

Examining the Structure of Internet-based Open Source Software Collaborations

Charles M. Schweik

Associate Professor

Center for Public Policy and Administration,
Department of Natural Resources Conservation and
Associate Director, National Center for Digital Government
University of Massachusetts, Amherst

Robert English

Graduate Student

Center for Public Policy and Administration
University of Massachusetts, Amherst

What is Free/Libre and Open Source Software (FOSS)?

- Software with licenses that give users the freedom to:
 - (1) run/use the program,
 - (2) read and modify the program, and
 - (3) redistribute copies of either the original or modified program
- Teams of paid and/or volunteer developers
- FOSS projects are a form of Internet-based commons, producing public goods

What is the population of FOSS projects?

Table 1: Numbers of FOSS projects in Sourceforge.net organized by development team size (April, 25, 2005)

Source: FLOSSmole (2005)

Greater than 25 developers	224
Between 11 and 25 developers	1573
Between 5 and 10 developers	5532
Between 1 and 4 developers	86373
Total number of FOSS projects on Sourceforge as of April, 2005	93702

- Sourceforge.net currently advertises 140,000+ projects
- Many projects do not use Sourceforge as a hosting site

Why should public administrators care about FOSS? - Technological reasons

- FOSS software in use in government agencies
- FOSS as a driver toward open standards and interoperability of software and data
- Avoidance of vendor “lock-in”
- Software sharing and reuse
- Sharing of resources toward common goal

Why should public administrators care about FOSS collaborations?

- Recent efforts to initiate FOSS collaborations:
 - Government Open Code Consortium (GOCC.org);
 - Open Source Software Institute (<http://www.oss-institute.org/>);
 - The Foundation for Free and Open Source Software in Government (free4gov.org);
 - EU's Interoperable Delivery of European eGovernment Services (<http://ec.europa.eu/idabc/en/home>)

Why should public administrators care about FOSS collaborations?

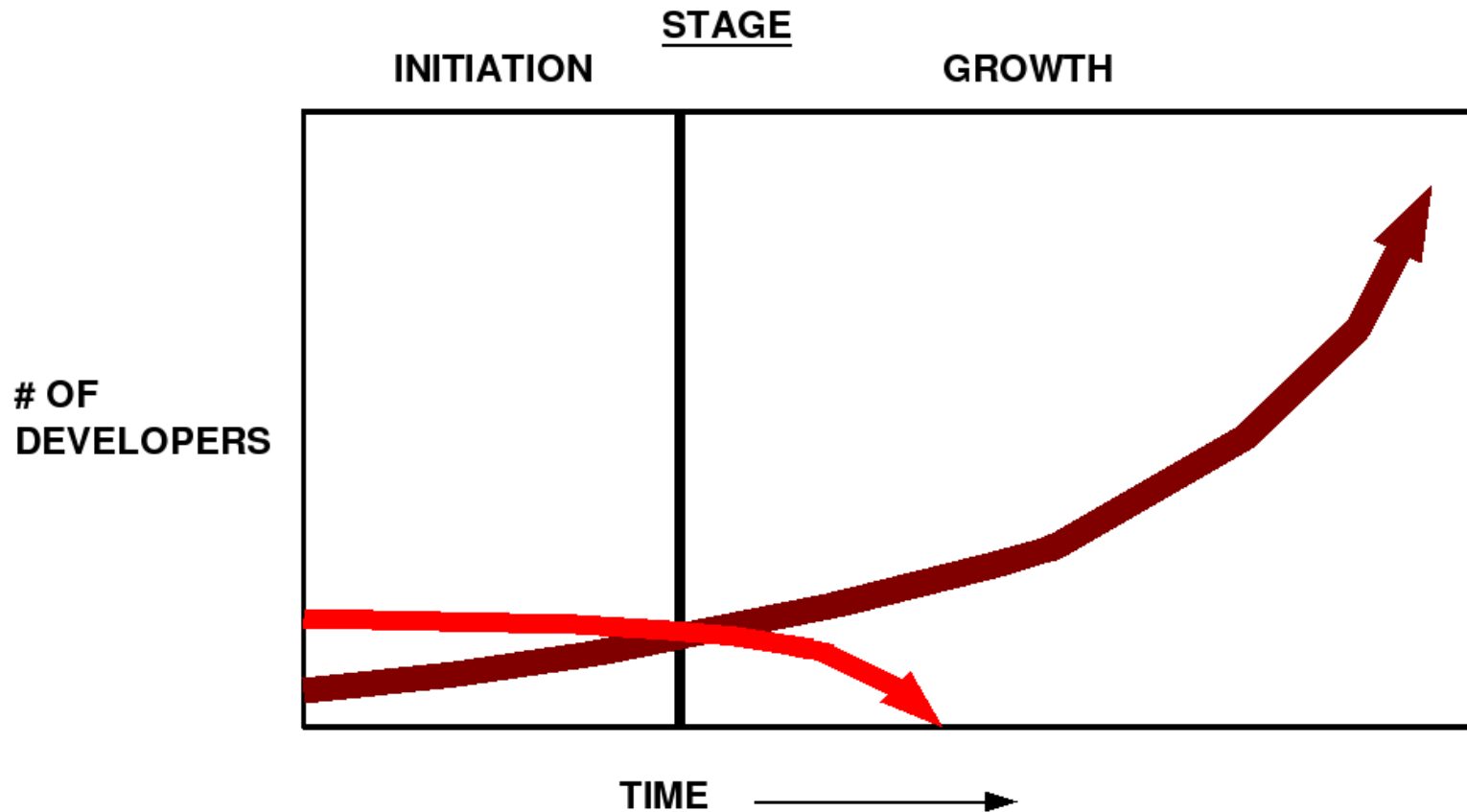
- The potential of FOSS collaboration in domains other than software
 - An existence proof: web growth 1994-2000
 - Creative Commons
 - Open Content collaboration in landuse change modeling

A 5-year research project studying FOSS and Open Content collaborations

Goals:

- 1) To identify “design principles” for FOSS collaborations;
- 2) To study open content collaborations and see if they exhibit similar characteristics

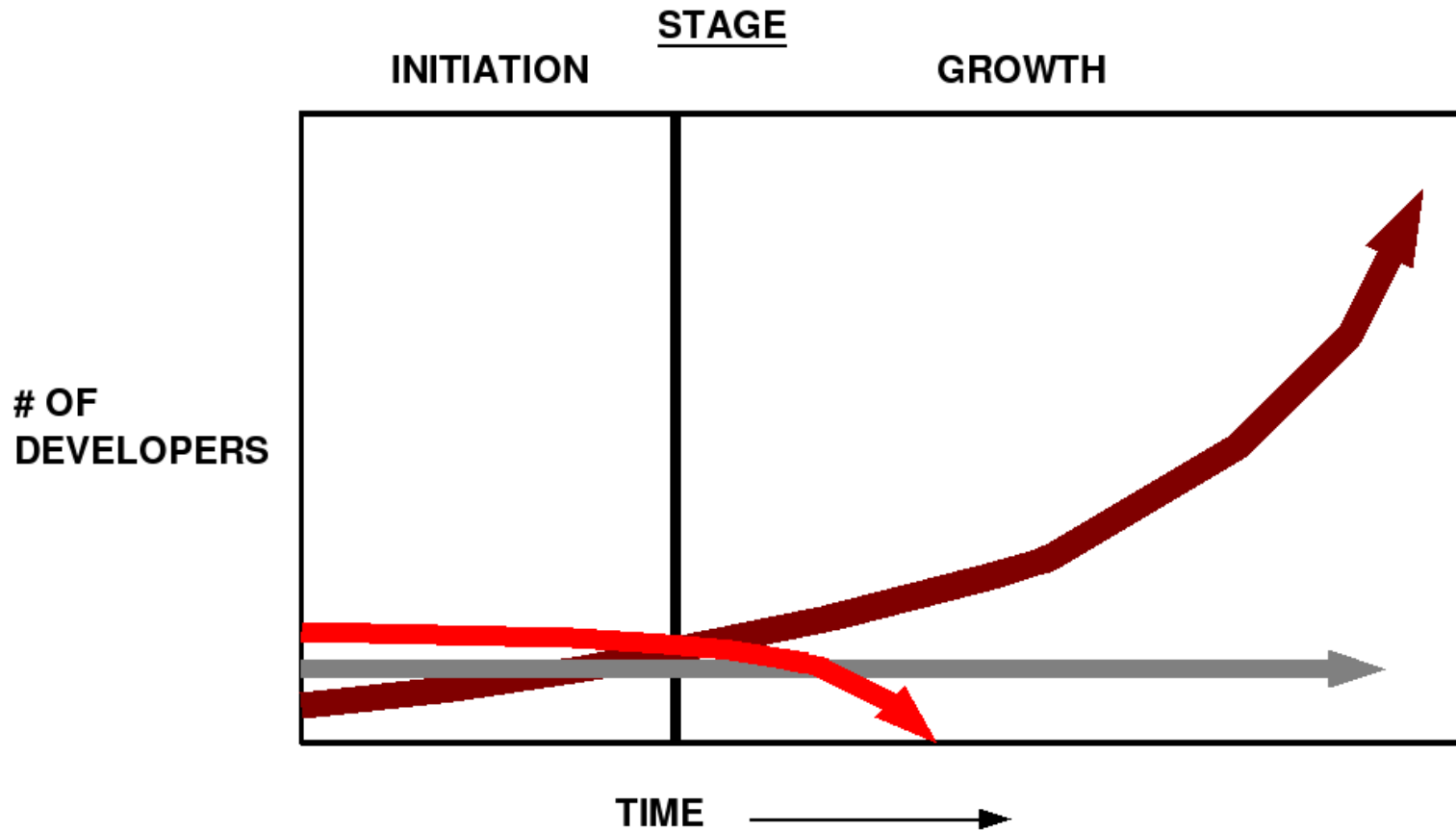
THE POSSIBLE SUCCESS/FAILURE TRAJECTORIES OF FOSS PROJECTS



Typical success: starts small, grows into large team and large user base

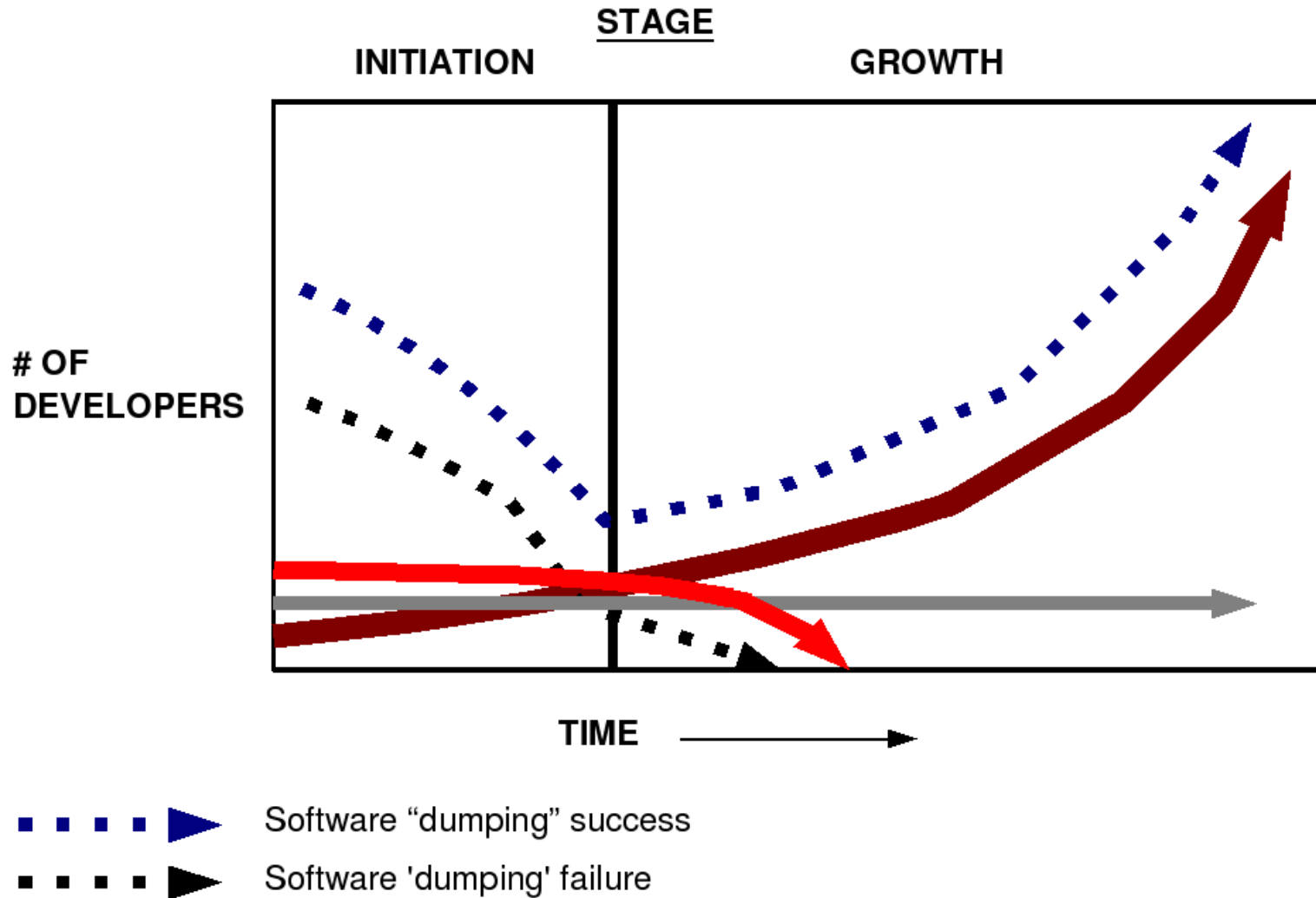
Typical failure: starts small, becomes abandoned

THE POSSIBLE SUCCESS/FAILURE TRAJECTORIES OF FOSS PROJECTS



➔ Another success, starts with small community, stays small (e.g., bioinformatics)

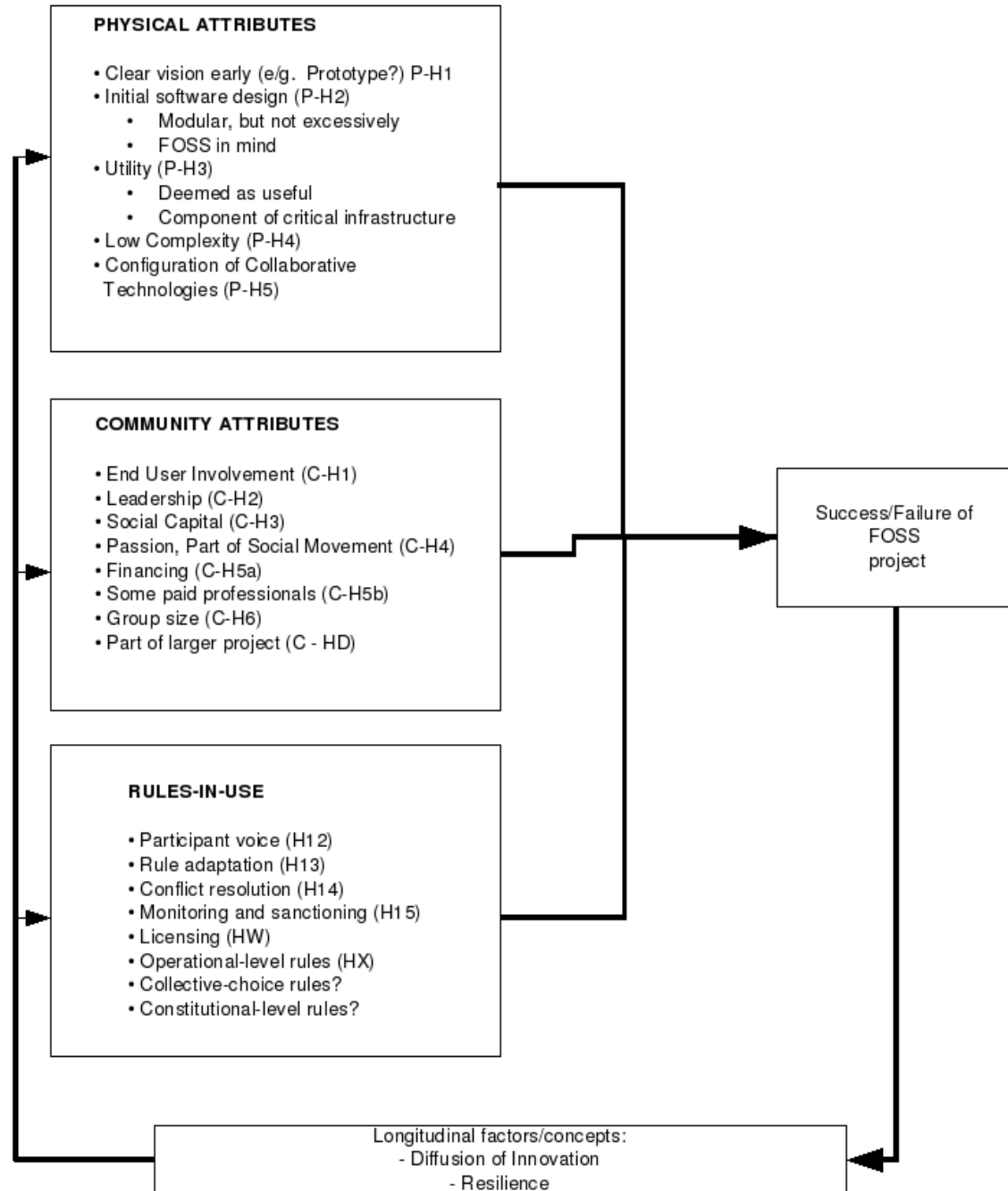
THE POSSIBLE SUCCESS/FAILURE TRAJECTORIES OF FOSS PROJECTS



FOSS Research Strategy

1. Literature review, develop preliminary success/failure model
2. Interviews with FOSS developers
3. Develop success/failure classification system for FOSS, implement and validate
4. Online survey of FOSS developers for quantitative model
5. Qualitative research of cases that look particularly interesting

A Preliminary Model of FOSS Success/Failure



FOSS Research Strategy

1. Literature review, develop preliminary success/failure model
2. Interviews with FOSS developers
3. Develop success/failure classification system for FOSS, implement and validate
4. Online survey of FOSS developers for quantitative model
5. Qualitative research of cases that look particularly interesting

FOSS DEVELOPER INTERVIEWS - METHODS

- 9 FOSS projects
- Case selection strata:
 - ◆ Project stage: “Initiation” and “Growth”
 - ◆ Number of developers
 - ✓ <5, between 5-10, between 11-25, >25
- Structured interviews, some focus on institutions
- Utilized FLOSSMole database from Syracuse U.

SELECTED FINDINGS – PHYSICAL ATTRIBUTES

- Important factors:
 - Modularity
 - Programming language
 - Collaborative platform (versioning system, bug tracking, IRC, listservs)
 - Well documented code, user documentation
- Not important
 - Particular type of collaboration platform

SELECTED FINDINGS – COMMUNITY ATTRIBUTES

- Important factors:
 - Developer abilities
 - Developers are endusers of software
 - Leadership: involved and provide vision for project
 - Trust between team members; face-to-face

- Unclear
 - Financing
 - Homogeneity/heterogeneity of participants (e.g., language?)

SELECTED FINDINGS – ROLE OF INSTITUTIONS

- Small projects: informal norms, Larger projects: more formalized rule systems (as expected)
- **Very few operational level rules across all group size classes;** Most rules that exist embedded in the web-based collaboration platform (e.g., CVS)
- No systems for rule breaking or conflict resolution
- Very flat hierarchies
- One developer: “Remarkable that it works, isn't it?”

IMPLICATIONS?

- Suggests that future FOSS collaborations should:
 - avoid imposing new rules to existing projects
 - be careful of developing overly bureaucratic system in new projects
- Small sample – obviously not generalizable

FOSS Research Strategy

1. Literature review, develop preliminary success/failure model
2. Interviews with FOSS developers
3. Develop success/failure classification system for FOSS, implement and validate
4. Online survey of FOSS developers for quantitative model
5. Qualitative research of cases that look particularly interesting

MOVING TOWARD A LARGER N STUDY: SUCCESS/FAILURE (Tragedy)

<i>Class/ Abbreviation</i>	<i>Definition(D)/Operationalization(O)</i>
Success, Initiation (SI)	D: Developers have produced a first release. O: At least 1 release (Note: all projects in the growth stage are SI)
Tragedy, Initiation (TI)	D: Developers have not produced a first release and the project is abandoned O: 0 releases AND ≥ 1 year since SF project registration
Success, Growth (SG)	D: Project has achieved three meaningful releases of the software and the software is deemed useful for at least a few users. O: 3 releases AND ≥ 6 months between releases AND does not meet the download criteria for tragedy detailed in the TG description below.
Tragedy, Growth (TG)	D: Project appears to be abandoned before producing 3 releases of a useful product. O: 1 or 2 releases and ≥ 1 year since the last release at the time of data collection OR < 11 downloads during a time period greater than 6 months starting from the date of the first release and ending at the data collection date
Indeterminate Initiation (II)	D: Project has yet to reveal a first public release but shows significant developer activity O: 0 releases and < 1 year since project registration
Indeterminate Growth (IG)	D: Project has not yet produced three releases but shows development activity or has produced 3 releases but has been in the growth phase for less than 6 months. O: 1 or 2 releases and < 1 year since the last release OR 3 releases and < 6 months between releases

Preliminary classification of 107,000 FOSS projects

<i>Class</i>	<i># of Projects (% of Total)</i>
Failed, initiation	35,589 (33)
Success, growth	15,782 (15)
Failed, growth	23,134 (21)
Indeterminant Initiation	15,073 (14)
Indeterminant, growth	18,169 (17)
Total	107,747

Note: SI is not listed because these successes are Growth Stage projects. Including SI would double count.

ACKNOWLEDGEMENTS

- Support for this project is provided by a grant from the U.S. National Science Foundation (NSFIIS 0447623)
- The findings, recommendations, and opinions expressed are those of the authors and do not necessarily reflect the views of this funding agency.

Papers related to this research

Schweik, C., T. Evans and J. Grove 2005. Open Source and Open Content: a Framework for Global Collaboration in Social-Ecological Research. *Ecology and Society* 10 (1): 33.

Schweik, C.M. and English, R. "Conceptualizing the Institutional Designs of Free/Libre and Open Source Software Projects." *First Monday* 12(2).

English, R. and Schweik, C.M. Accepted. "Identifying Success and Tragedy of Free/Libre and Open Source (FLOSS) Commons: A Preliminary Classification of Sourceforge.net projects." *Proceedings of the 29th International Conference on Software Engineering + Workshops*

Papers available at: <http://people.umass.edu/cschweik>

Email: cschweik@pubpol.umass.edu