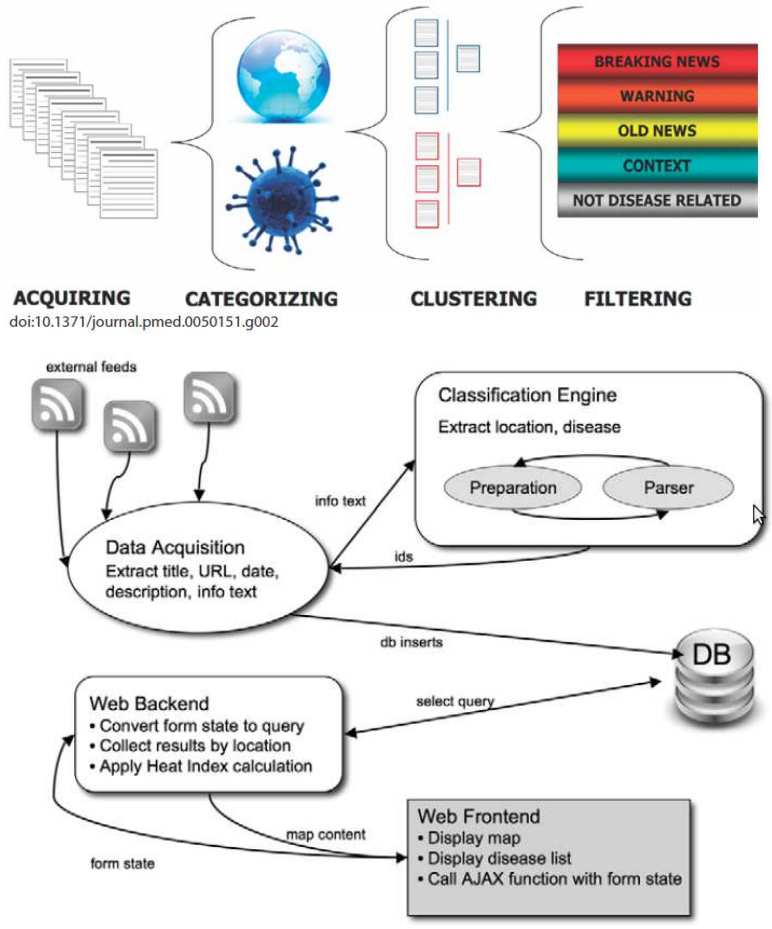


Figure: HealthMap architecture (Freifeld et al. *J. of the American Medical Informatics Association*:15, 150, 2008)

Conclusion and Outlook

- Web 2.0 is based on bottom-up processes
- Self-organized approaches has potential to manage NeLI's large semantic web system
- Both biological and statistical approaches can be used to exploit the social networking tools for data mining and information dissemination
- NeLI's infectious disease monitoring framework can potentially be built by putting semantic web technology into existing frameworks.