

Raw Data



Data Integration

People with their events		
• Person ID: 48 (identified by sf_user_id: 993558, real_name	ne: Jeff Johnson, sf_username: johnson1234, email: jgjohnso@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 Tue Jun 21 10:33:12 EDT 2005: Accents in macros Tue Jun 21 10:40:03 EDT 2005: Crashed on changi Tue Jun 21 16:32:26 EDT 2005: 1682440 on bugs 1 Thu Jul 07 21:40:07 EDT 2005: TypeInfo Editor do Thu Jul 07 22:07:01 EDT 2005: Author sorting for 1 Sat Jul 09 16:03:43 EDT 2005: 1699934 on bugs 12 Sat Jul 23 18:19:46 EDT 2005: Save changes dialog Wed Aug 31 20:39:21 EDT 2005: BibDesk 1.1.5 ha Wed Aug 31 20:53:58 EDT 2005: BibDesk crashes Tue Sep 06 11:31:35 EDT 2005: BibDesk crashes Tue Sep 06 11:31:35 EDT 2005: Preview pane stop Sat Oct 08 20:16:37 EDT 2005: 1791318 on bugs 1 Thu Oct 13 16:46:20 EDT 2005: Sorting not fully st Thu Oct 20 12:25:49 EDT 2005: Extraneous arrow Sat Oct 29 13:46:11 EDT 2005: Extraneous arrow Sat Oct 29 13:46:11 EDT 2005: Groups don't work 	912358 not displayed ng Preview Pane preference 1224896 esn't quit edit mode two-word last names 234554 g with no changes angs on New Publication with Crossref when entering "{" character in raw BibTeX by Booktitle ields appear first in sorted list ps updating after cancelled Close 308021 ticky twords/Group pane not sticky ts above scroll bar when search by content is selected	
Type of Event to graph	Time Series for 'bi	bdesk'
 Ougs_addinasion Cvs_action Spoke_on_mailing_list bugs_comment initiated_mailing_list_thread release downloads 	}	
 pageviews activity_percent activity_rank Which graphing library? • R 'its' • Scruffy	Cvs mail_speak	
Dates		
Truncate Format as YYYY-MM-DD eg 2004-01-10. Leave blank for all available. Start Date:		
End Date:	2003 2004	2005



Team Performance

implementation

Measure

Software

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? L**eadership Building:** How is leadership defined in open-source teams, and how do members rise into Coordination/Collaboration: What are the dynamics of interaction within open-source teams where

structure

Process

Software development process

Planning, requirements, coding,

release, maintenance, project

alization, decision making

itenance, knowledge

ership building, coordinati

IT-related process

management

ocial processes

aboration, team

Technology collaboration

company,

discontinuities,

motivation, skills

Technology use

Types of technology

used, technology features

Project characteristics

Software type, license

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?



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

> 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

• 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.





naelsciali





• 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.

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



All - analysis by each marked sentence Group - analysis by each decision group of sentences *SM* - software modification decision episodes *St* - strategic decision episodes

Precision - % marked that agree with Gold Standard

Recall - % of Gold Standard that are marked **Utility** - % marked that assist researcher in understanding Decision Episodes

DA - Decision Announcements

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

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