Lightning talk: Software Impact Measurement at the Netherlands eScience Center

Willem van Hage, Jason Maassen, Rob van Nieuwpoort
Netherlands eScience Center
{w.vanhage | j.maassen | r.vannieuwpoort}@esciencecenter.nl

1. Introduction

In this lightning talk we present a new initiative at the Netherlands eScience Center (NLeSC) about software impact measurement.

The NLeSC is a Dutch funding agency and scientific software development expertise center that supplies specialized manpower along with research funding for innovative eScience projects to all of the sciences. The main goal of the NLeSC is to enhance science with information technological innovation. This is accomplished by alleviating a number of problems that come with the adaptation of the scientific process as we know it to include new IT, such as missing scientific software tooling; the lack of skilled scientific software developers; missing specific technological expertise such as big data analytics, optimized data handling, and efficient computing; missing publication fora for scientific software; and the lack academic professional assessment by scientific software impact. Of all these urgent problems, the last might be the most important. If academics are assessed by their software, then the academic world could eventually reorganize itself so that all the other issues are solved.

2. eScience Technology Platform

To stimulate scientific software development, the NLeSC has set up the eScience Technology Platform (eStep), which consists of three parts: a software catalogue of eScience software, interfaces, libraries, tools workflows, applications, models and data sets; a knowledge base consisting of guides, reports and recommendations for scientific software development; and documents describing eScience research such as scientific publications and demos.

Currently, eStep is used by the NLeSC itself for internal dissemination of knowledge and for self-evaluation, by NLeSC partners like the funded research institutes to learn about the portfolio of the NLeSC and new trends in eScience, and by other users of the eScience software produced by the NLeSC to find out how to contribute or whom to contact for more information.

To help all of these users to access the information they need NLeSC launched a new Web site that presents all the interrelated topics in eStep, such as software packages, people that were involved in the development and use of these software packages, projects in which they were used, and organizations that participated in these projects. Users can browse, search, and filter the website. The analytics site that allows visitors to filter the contents of eStep is illustrated by the screenshot shown in Figure 1. An example software project page is shown in Figure 2.

The website is automatically generated from metadata records that can be submitted added to the website by the developers themselves, putting the responsibility at the scientific developers themselves. Currently, the only contributors to the website are NLeSC employees, but in the near future the website will become a shared effort by an increasing number of eScience research institutes.

3. Software Impact Test Bed

This year, the NLeSC will start with an experimental addition to the eStep website to aid scientific developers with the assessment of their software impact. The goal is that each software package described on the website, and added by the developers themselves, will get an automatically generated dashboard presenting various aspects of the software impact.

There are many different approaches to measure software impact based on different measurable quantities: direct and indirect software (report) citations, software dependency structure, software contribution, forking of the code base, usage in projects, endorsements or reviews, and software quality metrics.

The goal of the software impact measurement website is to serve as a test bed to assess the usefulness and reliability of all of these different approaches in practice. The NLeSC hopes to learn which aspects of software projects are best used to measure software impact, and to use this empirical knowledge to influence scientific organizations to start assessing researchers by these measures. Hopefully this will lead to an increased awareness at both funding agencies and research institutes that software drives innovative research, that better software means better research, and that appropriate key performance indicators can be used to improve the scientific performance.

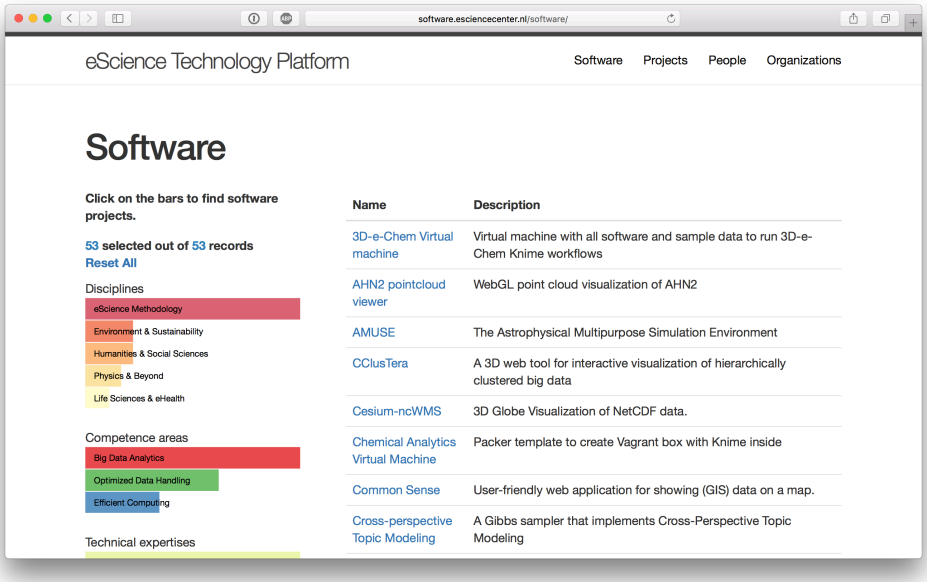

Figure 1: The eStep software project analytics webpage (screenshot June 2016).

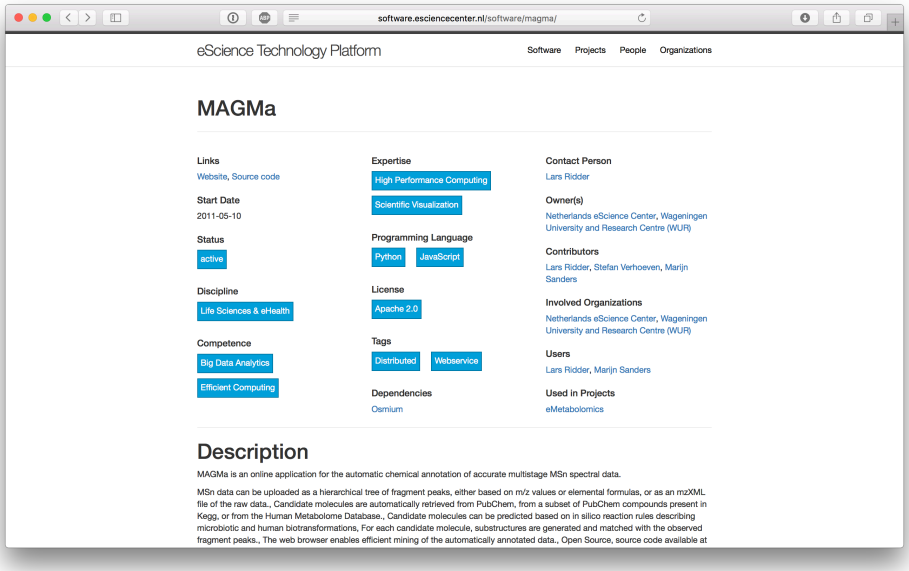

Figure 2: An example software project description. The screenshot shows the metadata of the MAGMa chemistry tool and a small part of the description of the tool. (screenshot June 2016