



GenOA week 2023
Università di Genova
23-26/10/2023

GitLab@Polito

Una piattaforma
istituzionale per
sviluppare e valorizzare
il software di ricerca

Mauro Paschetta, Riccardo Gnudi, Stefano
Di Carlo, Daniele Marchisio, Fabrizio Riente,
Giovanni Squillero, Enrico Venuto, Federica
Cappelluti

In collaborazione con l'Ufficio Trasferimento Tecnologico
e con la Commissione Open Access di Ateneo



Politecnico
di Torino

Hannay et al., 2009

Conferences > 2009 ICSE Workshop on Softwar...

How do scientists develop and use scientific software?

Publisher: IEEE [Cite This](#) [PDF](#)

Jo Erskine Hannay; Carolyn MacLeod; Janice Singer; Hans Petter Langtangen; Dietmar Pfahl; Greg Wilson [All Authors](#)

128 Cites in Papers 1796 Full Text Views

[R](#) [←](#) [©](#) [📁](#) [🔔](#)

[LEARN MORE >](#)

Abstract

Document Sections

1. Motivation
2. Research Method
3. Results
4. Discussion
5. Related Work

[Show Full Outline ▾](#)

Abstract:

New knowledge in science and engineering relies increasingly on results produced by scientific software. Therefore, knowing how scientists develop and use software in their research is critical to assessing the necessity for improving current development practices and to making decisions about the future allocation of resources. To that end, this paper presents the results of a survey conducted online in October-December 2008 which received almost 2000 responses. Our main conclusions are that (1) the knowledge required to develop and use scientific software is primarily acquired from peers and through self-study, rather than from formal education and training; (2) the number of scientists using supercomputers is small compared to the number using desktop or intermediate computers; (3) most scientists rely primarily on software with a large user base; (4) while many scientists believe that software testing is important, a smaller number believe they have sufficient understanding about testing concepts; and (5) that there is a tendency for scientists to rank standard software engineering concepts higher if they work in large software development projects and teams, but that there is no uniform trend of association between rank of importance of software engineering concepts and project/team size.

More Like This

OASIS: an open architecture scientific information system
Proceedings RIDE '96, Sixth International Workshop on Research Issues in Data Engineering
Published: 1996

Extracting and providing knowledge within an object-centered scientific information system for atmospheric research
Proceedings of 8th International Conference on Scientific and Statistical

J. E. Hannay et al., "How do scientists develop and use scientific software?," 2009 ICSE Workshop on Software Engineering for Computational Science and Engineering, Vancouver, BC, Canada, 2009, pp. 1-8, doi: 10.1109/SECSE.2009.5069155.

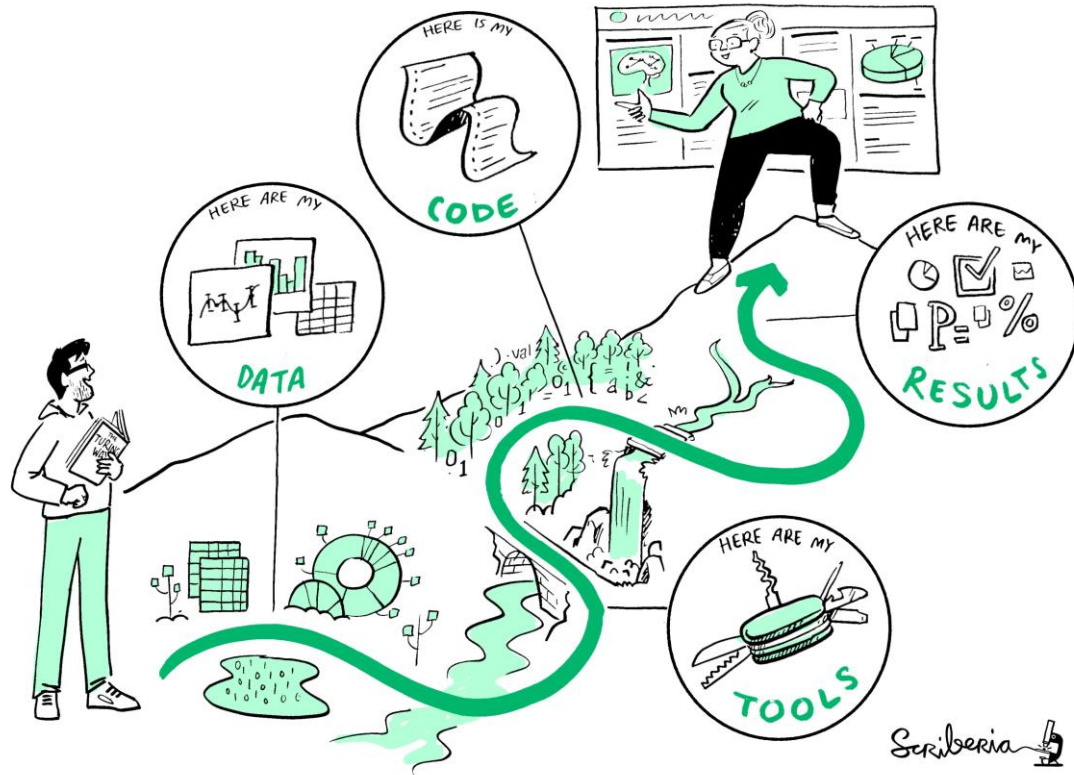
Campione di 1972 ricercatori

84.3%: sviluppare software è molto importante per la propria ricerca

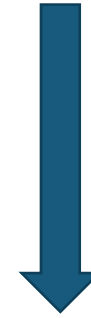
91.2%: utilizzare software è molto importante per la propria ricerca

Pinto, Wiese & Dias replicano lo studio nel 2022, ottenendo percentuali ancora più alte.

I SOFTWARE, insieme a PUBBLICAZIONI e DATI, sono diventati uno dei prodotti più importanti della ricerca.



<https://the-turing-way.netlify.app/reproducible-research/reproducible-research.html>



Oggetto delle pratiche di Scienza Aperta

Open Science e software in Horizon Europe

Evaluation of proposals and Open Science

“Excellence” criterion (methodology)

- Evaluation of the quality of open science practices
- Up to 1 page to describe Open Science practices + up to 1 page to describe research data/output management

“Quality and efficiency of implementation” criterion

(capacity of participants and consortium as a whole + list of achievements)

- Explain expertise on Open Science
- List publications, **software** data, etc, relevant to the project with qualitative assessment and, where available, persistent identifiers

Publications are expected to be open access; datasets are expected to be FAIR and ‘as open as possible, as closed as necessary’. **Significance of publications to be evaluated on the basis of proposers’ qualitative assessment** and not per Journal Impact Factor



See detailed slides by an EC officer dated June 30, 2021: https://pure.mpg.de/rest/items/item_3328552_1/component/file_3328558/content

Informatics Europe

Associazione europea dei dipartimenti universitari e dei laboratori di ricerca nel campo dell'informatica

Open access recommendations for the scientific community

**Open Access:
Status and Recommendations**

An Informatics Europe Report

Prepared by:


- Manuel Carro, Universidad Politécnica de Madrid and IMDEA Software Institute.
- Stefano Bistarelli, University of Perugia.
- Roberto Di Cosmo, Software Heritage, INRIA, and Université Paris Cité.
- Alfonso Pierantonio, Università degli Studi dell'Aquila.

Open Access: Status and Recommendations
May 2023

Published by:

Informatics Europe
Binzmühlestrasse 14/54
8050 Zurich, Switzerland
www.informatics-europe.org
administration@informatics-europe.org

© Informatics Europe, 2023, BY-SA.



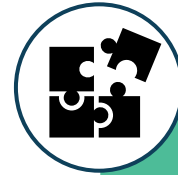
6. Participate in **international activities** on the emerging issues of Open Data and Open Source in research; disseminate best practices for archiving, identification, description, and citation for software produced by researchers using open and mutualized infrastructures such as Software Heritage [15]; contribute to the ongoing international reflection on the evaluation of software production in career progression and funding allocation, avoiding as much as possible to reproduce the errors incurred in scientific publishing [16], such as the misuse of numerical indicators (e.g., the number of commits made on a project).

<https://t.co/4DsjPIEtQ1>

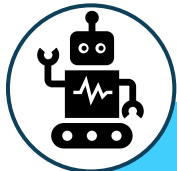
Caratteristiche SOFTWARE



Natura evolutiva



Architettura composita



Codificazione in linguaggio macchina



Il codice sorgente, comprensibile dall'uomo, è un sapere prezioso e va conservato e reso pubblico

SOFTWARE



Strumento



Oggetto di studio

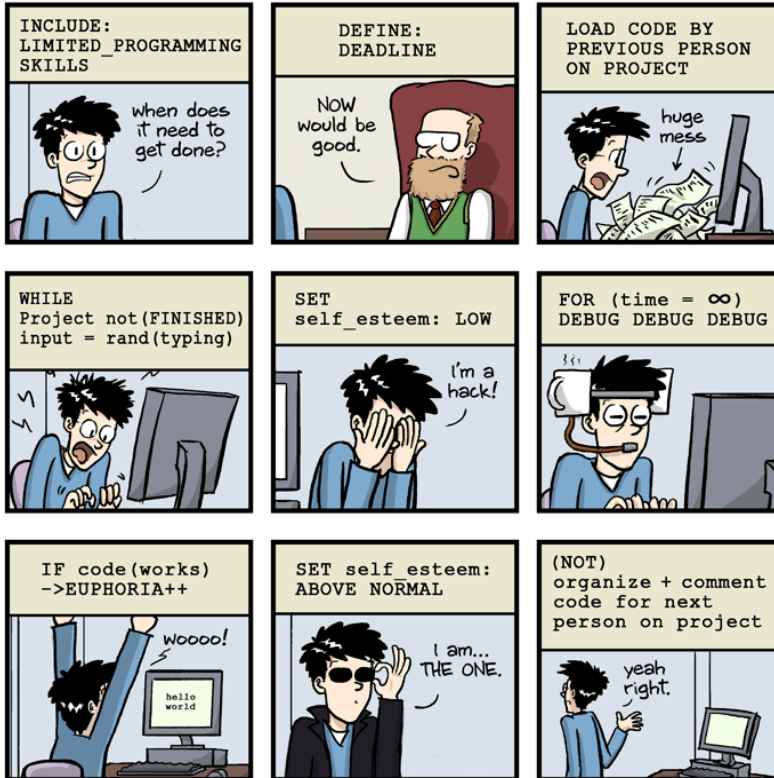


Risultato



Quali sono le necessità di un RICERCATORE?

PROGRAMMING FOR NON-PROGRAMMERS



JORGE CHAM © 2014

WWW.PHDCOMICS.COM

"Piled Higher and Deeper" by Jorge Cham
www.phdcomics.com
<http://phdcomics.com/comics.php?f=1690>

- **Infrastruttura** e **strumenti** per impostare un flusso di lavoro riproducibile
- **Formazione** e **supporto** per la scrittura e il mantenimento di codice riproducibile
- Linee guida sui **flussi di lavoro** e informazioni sulle **licenze**
- Guida allo **sfruttamento commerciale** del software

Progetto Pilota – Piattaforma “GitLab@PoliTO”

OBIETTIVI

Supporto ai ricercatori
nello **sviluppo** e
nell'**utilizzo** dei
software

Corretta gestione del
software per
valorizzare e
conservare i risultati
scientifici

Favorire il rispetto
dei principi FAIR
tramite un workflow
il più possibile **FAIR-
by-design**

- **Trusted repository** istituzionale e vetrina d'ateneo (**F**indable)
- **Accessibilità** ai fini di sviluppo e di utilizzo del software (**A**ccessible), in base alla natura del software (commerciale o non).
- Flussi standardizzati, controllo di versione, documentazione associata (**I**nteroperable)
- Gestione delle **licenze** (**R**eusable) tramite un flusso di gestione standardizzato ed efficiente

STRUMENTO – Perché GitLab

Software per il **controllo di gestione** basato su **GIT** (software per la gestione di versioni più diffuso al mondo)

Flessibilità di accesso

Personalizzabile per i diversi progetti (es. livello di apertura e partecipazione esterni)

Appropriato per gestione di dati **confidenziali** e **personali**

Ambiente di sviluppo **collaborativo**

Open Source

Permette la gestione di su **web** o su **piattaforme** in cloud (private o pubbliche)

Offre diverse tipologie di licenza adatte per ricerca a fini **istituzionali** o **commerciali**

E' possibile avere un'istanza istituzionale, gestita **localmente**, con **login centralizzato**



GitLab

<https://about.gitlab.com/>

STRUTTURA TECNICA

Ambiente di sviluppo - Repo1

Istanza GitLab su server di ateneo.

- Code management secondo linee guida OS.
- Sgravio risorse hardware (server) per gruppi di ricerca.

release

pubblicazione

Ambiente di gestione IPR e pubblicazione prodotto - Repo2

• **Registrazione codice**

Creazione di file di testo o pdf contenenti gli *hash* delle release con scadenza temporale fissa. Marcatatura temporale dei file da parte dell'Ufficio Trasferimento Tecnologico tramite firma elettronica.

• **Applicazione e gestione licenze FOSS**

In base delle esigenze (*as open as possible, as closed as necessary*) e con il supporto dell'Ufficio Trasferimento Tecnologico.

Archivio pubblico unico

- Titolo, descrizione e autori.
- Proprietà: autori o ateneo.
- Licenza FOSS applicata o possibili licenze commerciali.
- Codice sorgente per FOSS → integrazione **Software Heritage**
- Download previo pagamento per azioni commerciali

A che punto siamo?

Ambiente di sviluppo - Repo1 - Sviluppo e controllo accessi

2019-2020: primi test di utilizzo

Su un istanza di tipo «Educational» nell'ambito di alcuni corsi di laurea magistrale

2021: disseminazione e raccolta feedback di interesse

2022: creata 2a istanza con licenza *Free*

Adatta anche a progetti commerciali

Ambiente di gestione IPR e pubblicazione prodotto - Repo2

Sviluppo e test del repository n.2 per la registrazione dei software, applicazione licenze e pubblicazione



What's
Next?

VANTAGGI di una piattaforma istituzionale



Interoperabilità e uniformità

interdipartimentale grazie ad uno strumento unico per lo sviluppo



Valorizzazione e visibilità del software prodotto in ateneo → valutazione, impatto ricerca, ecc.



Facilitazione della **riproducibilità** e **integrità** del software già in fase di sviluppo e a qualsiasi livello (studenti, dottorandi, ricercatori)



Strumento professionale di valore in collaborazioni esterne



Collaborazione trasparente con interni e/o esterni



Conservazione a lungo termine del patrimonio software prodotto in ateneo = **Repository**



Maggiore facilità nel gestire esigenze di **tutela** e **condivisione** del codice prodotto → Open Science e IPR

Ringraziamenti

Per il supporto «culturale» si ringrazia :



Hanno contribuito:

Dr. **Riccardo Gnudi**

Prof.ssa **Federica Cappelluti**

Prof. **Stefano Di Carlo**

Prof. **Daniele Marchisio**

Prof. **Fabrizio Riente**

Prof. **Giovanni Squillero**

Dr. **Enrico Venuto**

Dr. **Shalini Kurapati**

Dr. **Shiva Loccisano**