

Open source metrics: where software engineering and economics meet

Rishab Aiyer Ghosh, 25/9/2003, Brussels

for the meeting on Open Source Software with Commissioner Liikanen

The demonetisation of productivity

The growth of productive time spent in online activity is leading to an increasing “demonetisation” of the economy. Open source software is the most obvious example. While business interest and commercial exploitation in recent years has shown that open source software development is clearly productive economic activity, more than half of it remains without the direct involvement of businesses, and most open source software is produced without explicit monetary rewards.

This demonetisation is bad for economists and decision makers. As the tools decision makers and economists use are generally blind to non-monetary economic activity, an increasingly large and fast growing section of the economy is now invisible.

Hidden software, theoretical engineering

In software engineering, applied research (usually academic) has long been hampered by the lack of good access to actual development activity, which has been proprietary and thus usually closed to research. This makes it hard to develop and evaluate theoretical models based on empirical evidence, and to test that engineering techniques that should work in theory actually work in practise – despite the fact that many theoretical models are widely used in industry almost as an act of faith. The COCOMO models for estimating the development cost of software were most famously based on studies of as few as 63 projects, studies which by their nature could not be reproduced by another research group.

Software engineering research would develop much further if theory and practise could meet in the same environment.

Non-monetary measurement and empirical software engineering

Open source software development provides an excellent open environment for the empirical, evidence-based exploration and development of software engineering models and techniques. Not only does the open access of this model provide researchers the possibility of testing their theories in practise, it also provides the empirical basis to form theoretical development models in the first place.

Tools and methods used to study open source software development from a software engineering viewpoint also provide an excellent opportunity for economists to measure the demonetised, previously invisible productivity of open source software projects, and also to analyse the organisation and production methods of software at a level of detail probably unmatched by any other field of economic activity. This is because almost every single act of production, direct or indirect, is documented, recorded and published somewhere in the open source development process, suitable for study. And although this sort of measurement may not, initially, be in monetary terms, it does represent human time and effort spent on productive activity, and can be “remonetised” at least for the purposes of measurement.

What open software can tell us

Although open source software is not uniformly structured, current research already shows how it is possible to extract at a high level of detail: who created which lines of code at what time while collaborating with which other developers linked to which other software projects, responding to which bug reports after how many lines of discussion in which discussion group... all feeding back into software engineering models and economic productivity studies.