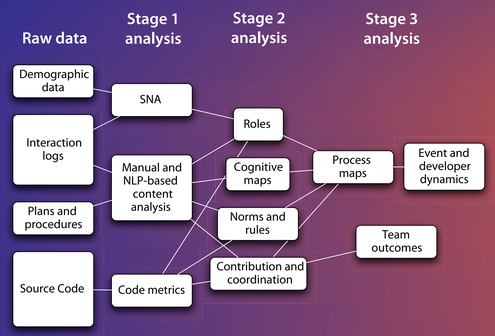
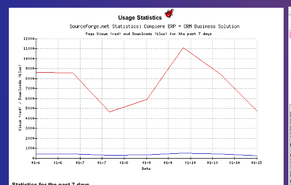
Investigating the Dynamics of FLOSS Development Teams

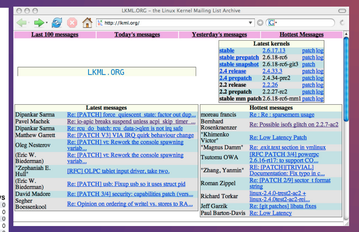
What are the dynamics through which self-organizing, distributed teams develop and work?
Research partially funded by NSF grant 05–27457, with prior support from 03–41475 & 04–14468

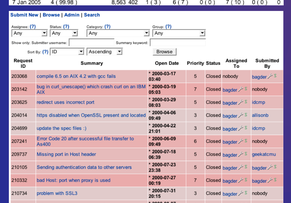
Analysis Plan

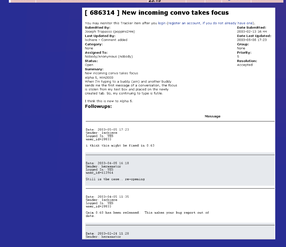


Raw Data









Data Integration

People with their events

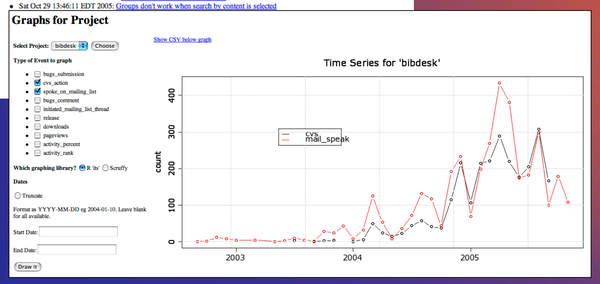
- Person ID: 48 (identified by sf_user_id: **993558**, real_name: **Jeff Johnson**, sf_username: **johnson1234**, email: **jcjohnso@wisc.edu**)

Participated in 34 events (of all kinds)

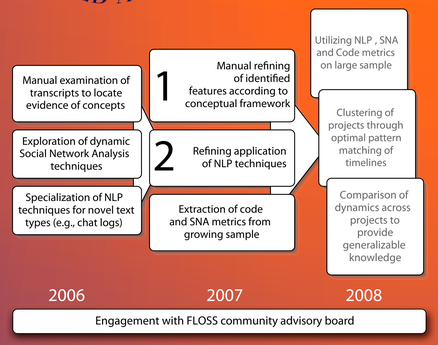
- bugs_submission: 22
- bugs_comment: 12
- release: 0
- cvs_action: 0
- initiated_mailing_list_thread: 0
- spoke_on_mailing_list: 0

Events in datetime order:

- Mon Mar 08 22:01:13 EST 2004: 1053361 on bugs 912358
- Tue Jun 21 10:33:12 EDT 2005: Accents in macros not displayed
- Tue Jun 21 10:40:03 EDT 2005: Crashed on changing Preview Pane preference
- Tue Jun 21 16:32:26 EDT 2005: [682448] on bugs 1224896
- Thu Jul 07 21:40:07 EDT 2005: TypeInfo Editor doesn't quit edit mode
- Thu Jul 07 22:07:01 EDT 2005: Author sorting for two-word last names
- Sat Jul 09 16:03:43 EDT 2005: [699934 on bugs 1234554
- Sat Jul 23 18:19:46 EDT 2005: Save changes dialog with no changes
- Wed Aug 31 20:39:21 EDT 2005: BibDesk 1.1.5 hangs on New Publication with Crossref
- Wed Aug 31 20:53:58 EDT 2005: BibDesk crashes when entering "{" character in raw BibTeX
- Tue Sep 06 11:31:35 EDT 2005: Gaps in list sorted by Booktitle
- Thu Sep 29 09:52:36 EDT 2005: Blank Container fields appear first in sorted list
- Wed Oct 05 19:32:13 EDT 2005: Preview pane stops updating after cancelled Close
- Sat Oct 08 20:16:37 EDT 2005: 1791318 on bugs 1308021
- Thu Oct 13 16:46:20 EDT 2005: Sorting not fully sticky
- Thu Oct 20 12:25:49 EDT 2005: Width of new Keywords/Group pane not sticky
- Mon Oct 24 16:51:47 EDT 2005: Extraneous arrows above scroll bar
- Sat Oct 29 13:46:11 EDT 2005: Groups don't work when search by content is selected



Research Plan



Key Findings To Date

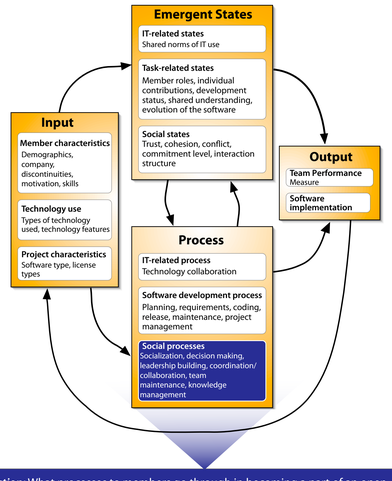
- We have developed National Language Processing (NLP) text extraction rules to reliably identify 63-72% of decision triggers, and 71-85% of decision announcements.
- There is a positive relationship between the level of participation and group effectiveness, with effective projects exhibiting high involvement from peripheral members.
- Some open-source projects that appear decentralized are actually centralized, but with individuals at the center changing their roles over time.
- In studying shared-mental models, norms and rules are highly shared within groups. However, it seems that tenure in projects can cause those to change within individuals, indicating dynamic processes over time.
- Second-level leadership in open-source teams is enabled by first-level leadership. First-level includes task coordination and contribution, group maintenance, and boundary spanning; while second-level is behavior that influences changes in structure that guides group action.

1 IMOI model of FLOSS literature review

We are reviewing the research on FLOSS development and use in an effort to assess the state of the literature. We have reviewed 586 papers in the first round, which were collected from Web of Science, ABI/Inform, journal special issues, AOM and AIS conferences, IntOSS (IFIP) conferences, ICSE workshops and opensource.mit.edu.

We are currently organizing the review around an input-mediator-output-input (IMOI) model. The 295 papers that were found both empirical and relevant are being categorized into issues pertaining to

- inputs (member characteristics, technology use and project characteristics)
- processes (social processes, software development processes and IT-related processes)
- emergent states (social states, task related states and IT-related states)
- output (team performance and FLOSS implementation)



Socialization: What processes to members go through in becoming a part of an open-source team, and how are different levels of membership defined?
Decision Making: How are decisions made in matrix environment such as open-source teams where leadership is not always clearly defined?
Leadership Building: How is leadership defined in open-source teams, and how do members rise into such positions?
Coordination/Collaboration: What are the dynamics of interaction within open-source teams where roles are not always clearly defined?
Team Maintenance: How do open-source teams form, and what keeps these teams together?
Knowledge Management: Where do open-source teams store their group knowledge, and how is it accessed by members?
Trust: How is trust formed within open-source teams?
Conflict: What types of conflicts arise within open-source teams, and how are such conflicts resolved?
Level of Commitment: Considering the often voluntary nature of open-source teams, what amounts of commitment do members invest?

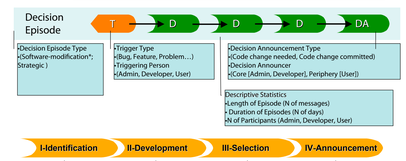
1 Decision making processes in FLOSS teams

We analyzed the email listservs and fora for 6 FLOSS project teams using Decision Episodes as the unit of analysis. The analysis data consist of 360 Decision Episodes from the beginning, middle, and end periods of email interactions.

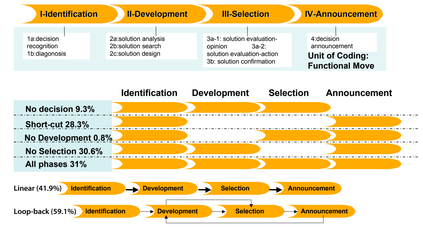
Case Selection

Product Complexity	Intended Users	Team Success	
		More Successful—>*Less Successful*	
Instant Messaging Client (IM)	Company	Gaim	aMSN	Fire
Enterprise Resource Planning System (ERP)	Individual	Compiere	OFbiz	webERP

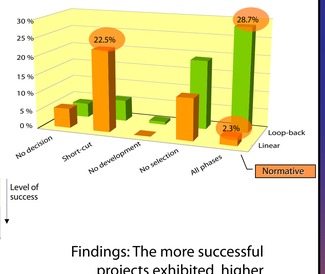
Coding Scheme I Decision Episode



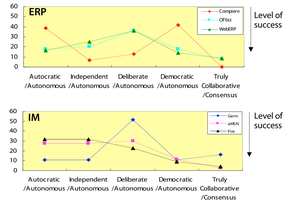
Coding Scheme II Normative Decision Making Process



Findings: Generally, the decision making processes in the open source projects are not normative.



Decision Styles

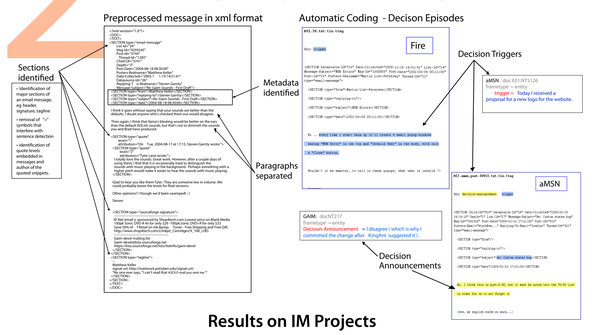


Findings: The more successful projects exhibited higher peripheral involvement in their decision making processes than the less successful projects.

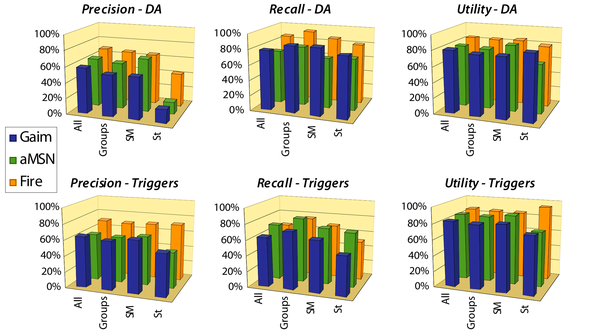
2 Natural language processing for content analysis

Automating manual coding processes

After preprocessing, manual processes were automated by using manual coding as the Gold Standard
- IM projects data were analyzed for decision announcements and triggers with the sentence as a unit of analysis
- Manual coding was validated by 2 or 3 coders, and used as training data for rule-writing and evaluation of results
- Rules utilize NLP features such as morphology, word form, syntax, semantics, discourse, and pragmatics



Results on IM Projects

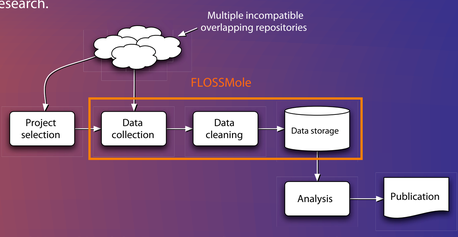


Key
All - analysis by each marked sentence
Group - analysis by each decision group of sentences
SM - software modification decision episodes
St - strategic decision episodes
DA - Decision Announcements

Metrics
Precision - % marked that agree with Gold Standard
Recall - % of Gold Standard that are marked
Utility - % marked that assist researcher in understanding Decision Episodes

Products

- floss.syr.edu website with copies of papers and analyses
- FLOSSmole, a collaborative project to gather, share and store comparable data and analyses of free and open source software development for academic research.



Plans

- **Shared Mental Models:** Transcripts of 27 interviews with members of the open-source community are being analyzed through cognitive mapping, and the integration of Natural Language Processing is being considered. Results to date indicate a degree of sharing of mental models, and we are looking for casual links within the community.
- **Leadership:** A detailed coding scheme has been developed that considers the concepts of first and second order leadership, including respective indicators. This code book is currently in the process of validation and the subsequent content analysis is planned for inclusion in the doctoral dissertation of one of our graduate assistants.
- **Group Maintenance:** A detailed coding scheme drawing on politeness theory, organizational citizenship behavior, and social presence theory literature has been developed. The code book is being validated and refined through a pilot analysis on episodic data from two open-source teams. Once finalized, our content analysis will continue on additional episodic data.

People

Partners

- Barbara Scozzi, Politecnico di Bari, Italy
- Megan Conklin, Elon University, North Carolina

Citations

Scozzi, B., Crowston, K., Eseryel, U. Y., & Li, Q. (Accepted). Shared mental models among open source software developers. Hawai'i International Conference on System Science, Big Island, Hawai'i, Jan 2008.

Crowston, K., Wei, K., Li, Q., Eseryel, U. Y., & Howison, J. (2007) Self-organization of teams in free/libre open source software development. Information and Software Technology Journal, Special issue on "Understanding the Social Side of Software Engineering, Qualitative Software Engineering Research", vol. 49, p. 564. Published.

Heckman, R., Crowston, K., Eseryel, Y., Allen, E., & Li, Q. (2007) Emergent Decision-making Practices in Free/Libre Open Source Software (FLOSS) Development Teams. Third International Conference on Open Source Systems, IFIP 2.13 Working Conference, Limerick, Ireland. Published.

Heckman, R., Crowston, K., Misiolek, N. & Eseryel, Y. (2007) Emergent leadership in self-organizing virtual teams. Proceedings of the International Conference on Information Systems (ICIS 2007), Montreal, Quebec, Canada, Dec 2007. Accepted with revisions.

Crowston, K., Conklin, M. & Howison, J. (2007) A Proposed Data and Analysis Archive for Research on Free and Open Source Software and Its Development. Third International Conference on e-Social Science. Accepted as poster.

Crowston, K. (2007) The bug fixing process in proprietary and free/libre open source software: A coordination theory analysis, bibl. M. E. Sharpe, Inc, Armonk, NY. Book Published of Collection: Grover, V. and Markus, M. L., Business Process Transformation.

Plus 6 publications from 2006