

How to Gather FLOSS Metrics

Megan Conklin¹, Jesus M. Gonzalez-Barahona², and Gregorio Robles³

1 Elon University, Department of Computing Sciences, Elon, NC 27244
mconklin@elon.edu,

WWW home page: <http://facstaff.elon.edu/mconklin>

2 Universidad Rey Juan Carlos, Grupo de Sistemas y Comunicaciones,
c/Tulipan s/n E-28933 Mostoles, Spain
{jgb,grex}@gsyc.escet.urjc.es,

WWW home page: <http://libresoft.urjc.es/>

Abstract. In this half-day tutorial, participants will gain hands-on exposure to key technologies for data collection about open source projects.

1 Program and Objectives

The tutorial will begin with reviews of the main source code repositories, including popular code forges such as Sourceforge, and techniques for collecting data directly from the forges as well as from aggregation projects such as FLOSSmole¹. The tutorial will then discuss tools designed for analyzing the data found on forges, such as CVSanaly², Pyternity, and SLOCCount, among others. Most importantly, participants will have a chance to analyze data with the help of the presenters. Teams of participants will solve open-ended analysis problems collaboratively and in real-time during the workshop. Finally, participants will have opportunities to discuss with the presenters what sort of data collection and analysis tools they would like to see built in the future.

Tutorial program:

- Briefly introduce overall problem of data collection
- Introduce tools: FLOSSmole, CVSanaly, Pyternity, SLOCCount, etc
- Distribute data sets, pose problems for real-time assessment
- Share results
- Discuss future prospects

¹ <http://ossmole.sf.net>

² <http://cvsanaly.tigris.org/>

2 Background of Presenters

Megan Conklin is an assistant professor in the Department of Computing Sciences at Elon University. Her primary research focus is on data mining and large database systems, particularly for software engineering data. She was co-organizer of the 2006 WoPDaSD workshop at the International Conference on Software Engineering (along with Gregorio Robles and Jesus Gonzalez-Barahona). She has published a number of papers on tools for analyzing open source projects, and has spoken about open source data collection at such diverse events as the Mining Software Repositories workshop at ICSE and the O'Reilly Open Source Convention. She has a PhD in computer science from Nova Southeastern University.

Jesus M. Gonzalez-Barahona teaches and researches at Universidad Rey Juan Carlos, Mostoles (Spain). His research interests include libre software engineering, and in particular quantitative measures of libre software development and distributed tools for collaboration in libre software projects. In this area, he has published several papers, and is participating in some international research projects (more info at <http://libresoft.urjc.es>). He is also one of the promoters of the idea of a European masters program on libre software, and has specific interest in education relating to that area.

Gregorio Robles is Associate Professor at the Universidad Rey Juan Carlos in Madrid, Spain. He earned a degree in electrical engineering from the Universidad Politécnica de Madrid (studying his last year and submitting his master thesis at the Technical University of Berlin, DE) and obtained his PhD in 2006. His research work is centered in the study of libre software development from an engineering point of view, especially with regard to quantitative and empirical issues. Related, non-technical matters have also been of interest: volunteer-driven software development and social network analyses of the libre software phenomenon. He has developed or collaborated in the design of programmes to automate the analysis of libre software and the tools used to produce them. He was also involved in the FLOSS study on libre software financed by the European Commission IST programme, and was involved in other European-funded projects such as CALIBRE or FLOSSWorld. He has also had the opportunity to attend the following universities as a research visitor: Wirtschaftsuniversität Wien (AT, 2 months), MERIT/University of Maastricht (NL, 4 months), the University of Lincoln (UK, 3 months) and the Technical University Munich (DE, 5 months).