

**HELMHOLTZ**

Open Science

Helmholtz Open Science Briefing

# Helmholtz Open Science Forum Forschungssoftware

Report

## Impressum

Die Onlineversion dieser Publikation finden Sie unter:

<https://doi.org/10.48440/os.helmholtz.045>

## Verfasser:innen

Heinz Pampel, Nina Leonie Weisweiler, Roland Bertelmann, Wolfgang zu Castell, Sünje Dallmeier-Tiessen, Ronny Gey, Tobias Huste, Nicole Jung, Amir Kalali, Uwe Konrad, Christian Meeßen, Sebastian Müller, Dennis Oliveira, Tobias Schlauch

## Herausgeber

Helmholtz Open Science Office

## Redaktion

Heinz Pampel, Nina Leonie Weisweiler, Roland Bertelmann, Christoph Bruch, Lea Maria Ferguson, Antonia C. Schrader, Paul Schultze-Motel

## Kontakt

Helmholtz Open Science Office  
c/o Helmholtz-Zentrum Potsdam  
Deutsches GeoForschungsZentrum GFZ  
Telegrafenberg, 14473 Potsdam  
E-Mail: [open-science@helmholtz.de](mailto:open-science@helmholtz.de)

## Stand

03.05.2022. Version. 1.0

## Lizenz

Alle Texte dieser Veröffentlichung, ausgenommen Zitate, sind unter einem Creative Commons Attribution 4.0 International (CC BY 4.0) Lizenzvertrag lizenziert. Siehe: <https://creativecommons.org/licenses/by/4.0>.



# HELMHOLTZ

## Open Science

### Inhalt

Einführung .....	2
Hintergrund: Open Source in der Forschung .....	3
Programm .....	7
Vortragsfolien .....	8

#### Abstract

Das Helmholtz Forum Forschungssoftware, welches gemeinsam von der Task Group Forschungssoftware des AK Open Science und dem HIFIS Software Cluster getragen wird, veranstaltete am 7. April 2022 ein Helmholtz Open Science Forum zum Thema Forschungssoftware.

Die Veranstaltung wurde vom Helmholtz Open Science Office organisiert. Das virtuelle Forum widmete sich drei Aspekten beim offenen und nachhaltigen Umgang mit Forschungssoftware in der Helmholtz-Gemeinschaft: Policy, Praxis sowie Infrastrukturen und Tools.

Die Veranstaltung war die zweite einer Reihe von Helmholtz Open Science Foren zum Thema. Die erste Veranstaltung fand im Mai 2021 unter dem Titel „Policies für Forschungssoftware“ statt. Vorliegender Bericht dokumentiert die Veranstaltung.

# HELMHOLTZ

## Open Science

### Einführung

Mit der voranschreitenden Digitalisierung von Forschung und Lehre steigt die Zahl an Software-Lösungen, die an wissenschaftlichen Einrichtungen entstehen und zur wissenschaftlichen Erkenntnisgewinnung genutzt werden. Die unter dem Stichwort Open Science geforderte Zugänglichkeit und Nachnutzung von wissenschaftlichen Ergebnissen kann in vielen Fachgebieten nur sichergestellt werden, wenn neben den Forschungsdaten auch der Programmcode offen zugänglich gemacht wird.

Für die Nachvollziehbarkeit und Reproduzierbarkeit von wissenschaftlichen Ergebnissen ist die Referenzierung und Zugänglichmachung der genutzten bzw. entwickelten Software unerlässlich. Auch für die Nachvollziehbarkeit von Datenanalysen sowie für die Nachnutzung von Forschungsdaten ist die gleichzeitige Bereitstellung korrespondierender Software in vielen Fällen von großer Bedeutung.

An den Helmholtz-Zentren gibt es bereits seit einigen Jahren vielfältige Initiativen im Themenfeld Forschungssoftware. Die Task Group Forschungssoftware des Arbeitskreises Open Science in der Helmholtz-Gemeinschaft widmet sich seit dem Jahr 2016 der Zugänglichkeit und Nachnutzung von Forschungssoftware im Kontext von Open Science.<sup>1</sup>

Mit dem Start des Software Services Cluster der Plattform Helmholtz Federated IT Services (HIFIS)<sup>2</sup> im Rahmen des Helmholtz-Inkubators Information and Data Science im Jahr 2019 wird das Ziel verfolgt eine leistungsfähige IT-Serviceplattform zu schaffen, die Wissen aus allen Helmholtz-Zentren verbindet. Sowohl der Arbeitskreis als auch die HIFIS-Plattform widmen sich dem nachhaltigen Umgang mit Forschungssoftware in der Helmholtz-Gemeinschaft, was durch eine hohe personelle Überschneidung der Stakeholder dokumentiert wird.

Um das Thema in Helmholtz weiter koordiniert voranzutreiben, bilden die Task Group Forschungssoftware und die Software Services Plattform von HIFIS seit 2020 das vom Helmholtz Open Science Office<sup>3</sup> betreute Forum Forschungssoftware in der Helmholtz-Gemeinschaft.

Das Forum Forschungssoftware veranstaltete am 7. April 2022 ein zweites Helmholtz Open Science Forum zum Thema Forschungssoftware. Die Veranstaltung wurde vom Helmholtz Open Science Office organisiert. Die virtuelle Veranstaltung widmete sich drei Aspekten des offenen und nachhaltigen Umgangs mit Forschungssoftware in der Helmholtz-Gemeinschaft:

---

<sup>1</sup> Siehe: <https://os.helmholtz.de/open-science-in-der-helmholtz-gemeinschaft/open-research-software/>

<sup>2</sup> Siehe: <https://www.hifis.net>

<sup>3</sup> Siehe: <https://os.helmholtz.de>



# HELMHOLTZ

## Open Science

- **Policy:** Beleuchtet wurde der Stand der Policy-Entwicklung anhand von Praxisbeispielen aus zwei Helmholtz-Zentren.
- **Praxis:** Vorgestellt wurden drei exemplarische Software-Projekte, die Aspekte der Offenheit und Nachhaltigkeit in besonderer Weise berücksichtigen und bereits Breitenwirkung entfaltet haben.
- **Infrastrukturen & Tools:** Vorgestellt wurden Trainingsangebote von HIFIS, Repositorien und weitere Angebote rund um die Entwicklung und Publikation von Forschungssoftware. In partizipativen Sessions wurden anschließend aktuelle Herausforderungen und Chancen dieser drei Aspekte adressiert.

Die Veranstaltung war die zweite einer Reihe von Helmholtz Open Science Foren zum Thema. Die erste Veranstaltung fand im Mai 2021 unter dem Titel „Policies für Forschungssoftware“ statt.

Mit etwas über 100 Teilnehmenden war die Veranstaltung sehr gut besucht und bot interessierten Personen aus verschiedenen Arbeitsbereichen rund um das Thema Forschungssoftware eine Plattform für Information und Vernetzung zum Thema.

## Hintergrund: Open Source in der Forschung<sup>4</sup>

Mit der voranschreitenden Digitalisierung von Forschung und Lehre steigt die Zahl an Software-Lösungen, die an wissenschaftlichen Einrichtungen entstehen. Ob zur Analyse großer Datenmengen, zur Steuerung von wissenschaftlichen Geräten oder zur Visualisierung von wissenschaftlichen Ergebnissen - Forschungssoftware ist heute ein zentrales Element im wissenschaftlichen Erkenntnisprozess.

Damit Open Science<sup>5</sup> funktioniert, wissenschaftliche Ergebnisse also offen verfügbar und reproduzierbar sind, müssen in vielen Fachgebieten auch die Programmcodes offen zugänglich sein. Dieser Herausforderung widmet sich Helmholtz unter dem Begriff Open Research Software. Bereits 2007 wurde ein Positionspapier erarbeitet.<sup>6</sup> Seit 2020 agiert das Helmholtz Forum Forschungssoftware als Netzwerk zu dem Thema.

---

<sup>4</sup> Vorliegender Text wurde von Heinz Pampel im Mai 2022 in einer ersten Version auf [helmholtz.de](https://os.helmholtz.de) veröffentlicht. Der überarbeitete und aktualisierte Text ist Teil einer Artikelreihe zum Thema Open Science in Helmholtz und gibt einen Überblick über die Aktivitäten im Bereich Open Research Software.

<sup>5</sup> Siehe: <https://os.helmholtz.de>

<sup>6</sup> Siehe: <https://os.helmholtz.de/?id=2766>

# HELMHOLTZ

## Open Science

In diesem Verbund arbeiten Vertreter:innen aus Wissenschaft und Informationsinfrastruktur an der Förderung von Open Research Software. Partner ist das Software Services Cluster der Plattform Helmholtz Federated IT Services (HIFIS). HIFIS unterstützt Forscher:innen im Umgang mit Software unter dem Motto „Helmholtz IT for Science“.<sup>7</sup> Zum Angebot zählen zum Beispiel Kurse wie „Let's make your script ready for publication“ und Einzelberatungen, aber auch digitale Infrastrukturen wie der HIFIS Transfer Service (HTS). Dieser sorgt für den reibungslosen Austausch großer Datensätze zwischen den Helmholtz-Zentren.

Ein nachhaltiger Umgang mit wissenschaftlicher Software wird auch von Förderorganisationen eingefordert. SO behandelt die Deutsche Forschungsgemeinschaft (DFG) das Themenfeld in ihren „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“ umfassend.<sup>8</sup> Diesen Kodex verankern derzeit alle Hochschulen und außeruniversitären Forschungseinrichtung in Deutschland. Eine der Anforderungen an die digitale Wissenschaftspraxis dort lautet: „Bei der Entwicklung von Forschungssoftware wird der Quellcode dokumentiert.“ Auch die Europäische Kommission verlangt in ihrem neuen Rahmenprogramm Horizon Europe nicht nur den offenen Zugang zu wissenschaftlichen Publikationen und Forschungsdaten, sondern auch die Bereitstellung der verwendeten Tools, die zur Nachnutzung von veröffentlichten Forschungsdaten nötig sind.<sup>9</sup>

Doch mit der Bereitstellung der Software allein ist es meist nicht getan. Notwendig sind auch neue Publikationsstrategien. Sie müssen einerseits die Qualitätssicherung des Codes garantieren und andererseits auch die Zitation der Software erlauben - erst dann ist sichergestellt, dass die Programmierer:innen adäquate Anerkennung im wissenschaftlichen Reputationssystem finden.

Vor diesem Hintergrund beginnen Forschungseinrichtungen digitale Infrastrukturen aufzubauen, sogenannte Repositorien, auf denen wissenschaftliche Software offen publiziert werden kann. So zum Beispiel am CERN in Genf: Dort wurde das Open-Access-Repositorium Zenodo implementiert.<sup>10</sup> In Zusammenarbeit mit der Entwicklungsplattform GitHub stellt Zenodo wissenschaftliche Programmcodes als referenzierbares Produkt der wissenschaftlichen Arbeit dauerhaft zur Verfügung. Dazu wird die Software mit einem Digital Object Identifier (DOI) adressiert und kann so in der Publikationsliste geführt werden - dadurch

---

<sup>7</sup> Siehe: <https://hifis.net>

<sup>8</sup> Deutsche Forschungsgemeinschaft. (2019). Guidelines for Safeguarding Good Research Practice. Code of Conduct. <https://doi.org/10.5281/zenodo.3923602>

<sup>9</sup> European Commission. (2021). EU Grants. AGA - Annotated Model Grant Agreement.

[https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/guidance/aga\\_en.pdf](https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/guidance/aga_en.pdf)

<sup>10</sup> Siehe: <https://docs.github.com/en/repositories/archiving-a-github-repository/referencing-and-citing-content>

# HELMHOLTZ

## Open Science

wird ein breiter Blick auf Forschungsleistung ermöglicht. Ähnliche Strategien verfolgen auch die Helmholtz-Zentren. Ihr Ziel ist es, Programmcodes als Produkt wissenschaftlicher Arbeit anzuerkennen, damit die wichtige Arbeit der Entwickler:innen von Forschungssoftware sichtbar wird.

Um die Qualität der Software sicherzustellen, tauschen sich die Forscher:innen über dedizierte Software-Journale aus, etwa über die Open-Access-Zeitschrift „Journal of Open Source Software“.<sup>11</sup> Die Gutachter:innen dieser Zeitschrift prüfen einen Code und dessen Dokumentation. Sind alle Kriterien erfüllt, wird die Dokumentation als sogenanntes Software-Paper veröffentlicht. Dieser Artikel referenziert dann die publizierte Software auf einem Repositorium. Um ein Forschungsergebnis entsteht so ein Wissensraum, der im Sinne der Open-Science-Idee nicht nur Daten und Interpretationen vernetzt, sondern auch verwendete Methoden und Werkzeuge.

Aktuell beschäftigt sich eine wachsende Zahl von Helmholtz-Zentren mit der Verankerung von Leitlinien zum Thema Forschungssoftware. Dazu wurden bereits 2017 Empfehlungen zur Implementierung von Leit- und Richtlinien zum Umgang mit wissenschaftlicher Software an den Helmholtz-Zentren erarbeitet.<sup>12</sup> Eine Muster-Richtlinie<sup>13</sup> für den nachhaltigen Umgang mit Forschungssoftware sowie eine Checkliste<sup>14</sup> bietet den Zentren eine Orientierung bei der Entwicklung solcher Policies.

In Kooperationen mit anderen Wissenschaftsorganisationen in Deutschland wurde 2018 außerdem eine „Handreichung zum Umgang mit Forschungssoftware“<sup>15</sup> sowie 2019 ein Positionspapier<sup>16</sup> erarbeitet.

In Zusammenarbeit mit sechs europäischen Forschungsorganisationen hat Helmholtz im G6-Netzwerk im Dezember letzten Jahres die Bedeutung der Forschungssoftware in dem „G6 statement on Open Science“ als Schlüssel für die moderne Forschung prominent betont. So heißt es dort: „Die G6-Mitglieder sind sich daher einig, dass Forschungssoftware denselben

---

<sup>11</sup> Siehe: <https://joss.theoj.org>

<sup>12</sup> Helmholtz-Gemeinschaft. (2019). Empfehlungen zur Implementierung von Leit- und Richtlinien zum Umgang mit Forschungssoftware an den Helmholtz-Zentren: Positionspapier. <https://doi.org/10.2312/os.helmholtz.008>

<sup>13</sup> Bach, F. et al. (2019): Muster-Richtlinie Nachhaltige Forschungssoftware an den Helmholtz-Zentren. <https://doi.org/10.2312/os.helmholtz.007>

<sup>14</sup> Messerschmidt, R. et al. (2021): Checkliste zur Unterstützung der Helmholtz-Zentren bei der Implementierung von Richtlinien für nachhaltige Forschungssoftware. <https://doi.org/10.48440/os.helmholtz.031>

<sup>15</sup> Katerbow, M., & Feulner, G. (2018). Handreichung zum Umgang mit Forschungssoftware. <https://doi.org/10.5281/zenodo.1172970>

<sup>16</sup> Konrad, U. et al. (2020). Digitale Dienste für die Wissenschaft - wohin geht die Reise?. <https://doi.org/10.5281/zenodo.4301924>

# HELMHOLTZ

## Open Science

Grundsätzen und strengen Anforderungen genügen muss, die Forscher an ihre Veröffentlichungen, Daten, Proben, Geräte und Infrastrukturen stellen.“<sup>17</sup>

Im Verein der Research Software Engineers in Deutschland (de-RSE)<sup>18</sup> vernetzen sich zudem die Entwickler:innen von Wissenschafts-Software. Der Verein hat jüngst mit dem campusSOURCE<sup>19</sup> Förderverein und dem Helmholtz Open Science Office, den „campusSOURCE Award 2022“ zur Förderung von Softwareentwicklung in der Forschung vergeben.<sup>20</sup> Mehrere der Gewinner kommen aus Helmholtz. Der erste Preis geht an eine Gruppe von Helmholtz Munich, die mit „cookietemple“ die Softwareentwicklung fördert. Der zweite und dritte Preis geht an das DLR, u. a. für den Standard „Citation File Format (CFF)“ zur Zitierung von Forschungssoftware.

Die Weichen für die Realisierung von Open Research Software sind gesetzt. Durch die immer stärkere Digitalisierung der Forschung wird es sich schnell zum Megathema von Open Science entwickeln.

---

<sup>17</sup> G6. (2021). G6 statement on Open Science. <https://os.helmholtz.de/open-science-in-der-helmholtz-gemeinschaft/nationale-internationale-vernetzung/g6-und-open-science/>

<sup>18</sup> Siehe: <https://de-rse.org>

<sup>19</sup> Siehe: <https://www.campussource.de>

<sup>20</sup> Siehe: <https://ev.campussource.de/publikationen/csa2022>

# HELMHOLTZ

## Open Science

### Programm

UHRZEIT	PROGRAMMPUNKT	REFERENT:INNEN	MODERATOREN
10:00 - 10:15	Begrüßung und Vorstellung des Forums	Uwe Konrad (HZDR, HIFIS), Wolfgang zu Castell (GFZ, AK Open Science)	Roland Bertelmann (Helmholtz Open Science Office)
10:15 - 10:30	Einführung - Stand und aktuelle Herausforderung	Sünje Dallmeier-Tiessen (CERN)	
<b>Themenblock Policy</b>			
10:30 - 10:50	Stand am HZDR	Uwe Konrad (HZDR)	
10:50 - 11:10	Stand am FZJ	Dennis Oliveira (FZJ)	
11:10 - 11:25	Pause		
<b>Themenblock Praxis</b>			
11:25 - 11:45	The Medical Imaging Interaction Toolkit (MITK)	Amir Kalali (DKFZ)	
11:45 - 12:05	Chemotion ELN	Nicole Jung (KIT)	
12:05 - 12:25	The mesoscale Hydrologic Model (mHM)	Sebastian Müller (UFZ)	
12:25 - 13:15	Pause		
<b>Themenblock: Infrastrukturen &amp; Tools</b>			
13:15 - 13:35	HIFIS	Tobias Huste (HZDR)	
13:35 - 13:55	HERMES	Tobias Schlauch (DLR)	
13:55 - 14:30	<b>Diskussion in drei parallelen Sessions</b>		
	Session 1: Policy-Erfahrungsautausch		Ronny Gey (UFZ)
	Session 2: Zukunft der Spotlights		Christian Meeßen (GFZ)
	Session 3: Infrastrukturen & Tools		Tobias Huste (HZDR)
14:30 - 14:45	Zusammenfassung und Verabschiedung		Roland Bertelmann (Helmholtz Open Science Office)

# HELMHOLTZ

## Open Science

### Vortragsfolien

1. Roland Bertelmann: Begrüßung
2. Sünje Dallmeier-Tiessen: Forschungssoftware
3. Amir Kalali: The Medical Imaging Interaction Toolkit
4. Nicole Jung: Chemotion Electronic Lab Notebook (ELN)
5. Sebastian Müller: The mesoscale Hydrological Model - mHM
6. Tobias Huste: Helmholtz Digital Services for Science — Collaboration made easy.
7. Tobias Schlauch HERMES: Automated software publication with rich metadata

**HELMHOLTZ**

Open Science

# Helmholtz Open Science Forum Forschungssoftware

07.04.2022, 10:00-14:45 Uhr

# Willkommen

---

- Die Veranstaltung wird von Helmholtz Forum Forschungssoftware veranstaltet.
- Das Helmholtz Forum Forschungssoftware wird gemeinsam von der Task Group Forschungssoftware des AK Open Science und dem HIFIS Software Cluster getragen.
- Die Veranstaltung wird vom Helmholtz Open Science Office organisiert.
- Erstes Forum zum Thema fand im Mai 2022 statt.
- Das Forum widmet sich drei Aspekten beim offenen und nachhaltigen Umgang mit Forschungssoftware in der Helmholtz:
  - Policy
  - Praxis
  - Infrastrukturen & Tools
- Wir freuen uns über Ihre Teilnahme!



# Organisatorisches

---

- Wir möchten Sie bitten Ihr Mikrofon während der Veranstaltung zu muten / stumm zu schalten, damit wir eine gute Tonqualität ermöglichen können.
- Bitte signalisieren Sie Wortbeiträge durch ein Ausrufezeichen im Chat, oder nutzen Sie die Chatfunktion für schriftliche Beiträge. Wir werden diese nach Möglichkeit berücksichtigen.
- Die Folien der Veranstaltung werden im Anschluss zugänglich gemacht.
- Die Veranstaltung findet mehrheitlich in deutscher Sprache statt. Fragen und Diskussionen sind natürlich auch in englischer Sprache möglich.

Uhrzeit	Programmpunkt	Referent:innen
10:00-10:15	Begrüßung und Vorstellung des Forums	Uwe Konrad (HZDR, HIFIS) Wolfgang zu Castell (GFZ, AK Open Science) Roland Bertelmann (Helmholtz Open Science Office)
10:15-10:30	Einführung - Stand und aktuelle Herausforderung Themenblock: Policy	Sünje Dallmeier-Tiessen (CERN)
10:30-10:50	Stand am HZDR	Uwe Konrad (HZDR)
10:50-11:10	Stand am FZJ	Dennis Oliveira (FZJ)
11:10-11:25	Pause Themenblock: Praxis	
11:25-11:45	The Medical Imaging Interaction Toolkit (MITK)	Amir Kalali (DKFZ)
11:45-12:05	Chemotion ELN	Nicole Jung (KIT)
12:05-12:25	The mesoscale Hydrologic Model (mHM)	Sebastian Müller (UFZ)
12:25-13:15	Pause Themenblock: Infrastrukturen & Tools	
13:15-13:35	HIFIS	Tobias Huste (HZDR)
13:35-13:55	HERMES Diskussion in drei parallelen Sessions	Tobias Schlauch (DLR)
13:55-14:30	Session 1: Policy - Erfahrungsaustausch	Ronny Gey (UFZ)
	Session 2: Zukunft der Spotlights	Christian Meeßen (GFZ)
	Session 3: Infrastrukturen & Tools	Tobias Huste (HZDR)
14:30-14:45	Zusammenfassung und Verabschiedung	Roland Bertelmann (Helmholtz Open Science Office)

# Themenblock: Policy

Pause bis 11:25 Uhr

# Themenblock: Infrastrukturen & Tools

Pause bis 13:15 Uhr

# Themenblock: Infrastrukturen & Tools

# Parallele Sessions

---

- Bitte ordnen Sie sich selbst dem jeweiligen Breakout-Raum zu!
  - **Breakout-Raum 1:** Policy - Erfahrungsaustausch
    - Ronny Gey (UFZ)
  - **Breakout-Raum 2:** Zukunft der Spotlights
    - Christian Meeßen (GFZ)
  - **Breakout-Raum 3:** Infrastrukturen & Tools
    - Tobias Huste (HZDR)
- Gerne können Sie die Räume wechseln.
- Sie werden um 14:30 Uhr automatisch wieder dem Plenum zugeordnet.



# Themenblock: Infrastrukturen & Tools

# Kommende Veranstaltungen

---

- 19. April: Helmholtz Open Science Online-Seminar zu „OpenAlex“
- 04. Mai: ORICD DE Workshop
- 12. Mai: Helmholtz Open Science Office zu „Open Science und Transfer“

# HELMHOLTZ

## Open Science

### Dialog

- E-Mail: [open-science@helmholtz.de](mailto:open-science@helmholtz.de)
- Twitter: [@helmholtz\\_os](https://twitter.com/helmholtz_os)
- Website: <https://os.helmholtz.de>
- Mailingliste für Mitarbeiter:innen von Helmholtz:  
[Helmholtz Open Science Professionals](#)
- [Helmholtz Open Science Newsletter](#)



# Forschungssoftware

**Dr Sünje Dallmeier-Tiessen**

CERN Scientific Information Service - Open Science

[sunje.dallmeier-tiessen@cern.ch](mailto:sunje.dallmeier-tiessen@cern.ch)





CERN - the world's biggest laboratory for particle physics.

**International Organization**  
established on 1 July 1953  
“Science for Peace”



**“...and the results of its  
experimental and theoretical  
work shall be published or  
otherwise made generally  
available”**

**CERN Founding Convention (1953)**

ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE  
CERN EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

C O N V E N T I O N

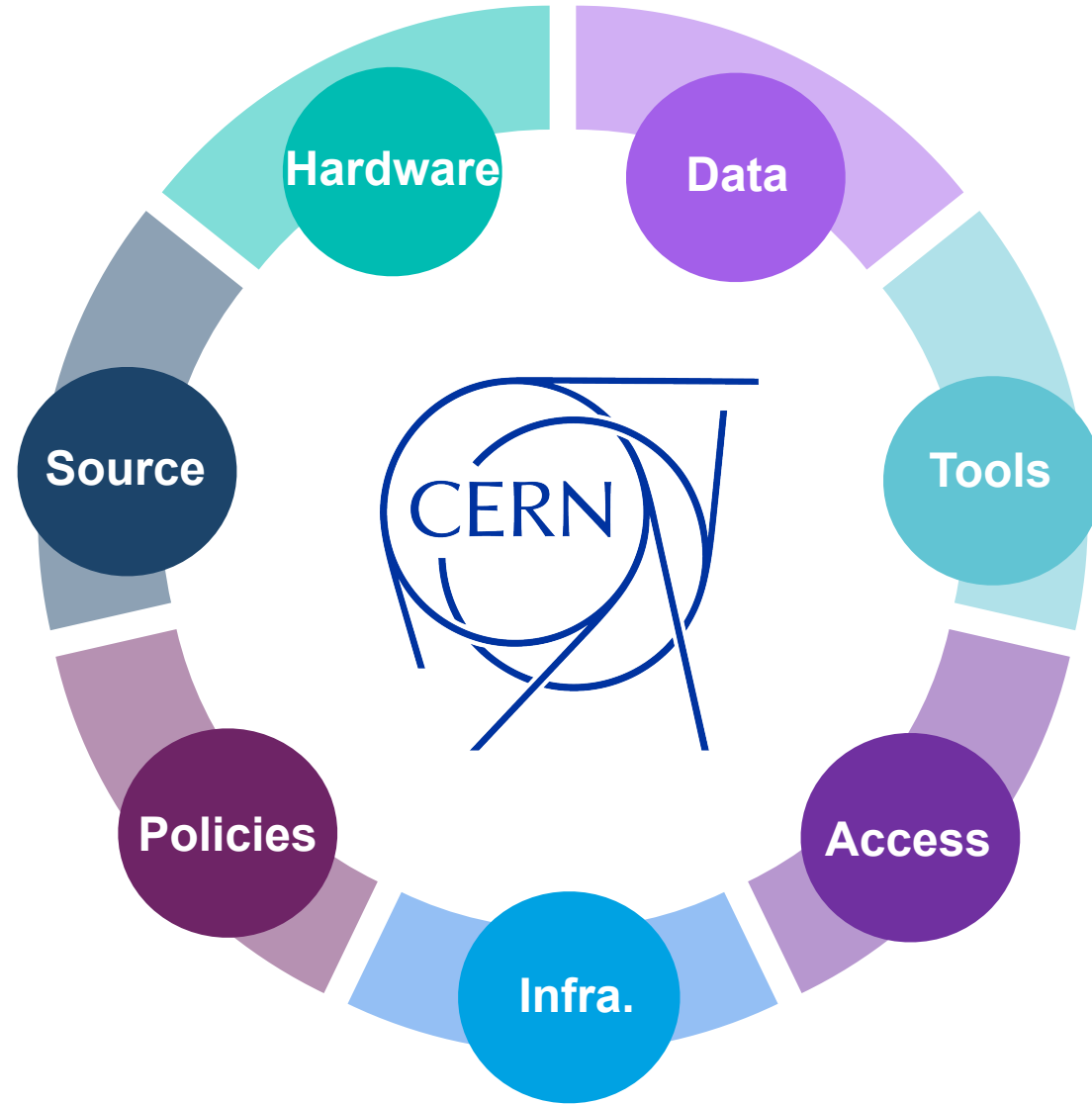
FOR THE ESTABLISHMENT OF A EUROPEAN ORGANIZATION  
FOR NUCLEAR RESEARCH

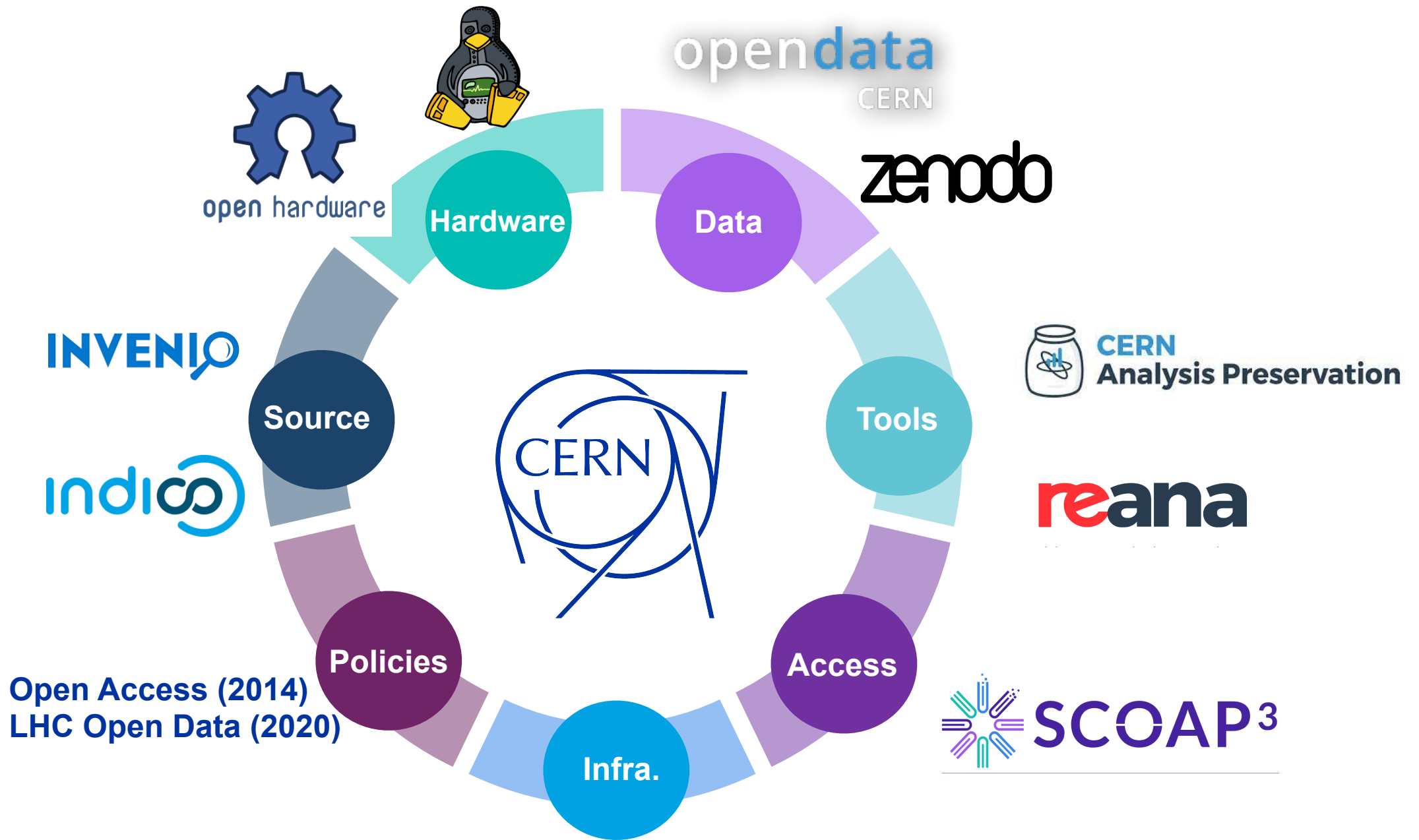
PARIS, 1st JULY, 1953



Illustration by Stephanie van de Sandt

# Open Science at CERN







# Research Software

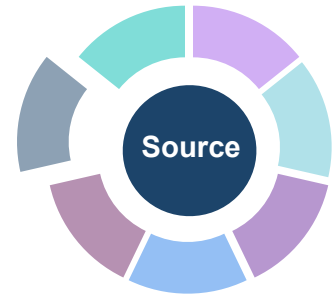
**Without data it's difficult to validate results.**

**But without code, we waste the opportunity to advance science.**

**— Neil Chue Hong, Director, UK Software Sustainability Institute**

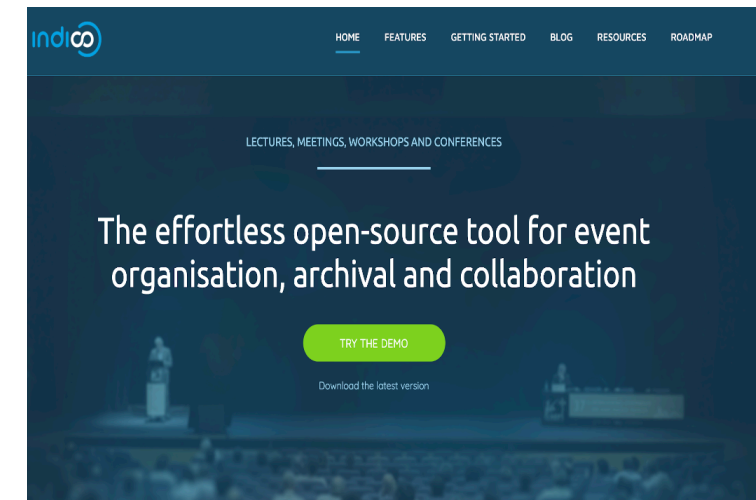
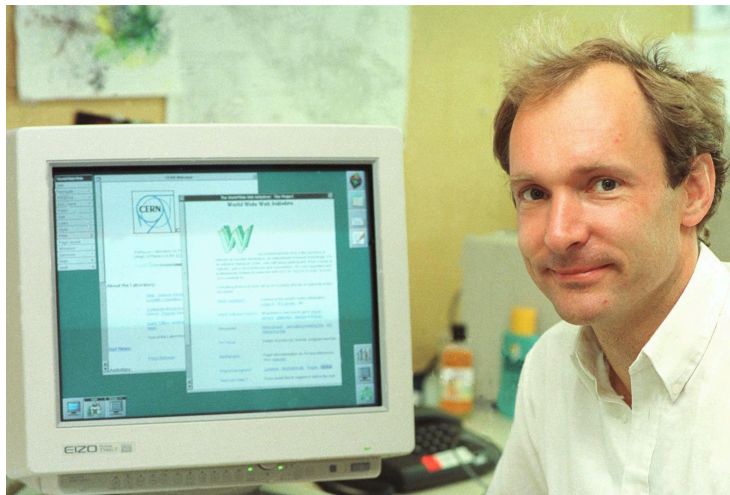
# Our Open-Source Philosophy

*“Recipients of technology should have access to all its building blocks, such as software code, schematics for electronics and mechanical designs, in order to study it, modify it and redistribute it to others”*

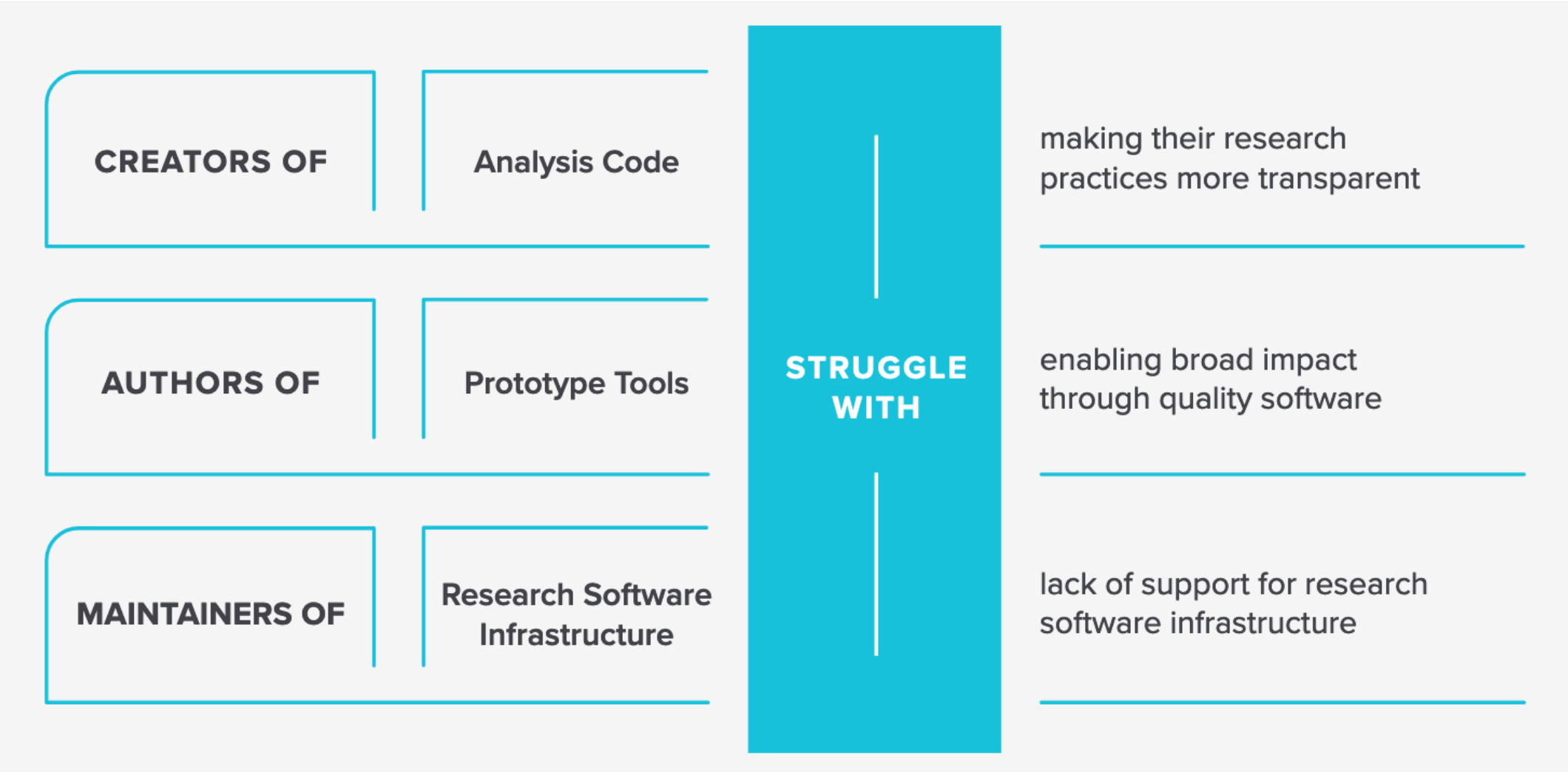


CERN Tradition of developing software and placing it in the public domain

- The www was created by Tim Berners-Lee at CERN
- Invenio is an open-source digital repository package
- Indico is an open-source online conferencing tool



# Research software: core challenges



# FAIR research software

## FAIR Principles for Research Software

The FAIR4RS WG define research software ([Gruenpeter et al., 2021](#)) as:

*Research Software includes source code files, algorithms, scripts, computational workflows and executables that were created during the research process or for a research purpose. Software components (e.g., operating systems, libraries, dependencies, packages, scripts, etc.) that are used for research but were not created during or with a clear research intent should be considered software in research and not Research Software. This differentiation may vary between disciplines.*

The FAIR4RS Principles are:

**F: Software, and its associated metadata, is easy for both humans and machines to find.**

F1. Software is assigned a globally unique and persistent identifier.

- F1.1. Components of the software representing levels of granularity are assigned distinct identifiers.
- F1.2. Different versions of the software are assigned distinct identifiers.

F2. Software is described with rich metadata.

F3. Metadata clearly and explicitly include the identifier of the software they describe.

F4. Metadata are FAIR, searchable and indexable.

**A: Software, and its metadata, is retrievable via standardized protocols.**

A1. Software is retrievable by its identifier using a standardized communications protocol.

- A1.1. The protocol is open, free, and universally implementable.
- A1.2. The protocol allows for an authentication and authorization procedure, where necessary.

A2. Metadata are accessible, even when the software is no longer available.

# Pragmatic next steps

## Four pillars: Archive, Reference, Describe, Credit

---

FAIR Guiding Principles do **not fit software** source code **well**

Focus on concrete issues to make software a first-class citizen:

Archive	ensure software artifacts are not lost
Reference	ensure software artifacts can be precisely identified
Describe	make it easy to discover / find software artifacts
Credit	ensure proper credit is given to authors

# Research software: a matter close to our heart

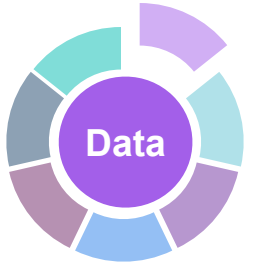
- Development, in particular collaborative development
- Quality Assurance and documentation
- Reusability
- Sustainability in times of budget cuts
- Licensing, e.g. Free Software Foundation
- Citation and reward, e.g. Zenodo workflow
- Training, Careers @CERN
- Currently: Policy development

## Let's work together!





# The 3 Pillars of Open Data



Sort by: **Best match** | asc. | Display: **detailed** | 20 results

Found 28 results.

**Configuration file for HLT step /cdqa/physics/Run2010HI/v1.5/HLT/V1**

The configuration file used in data taking and HLT data processing step in 2010. Run number after 151238 (appr.), software version

[Supplementaries](#) [Configuration](#) [CMS](#)

**CMS list of validated runs Cert\_161366-161474\_2760GeV\_PromptReco\_Collisions11\_JSON\_v2.txt**

This file describes which luminosity sections in which runs are considered good and should be processed when used as reference data for heavy-ion data analysis.

This list covers 2011 p-p dat...

[Environment](#) [Validation](#) [CMS](#)

zenodo

Search [ ] Upload Communities

lars.helm.nissen@cern.ch

Zenodo is continuing normal operation during the COVID-19 outbreak. All Zenodo staff are working remotely in accordance with preventive measures taken by CERN.

**COVID-19 related communities**

Need help? [Help](#) | [Contact us](#)

**Coronavirus Disease Research Community - COVID-19**

This community collects research outputs that may be relevant to the Coronavirus Disease (COVID-19) or the SARS-CoV-2. Scientists are encouraged to upload their outcome in this collection to facilitate sharing and discovery of information. Although Open Access articles and datasets are...

Curated by: Covid19\_Team, OpenAIRE

**Featured uploads related to COVID-19**

Want your dataset featured? [Contact us](#)

**Statistical review of Favipiravir versus Arbidol for COVID-19: A Randomized Clinical Trial**

Statistical review of Favipiravir versus Arbidol for COVID-19: A Randomized Clinical Trial

Wilkinson, Jack, Dahly, Damien

The following review has been prepared in collaboration with members of the MRC-NIHR Trials Methodology Research Partnership. The reviewers named above, and other, unnamed discussants of the...

Updated on April 3, 2020

**COVID-19 Open Research Dataset (CORD-19)**

Sullivan, Katherine, Kyle, Li, Luby, Lu, Wang, J, Hong

A full description of this dataset along with updated information can be found here. In response to the COVID-19 pandemic, the Allen Institute for AI has partnered with leading research groups to...

Updated on April 3, 2020

**Code for Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing**

Ferretti, Luca, Wymant, Chris, Fraser, Christophe

This code implements the COVID-19 mathematical analyses of Ferretti, Wymant et al. Science 2020. Namely, inference of the generation time interval for transmission pairs, solving the...

Updated on April 3, 2020

[Browse COVID-19 related research](#)

HEPData

Repository for publication-related High-Energy Physics data

Search on 9485 publications and 100131 data tables.

Search for a paper, author, experiment, reaction

Search Advanced

e.g. reaction: P P → LQ LQ X, title has "photon collisions", collaboration is LHCf or D0

**Data from the LHC**

**ATLAS** [View Data](#)

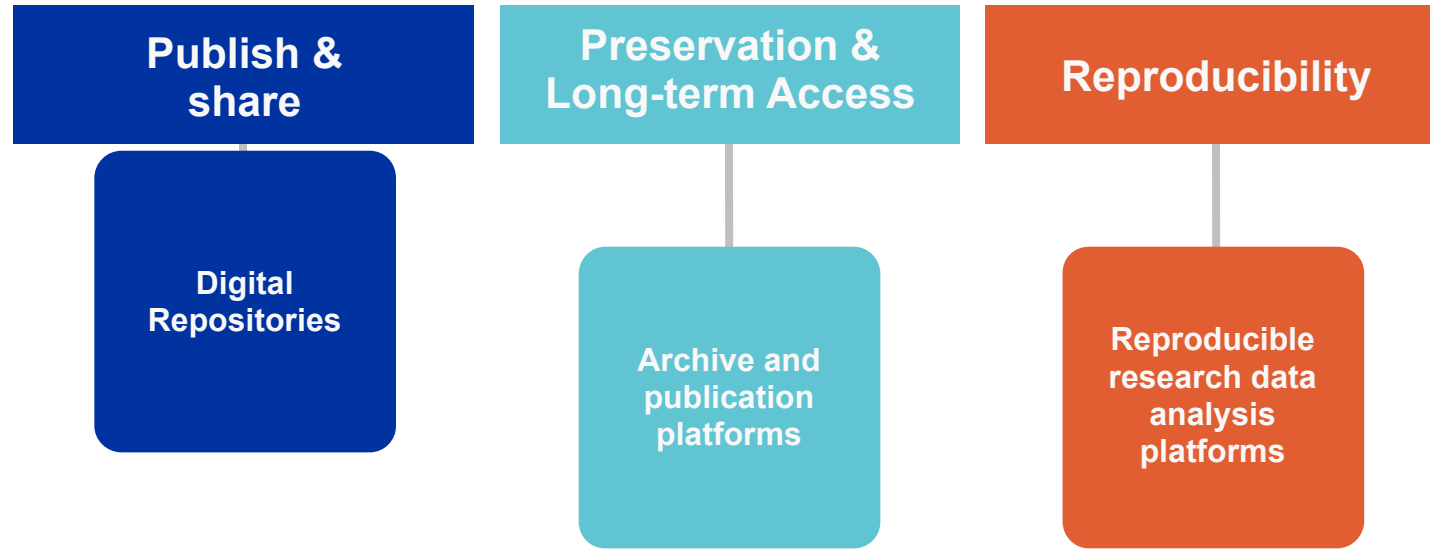
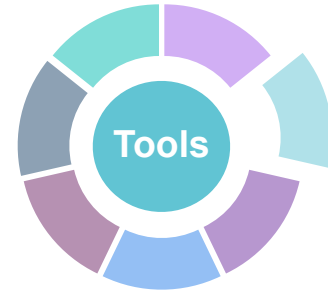
**ALICE** [View Data](#)

**CMS** [View Data](#)

**LHCb** [View Data](#)



# The 3 Pillars of Open Data



**CERN**  
Analysis Preservation






capture, preserve and reuse physics analyses



**reana**

Reproducible research data analysis platform

- Capture**  
  
 Collect and preserve elements needed to understand and rerun your analysis
- Collaborate**  
  
 Share your analysis and components with other users, your collaboration or group
- Reuse**  
  
 Run containerized workflows and easily reuse analysis components

- Flexible**  
 Run many computational workflow engines.  

- Scalable**  
 Support for remote compute clouds.  

- Reusable**  
 Containerise once, reuse elsewhere. Cloud-native.  


- Free**  
 Free Software. MIT licence. Made with ❤️ at CERN.  




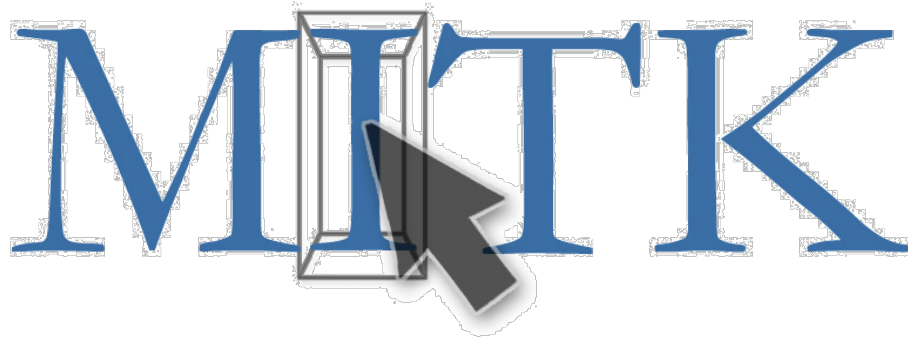
**dkfz.**

DEUTSCHES  
KREBSFORSCHUNGSZENTRUM  
IN DER HELMHOLTZ-GEMEINSCHAFT



Forschen für ein Leben ohne Krebs

# The Medical Imaging Interaction Toolkit



Amir Kalali  
Medical Image Computing (MIC), German Cancer Research Center, Heidelberg





## In-house development?

- All from scratch
- Design concepts
- Limited standardization & validation
- Medical imaging „ecosystem“ missing



## Commercial vendor software?

- Black-box
- No flexibility, configuration options
- Poor intersoftware-reproducibility amongst vendor softwares<sup>1</sup>

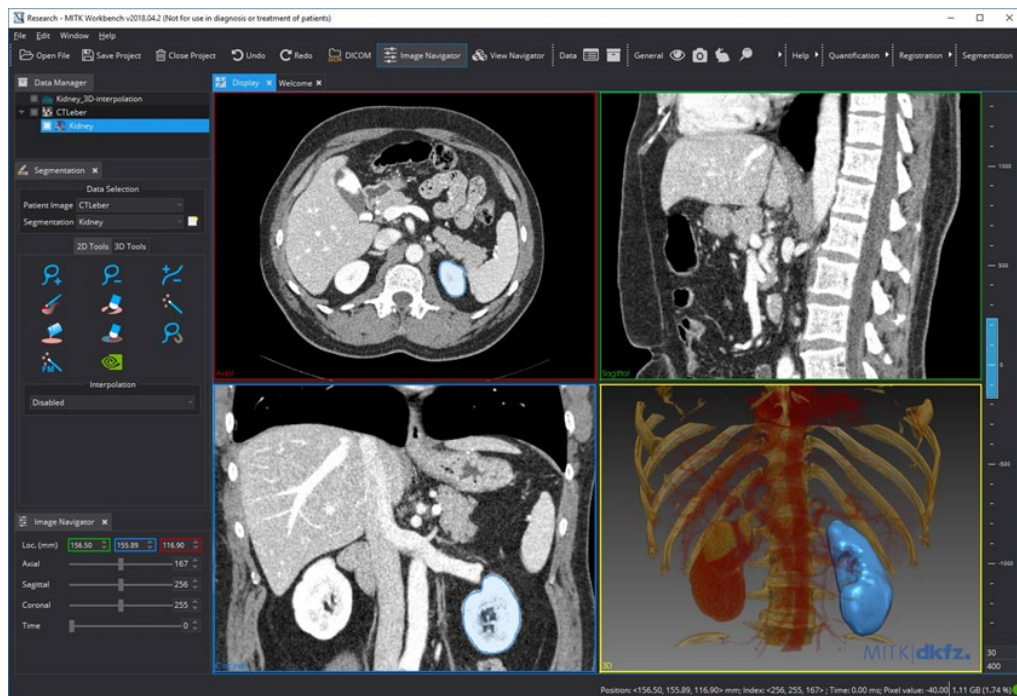


Open source software with a developer community



<sup>1</sup>Beuzit et al., *J Magn Reson Imaging*. 2016;43:1288-300.

## The Medical Imaging Interaction Toolkit (MITK)



- C++ framework / toolkit
  - MITK Workbench application
- Open source BSD-style license
- Highly modular
- DICOM import
  - CT, MR, US, PET, RT
- Visualization, Measurement, Image statistics
- Large segmentation toolbox
- Automated build process
  - Regular releases of installers  
(Windows, macOS, Ubuntu)

## Insights

Active development at DKFZ started in **2002**

More than **240 authors** contributed over  
**1.3 million lines of code** in nearly **90.000 commits**



## Paper citations

The medical imaging interaction toolkit  
*Wolf et al 2005*

442

The medical imaging interaction toolkit:  
Challenges and advances  
*Nolden et al 2013*

356



## Research software is mostly written by students

- Little to no professional experience
- Few dedicated software developers
- Typical time frame: 6 month – 3 years

**VS.**

## Requirements are highly demanding

- Scientific reproducibility
- Reusability, interoperability, and collaboration
- Complex integration scenarios



## Today MITK is developed by a core team

- Professional software development experience
- Using best practices and modern approaches
- Managing MITK's infrastructure
- Carefully reducing the complexity of the software



## DevOps as a set of practices

- Version control, issue tracking, code reviews
- Continuous integration
- Testing, packaging, releasing

**git****Version control***Open source code repository***GitHub****Single project management platform***Tickets, workboards, code review***Continuous integration***Build & unit test on multiple platforms***openstack.**



## Code guarantees for developers

*Quality levels, specified reviewer groups*



**git**

## Git branching model

*Modified Git-Flow, naming conventions*



## Task (lifecycle) management

*Issue tags, auto-cleanup*

## Master branch snapshots, Releases

*(In)official installers, public changelog*





## 2021 Week 39 (Very Early October)

Updated 170 Days Ago

Public

[← Previous changelog](#) • [Next changelog →](#)

### Third-party dependency changes

LZ4 is a new dependency providing extremely fast compression.

Dependency	Old version	New version
LZ4		1.9.3

### New features


- Creating new segmentations does not reset the scene anymore
- Accelerated 2-d interpolation of and undo/redo operations on huge segmentations
- Point sets can be visualized with fixed-size markers now (contribution by Sven Lafebre)
- Image and segmentations are selected automatically in the Segmentation views now
- Support latest versions of Visual Studio 2019
- Support latest GCC compiler versions in Ubuntu 20.04


### Bugfixes

- Fixed crash when using some segmentation utilities while segmentation interpolation is active
- Fixed crash when trying to load an empty DICOM directory
- Fixed loading of DICOM files from a directory that also contains other files with higher alphabetical order
- Fixed some interaction issues primarily related to the PACS interaction mode (ongoing work)
- Fixed GUI styling issues for active render windows

[← Previous changelog](#) • [Next changelog →](#)

 Edit Document

 View History

 Publish Draft


 Move Document

 Delete Document

 Printable Page

 Subscribe

 Mute Notifications

 Award Token

 Flag For Later

Tags

*None*

Subscribers

*None*

- Up
- Status
- People
- Build History
- Project Relationship
- Check File Fingerprint
- Open Blue Ocean

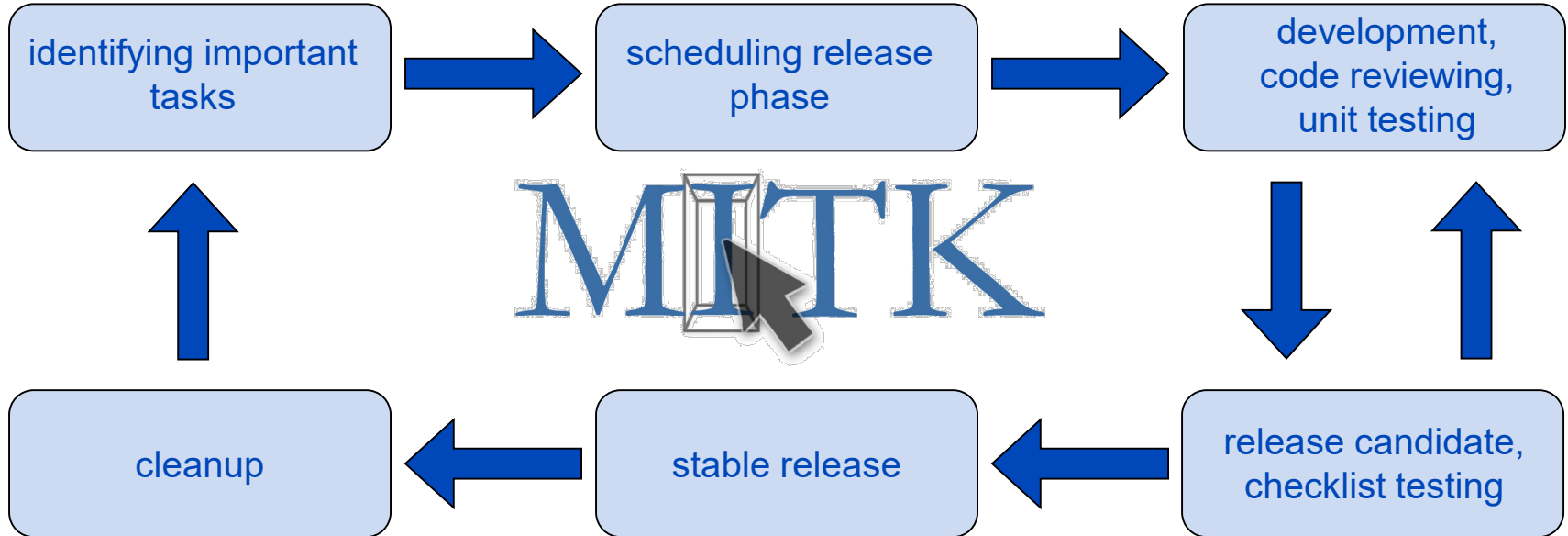
## Differential

All

S	W	Name ↓	Last Success	Last Failure	Last Duration	# Issues
✓	☁	macOS Catalina	21 hr <a href="#">#946</a>	22 hr <a href="#">#945</a>	4 min 45 sec	-
✓	🐧	Ubuntu 18.04	21 hr <a href="#">#950</a>	22 hr <a href="#">#947</a>	2 min 49 sec	-
✓	🐧	Ubuntu 20.04	21 hr <a href="#">#627</a>	22 hr <a href="#">#624</a>	3 min 19 sec	-
✓	⚙️	Windows	21 hr <a href="#">#938</a>	19 days <a href="#">#913</a>	12 min	-

# Continuous integration

MITK											
<span>Dashboard</span> <span>Calendar</span> <span>Previous</span> <span>Current</span> <span>Next</span> <span>Project</span>											
Testing data for this project can be found at: <a href="https://phabricator.milk.org/source/mitkdata/">https://phabricator.milk.org/source/mitkdata/</a>											
Continuous											5 builds
Site	Build Name	Update	Configure		Build		Test			Start Time	
		Revision	Error	Warn	Error	Warn	Not Run	Fail	Pass		
e071-jenkins	Windows Visual Studio 2019 16.11.3 Release All	6eae25	0	0	0	0	0	2 <sup>+2</sup>	375 <sub>-2</sub>	Mar 19, 2022 - 09:32 UTC	
e071-jenkins	macOS Big Sur 11.4 Clang 12.0.5 Release All	6eae25	0	0	0	51 <sup>+43</sup> <sub>-43</sub>	0	3 <sup>+2</sup>	374 <sub>-2</sub>	Mar 19, 2022 - 09:32 UTC	
e071-jenkins	Ubuntu 18.04 GNU 7.5 Release All	6eae25	0	0	0	3 <sup>+1</sup> <sub>-1</sub>	0	2 <sup>+2</sup>	433 <sub>-2</sub>	Mar 19, 2022 - 09:32 UTC	
e071-jenkins	Ubuntu 20.04 GNU 9.4 Release All	6eae25	0	0	0	10 <sup>+5</sup> <sub>-1</sub>	0	2 <sup>+2</sup>	433 <sub>-2</sub>	Mar 19, 2022 - 09:32 UTC	
e071-jenkins	macOS Catalina 10.15.7 Clang 12.0 Release All	6eae25	0	0	0	51 <sup>+40</sup> <sub>-40</sub>	0	2 <sup>+2</sup>	375 <sup>+1</sup> <sub>-2</sub>	Mar 19, 2022 - 09:32 UTC	
Nightly											5 builds
Site	Build Name	Update	Configure		Build		Test			Start Time	
		Revision	Error	Warn	Error	Warn	Not Run	Fail	Pass		
e071-jenkins	macOS Big Sur 11.4 Clang 12.0.5 Release All	145223	0	0	0	51 <sup>+1</sup> <sub>-1</sub>	0	2 <sup>+1</sup>	376 <sub>-1</sub>	Mar 19, 2022 - 00:00 UTC	
e071-jenkins	Windows Visual Studio 2019 16.11.3 Release All	145223	0	0	0	50	0	1	377	Mar 19, 2022 - 00:01 UTC	
e071-jenkins	macOS Catalina 10.15.7 Clang 12.0 Release All	145223	0	0	0	51 <sup>+1</sup> <sub>-1</sub>	0	0	378	Mar 19, 2022 - 00:00 UTC	
e071-jenkins	Ubuntu 18.04 GNU 7.5 Release All	145223	0	0	0	50 <sup>+1</sup> <sub>-1</sub>	0	0	436	Mar 19, 2022 - 00:00 UTC	
e071-jenkins	Ubuntu 20.04 GNU 9.4 Release All	145223	0	0	0	31 <sup>+1</sup> <sub>-1</sub>	0	0	436	Mar 19, 2022 - 00:00 UTC	
Release Continuous											4 builds
Site	Build Name	Update	Configure		Build		Test			Start Time	
		Revision	Error	Warn	Error	Warn	Not Run	Fail	Pass		
e071-jenkins	Windows Visual Studio 2019 16.11.3 Release WorkbenchRelease release/T28985-2022-Week-07	7872f0	0	0	0	0	0	0	306	Feb 18, 2022 - 08:28 UTC	
e071-jenkins	Ubuntu 20.04 GNU 9.3 Release WorkbenchRelease release/T28985-2022-Week-07	7872f0	0	0	0	36	0	0	362	Feb 18, 2022 - 08:28 UTC	
e071-jenkins	Ubuntu 18.04 GNU 7.5 Release WorkbenchRelease release/T28985-2022-Week-07	7872f0	0	0	0	4	0	0	362	Feb 18, 2022 - 08:28 UTC	
e071-jenkins	macOS Mojave 10.14.6 Clang 10.0.1 Release WorkbenchRelease release/T28473-2021-Week-17	f1783c	0	0	0	45	0	0	365	Apr 30, 2021 - 10:40 UTC	



## Current and future topics

- License management of dependencies
- Automated UI testing
- Migration to GitLab
  
- Establishing SE expertise
- Publication process (DOI)
- Community management



Danke für eure  
Aufmerksamkeit!



<https://www.mitk.org> | <https://www.dkfz.de/en/mic>

**dkfz.** DEUTSCHES  
KREBSFORSCHUNGSZENTRUM  
IN DER HELMHOLTZ-GEMEINSCHAFT



Forschen für ein Leben ohne Krebs

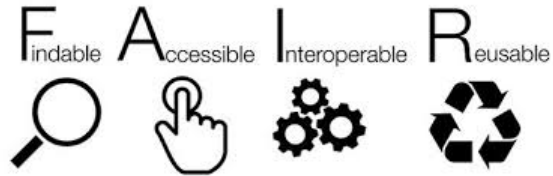
# Chemotion Electronic Lab Notebook (ELN)

**Stefan Bräse Group, Nicole Jung**  
**KIT Karlsruhe**



# RDM – Current Situation

**F**indable  
**A**ccessible  
**I**nteroperable  
**R**e-usable (Reproducible)



**not FAIR  
and  
not Open**



DATA FOR FUTURE GENERATIONS

Taken from: Auke Herrema – Het Bouwteam

# RDM – Current Situation

## Publications and research data

Journal articles and research data must be linked via persistent identifiers (DOI). Research data must be machine-readable and reusable.

4-Synthase (on Solid-Phase)

4.1 PS-methyl-tryptamine (3)

CN(C)Cc1ccc(C)cc1

Method: each (5.20 g, 1.05 mmol), 6.32 g (0.10 mol) was used for the synthesis. The reaction mixture was stirred at 60 °C for 16 hours. The colored crude material was washed according to 20091. Yield: 1.80 g (35%).

Peak	Time [min]	Retention [min]	Yield [g]	Purity [%]	Yield [%]	Purity [%]
1.1	2.20	2.20	1.80	7.20	7.20	7.20
1.2	4.00	4.00	1.50	1.50	1.50	1.50
1.3	5.70	5.70	1.50	5.70	5.70	5.70
1.4	6.70	6.70	1.50	6.70	6.70	6.70
1.5	7.60	7.60	1.50	7.60	7.60	7.60
1.6	1.10	1.10	1.50	1.10	1.10	1.10
1.7	1.70	1.70	1.50	1.70	1.70	1.70
1.8	1.70	1.70	1.50	1.70	1.70	1.70
1.9	1.70	1.70	1.50	1.70	1.70	1.70
2.0	1.70	1.70	1.50	1.70	1.70	1.70
2.1	1.70	1.70	1.50	1.70	1.70	1.70
2.2	1.70	1.70	1.50	1.70	1.70	1.70
2.3	1.70	1.70	1.50	1.70	1.70	1.70
2.4	1.70	1.70	1.50	1.70	1.70	1.70
2.5	1.70	1.70	1.50	1.70	1.70	1.70
2.6	1.70	1.70	1.50	1.70	1.70	1.70
2.7	1.70	1.70	1.50	1.70	1.70	1.70
2.8	1.70	1.70	1.50	1.70	1.70	1.70
2.9	1.70	1.70	1.50	1.70	1.70	1.70
3.0	1.70	1.70	1.50	1.70	1.70	1.70

4.2 PS-methyl-5-hydroxytryptamine (5)

CN(C)Cc1ccc(O)cc1

Method: each (5.20 g, 1.05 mmol), 6.32 g (0.10 mol) was used for the synthesis. The reaction mixture was stirred at 60 °C for 16 hours. The colored crude material was washed according to 20091. Yield: 1.80 g (35%).

not FAIR  
and  
not Open



PUBLICATIONS AND DATA  
Taken from: Auke Herrema – Het Bouwteam

---

Important part of the solution:  
Electronic Lab Notebooks

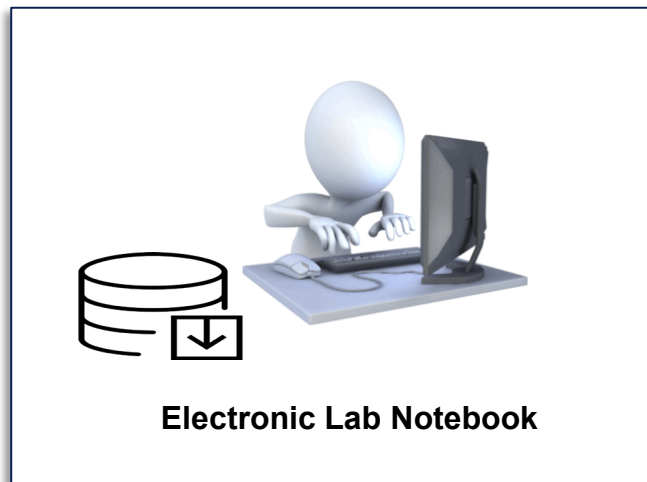
### Clarification

- The benefit of an ELN is not to have data in a digital form
- The benefit of an ELN comes with the option to generate systematically standardized data



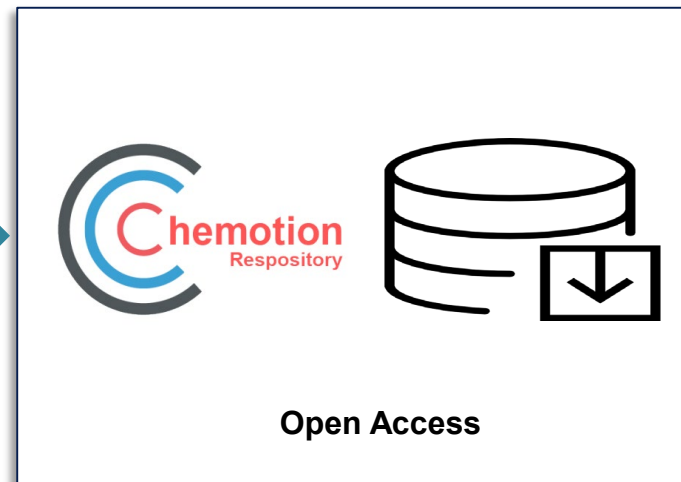
# Smart Lab – From ELN to Repository

## Documentation



select & transfer

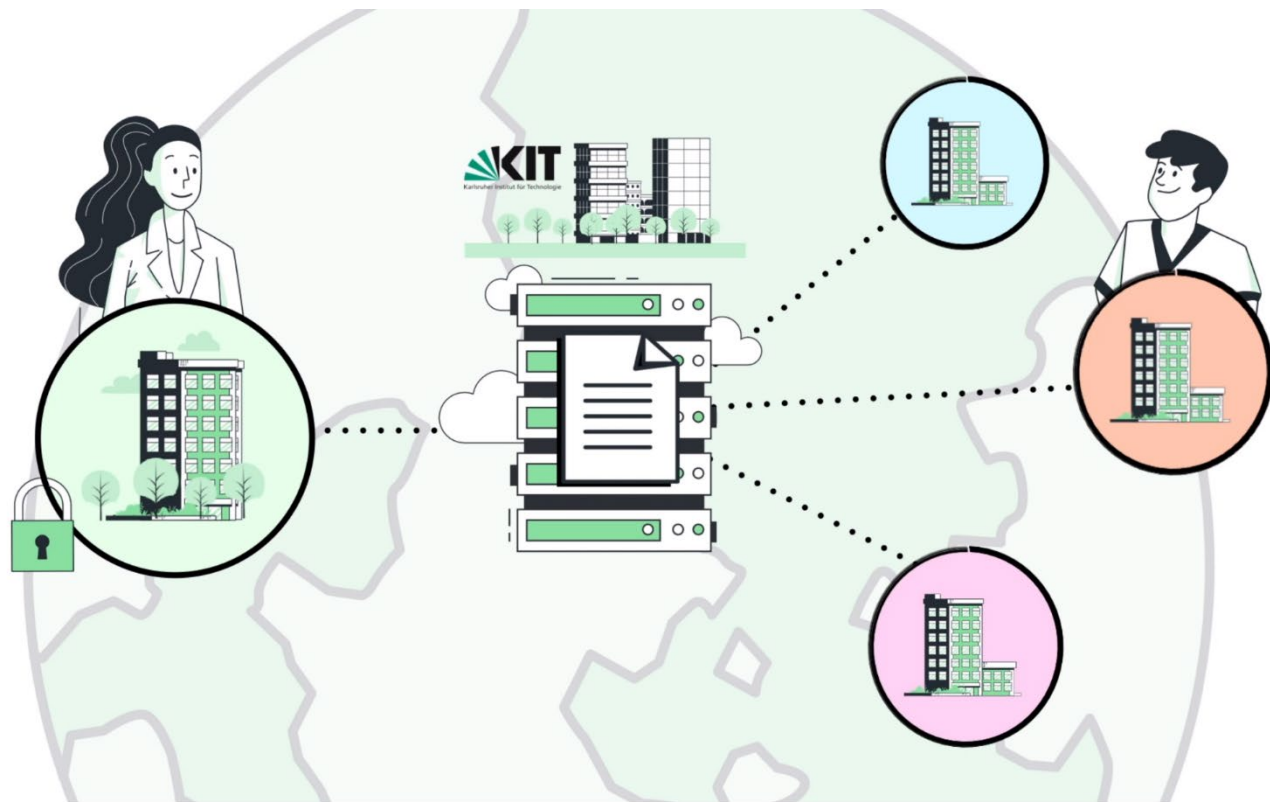
## Publication



Gif-Source: <https://gifer.com/en/PsKr>

**A few clicks to Open Data**

# Smart Lab – From ELN to Repository



# Electronic Lab Notebook – systematic view

---

## Different Parts:

- Access functions and infrastructure aspects
- Management
- Sharing and collaborative work
- Documentation
- Structure - dependencies
- Information
- Data analysis and conversion





# Chemotion – ELN for chemical synthesis

Chemotion(Testing) - All - IUPAC, InChI, SMILES, RIn

(Testing Instance) te 1

Collections: 6(0) 2(0) 21(0)

chemotion-repository.net

SPP Test Reactions

Test Collection Manuel Tsotsalas

SurMOF Processes

Test reaction Nicole

Test Collection Alex

Test Steffen

Evgenia

Favoriten

Test collection

Richard Thelen

My shared collections

Shared with me

Synchronized with me

Inbox

tet-R27

From To

tet-R27

tet-R29

tet-R28

tet-R26

tet-R25

tet-R24

1

Show 15

tet-R27

1.0

Scheme Catalysis

Starting materials

Ref	L/ST/RAmount	Conc	Equiv
A tet-101 NIW-c1122-A (E)-3-(3,3-diisopropyltriaz-1-en-1-yl)-	75.50 mg 0.00 ml 0.3607 mmol	120.2 mmol/l	1.000

Reactants

Reagents	L/ST/RAmount	Conc	Yield
B Cesium Carbon...	141.0 mg 0.00 ml 0.4329 mmol	144.3 mmol/l	1.200
C bromomethylbenz...	123.4 mg 0.0858 ml 0.7215 mmol	240.5 mmol/l	2.000

Products

L/ST/RAmount	Conc	Yield
P1 tet-99 NIW-c1146-A (E)-1-benzyl-5-(3,3-diisopropyltriaz-1-en-1-yl)-	36.60 mg 0.00 ml 0.1229 mmol	40.97 mmol/l 34%
P2 tet-100 NIW-c1146-B (E)-1-benzyl-3-(3,3-diisopropyltriaz-1-en-1-yl)-	28.90 mg 0.00 ml 0.09652 mmol	32.17 mmol/l 27%

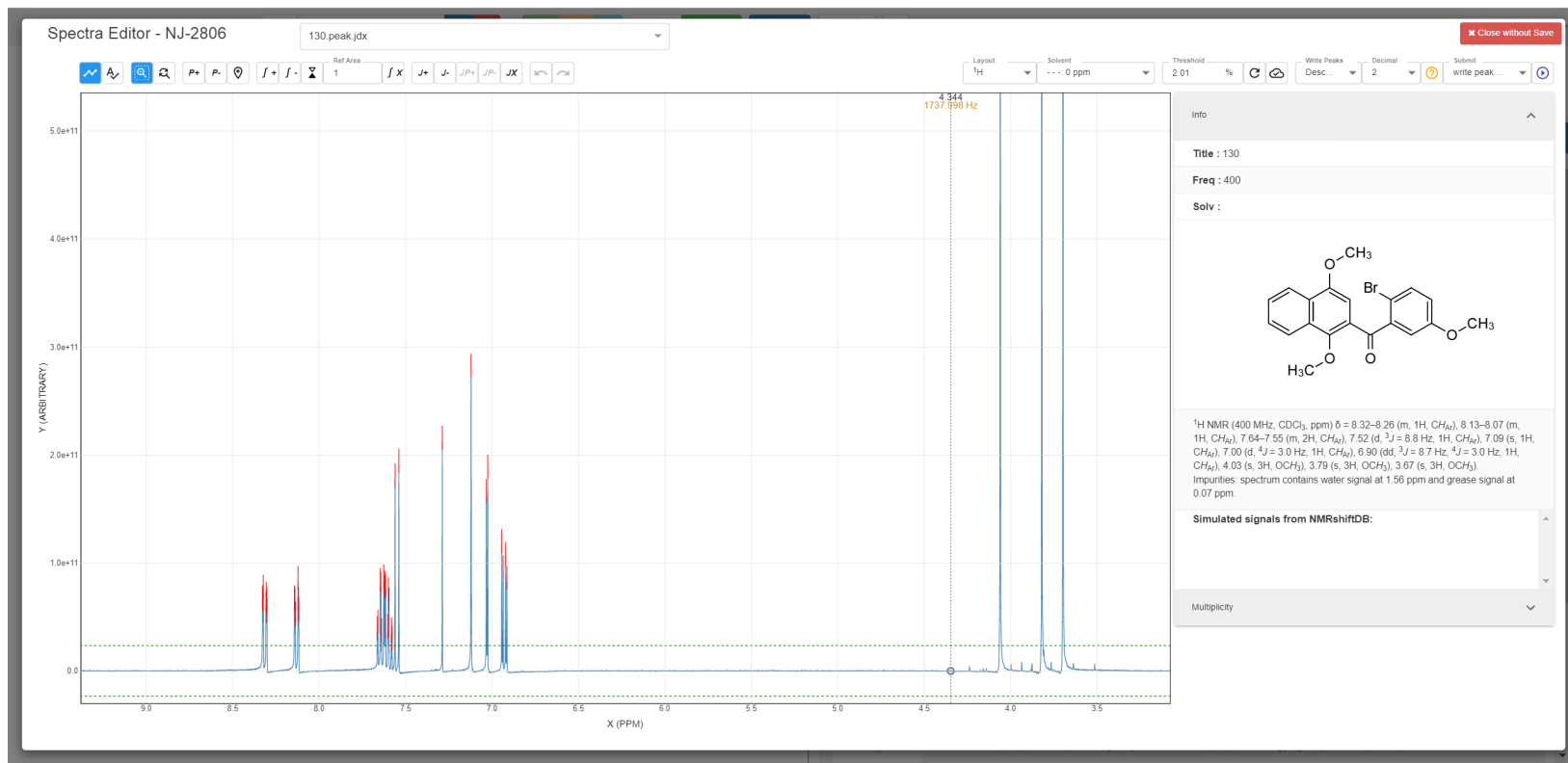
Solvents

Default solvents	T/R	Label	Vol	Vol ratio
S1 DMSO	1	DMSO	3.00 ml	100.0%

Conditions

Name: tet-R27 Status: Successful Temperature: 0-21 °C

# Visualize data in the ELN



# Chemotion – ELN for chemical synthesis

VLL-R14

VLL-R14

1H-quinoxalin-2-one + tetrabutylazanium bromide + phosphorus pentoxide  $\xrightarrow[100\text{ }^\circ\text{C, 2.5 hr}]{\text{Toluene}}$  2-bromanylquinoxaline (96%)

Scheme Properties Analyses References

Starting materials		Ref	L/S	T/R	Amount	Conc			
+	A VLL-33 1H-quinoxalin-2-one	⊙	s	t	2000 mg	0.00 ml	13.68 mmol	195.5 mmol/l	1.000

Reactants		Reagents		Amount	Conc				
+	B tetrabutylazanium;brom...	○	s	t	5735 mg	0.00 ml	17.79 mmol	254.1 mmol/l	1.300
+	C phosphorus pento...	○	s	t	6216 mg	0.00 ml	21.90 mmol	312.8 mmol/l	1.600

Products		L/S	T/R	Amount	Conc			
+	P1 VLL-32 VLL-072-... 2-bromanylquinoxaline	s	r	2749 mg	0.00 ml	13.15 mmol	187.9 mmol/l	96%

⌵ Solvents

⌵ Conditions

Name: Name...

Status: Successful

Temperature: 100 °C

# Chemotion – a flexible structure

The screenshot displays the Chemotion software interface. On the left is a sidebar with a 'Collections' panel containing various categories like 'chemotion.net', 'Test Collection Data', 'MOF examples', etc. The main workspace shows a list of reactions, with 'NJ-R1515' selected. A detailed view of this reaction is shown on the right, featuring a chemical reaction scheme where 1H-quinoxalin-2-one reacts with POCl<sub>3</sub> to form 2-chloroquinoxaline in 49% yield. Below the reaction is a data table with the following content:

Starting materials	Ref	L	S	T	R	A	Amount	Conc	Equiv
NJ-2657 1H-quinoxalin-2-one		S	L	T	R	A	100.0 mg 0.00 ml	0.6842 mmol	n.d. mmol/l 1.000

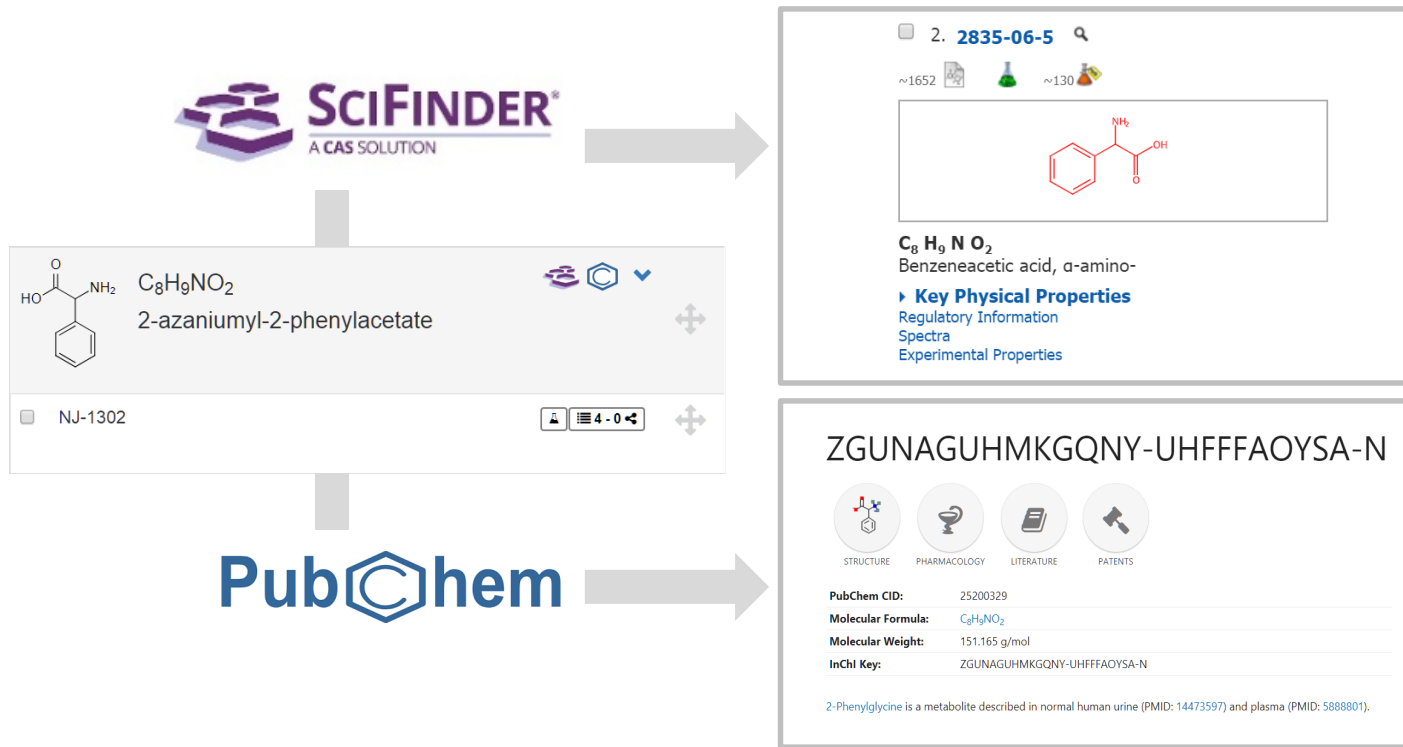
Reactants	Reagents	L	S	T	R	A	Amount	Conc	Equiv
ClOP		S	L	T	R	A	104.9 mg 0.00 ml	0.6842 mmol	n.d. mmol/l 1.000

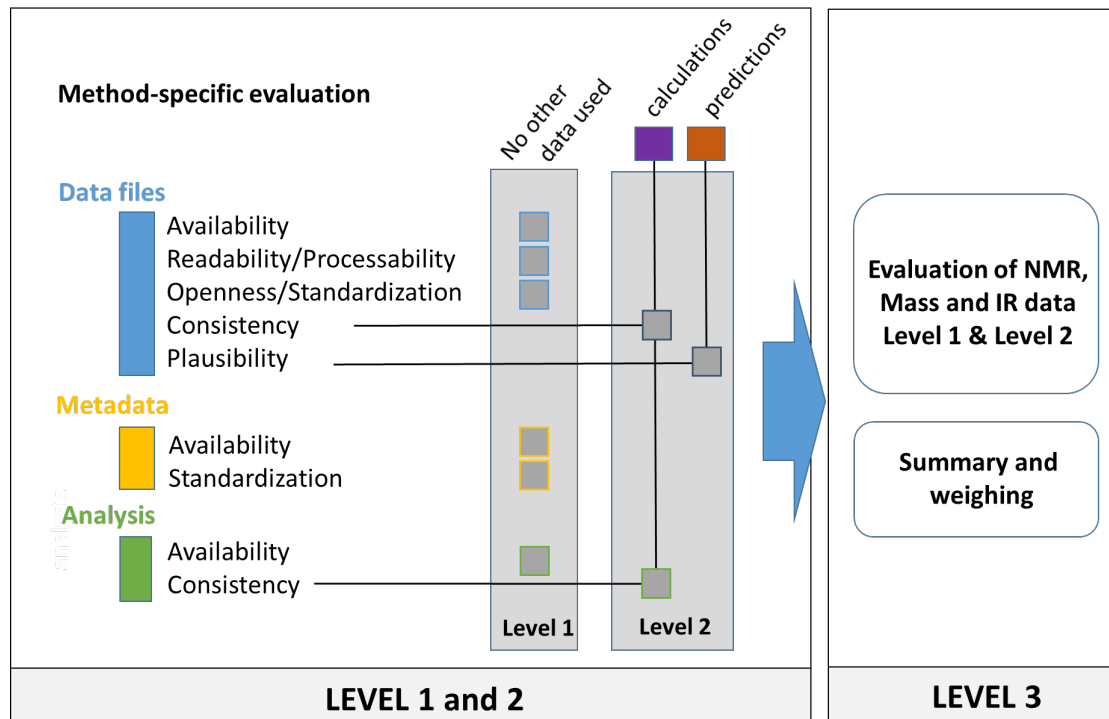
Products	L	S	T	R	A	Amount	Conc	Yield
NJ-2658 NJ-R1515... 2-chloroquinoxaline	S	L	T	R	A	55.00 mg 0.00 ml	0.3342 mmol	n.d. mmol/l 49%

Below the table are fields for Name, Status, Temperature, Start, Stop, Duration, Type (Name Reaction Ontology), Role, and Description.

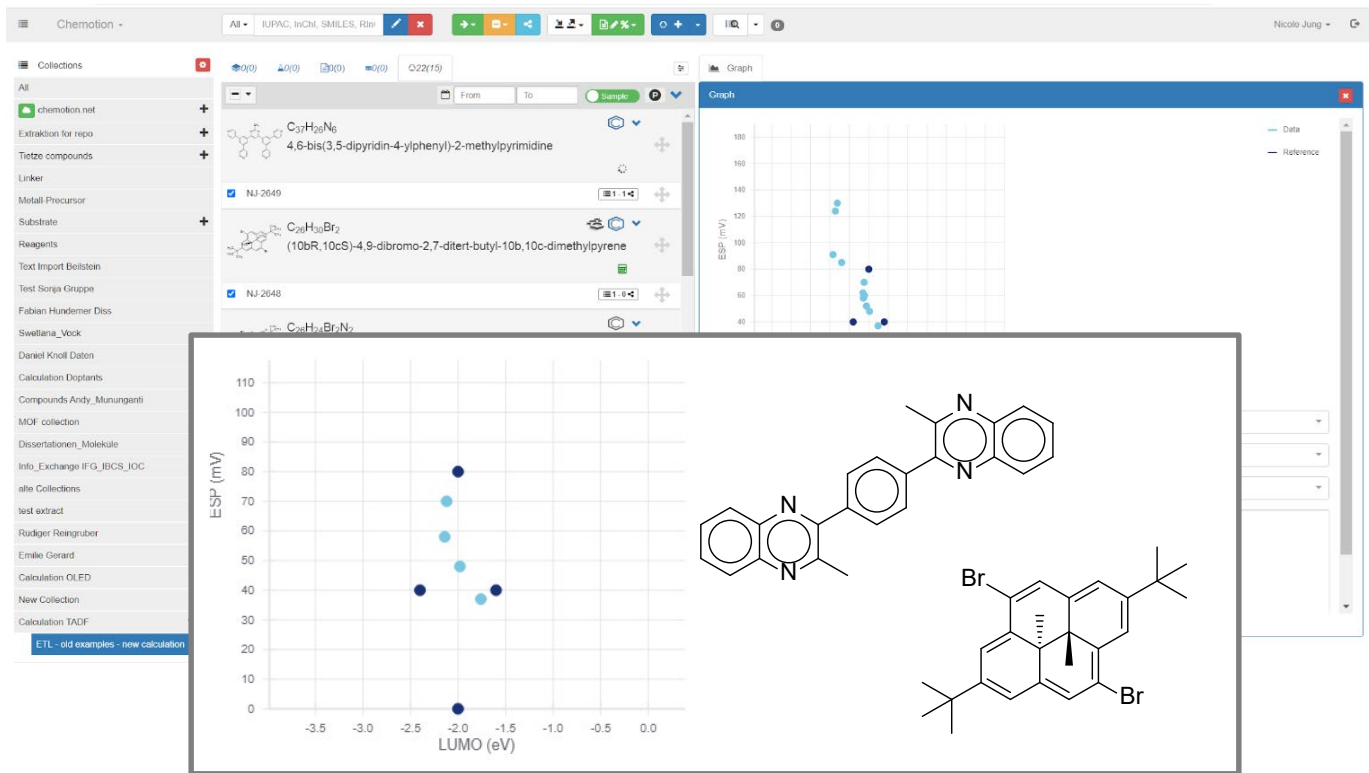
# Chemotion – link to external databases



# Chemotion – Quality aspects



# Chemotion – calculation and simulation



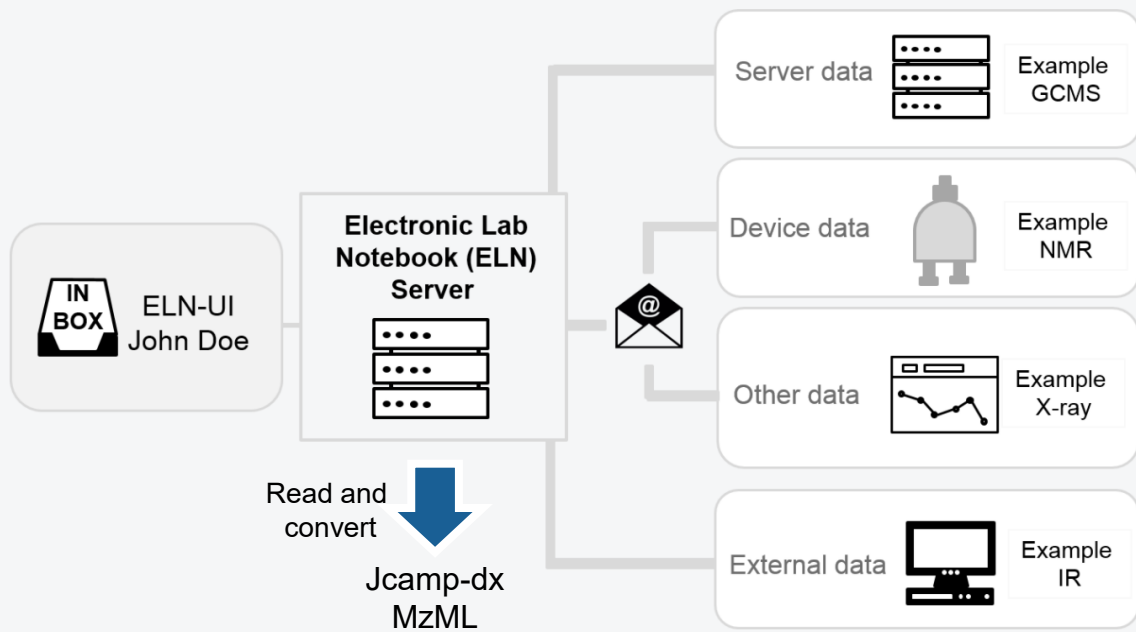
# Chemotion – integration of ML tools

The screenshot displays the Chemotion web interface. The browser address bar shows `complat-eln.ioc.kit.edu/mydb/collection/2206/`. The interface is divided into several sections:

- Left Panel:** A sidebar with a "Collections" menu. The "Linkor" collection is selected and highlighted in blue. Other collections listed include "chemotion.net", "Extraktion for repo", "Tietze compounds", "Metall-Precursor", "Substrate", "Reagents", "Text Import Beilstein", "Test Sonja Gruppe", "Fabian Hundemer Diss", "Svetlana\_Veck", "Daniel Knoll Daten", "Calculation Doptants", "Compounds Andy\_Mununganti", "MOF collection", "Dissertationen\_Molekule", "Info\_Exchange IFG\_IBCS\_IOC", "alte Collections", "test extract", "Rüdiger Reingruber", "Emilie Gerard", "Calculation OLED", "New Collection", "Calculation TADF", and "QC tool test cases".
- Main Table:** A table listing chemical compounds with their molecular formulas and names. The second entry, "4-(cyclobutanecarbonylamino)-2-nitrobenzoic acid", is highlighted with a yellow circle. The table includes columns for chemical structure, molecular formula, name, and a "Sample" button.
- Right Panel:** A "Synthesis Prediction" tool. It features a "Content" field, a "Predict Starting Materials" dropdown menu, an "Input: Product (accept max 1 molecule)" text area containing the message "Please select samples.", and an "Output: Starting Materials" text area.



# Chemotion – device integration



## Procedures for systematic capture and management of analytical data in academia

Jan Potthoff<sup>a</sup>, Pierre Tremouilhac<sup>b</sup>, Patrick Hodapp<sup>b</sup>, Bernhard Neumair<sup>a</sup>, Stefan Bräse<sup>b,c,d</sup>, Nicole Jung<sup>b,c,d</sup>

<sup>a</sup> Institute of Organic Chemistry, Karlsruhe Institute of Technology, Hermann von Helmholtz Platz 1, 76344 Eggenstein-Leopoldsdorf, Germany  
<sup>b</sup> Institute of Robotics and MeCS, Karlsruhe Institute of Technology, Hermann von Helmholtz Platz 1, 76344 Eggenstein-Leopoldsdorf, Germany  
<sup>c</sup> Institute of Organic Chemistry, Karlsruhe Institute of Technology, Prof. Huber Weg 6, 76131, Karlsruhe, Germany

### ARTICLE INFO

**Article history:**  
 Received 2 August 2018  
 Received in revised form 2 January 2019  
 Accepted 29 January 2019  
 Available online 15 February 2019

**Keywords:**  
 Research data management  
 Information and management systems  
 Analytical data  
 ELN

### ABSTRACT

Data management in universities is a challenging endeavor in particular due to the diverse infrastructure of devices and software in combination with limited budget. Nevertheless, in particular the analytical measurements and data sets need to be stored if possible digitally and in a well-organized manner. This manuscript describes how scientists can achieve a data management workflow focusing on data capture and storage by small adaptations to commonly used systems. The presented method includes data transfer systems from ubiquitous devices like NMR instruments, GC (MS) or LC (MS), IR and Raman, or mass spectrometers to a central server and the visualization of the available data files in an electronic lab notebook (ELN). The given instruments were chosen according to the needs of synthetic chemists, in particular devices needed to organic, inorganic and polymer chemistry where single data files in the range of several megabytes per data set are produced. Altogether, three different data transfer systems were elaborated to allow a flexible handling of different devices running with different proprietary software: The first procedure shows data capture via the use of a mail server as data exchange point. With the second procedure, data are automatically mirrored from a local file folder to a central storage server where new files are monitored and processed. The third procedure was designed to transfer data with manual support to a central server which is supported to register new information. All components that are necessary to install and use the herein elaborated functions are available as Open Source and the designed workflow are described step by step to facilitate the adoption of procedures in other universities accordingly if desired.

© 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

### 1. Introduction

Systematic data collection and management is of high importance in research institutions, in particular due to the increasing requirements to develop and meet research data management plans. Systems that support researchers in this aim contribute to the improvement of the overall data structure and help scientists to keep the overview of analytical files. Additionally, management systems for analytical data allow the systematic storage of digital

data which is of high importance e.g. for the accessibility of research data in compliance with the FAIR data [1] principles. In contrast to this need, the systematic capture and storage of analytical data in natural sciences and in experimental or analytical laboratories can be a challenging endeavor even for non-big data research due to the manifold sources of data and the use of different devices for their acquisition. Therefore, larger companies usually invest in Laboratory and Information Management Systems (LIMS) which allow to control the data workflow, data tracking, and data transfer. Some of the well-known systems are STARLIMS (Abbott Informatics Corporation) [2], Lintology (ABC Informaty) [3], Biowis LIMS (Accepio) [4], SampleManager (Thermo Fisher Scientific) [5], webLIMS (Labaya) [6] or LaboLIMS (Computing Solutions) [7]. A LIMS allows to reduce the errors that are inevitably associated with manual data handling. It prevents the loss of data, and improves the work efficiency. In particular with a high throughput of data, the

\* Corresponding author. Institute of Organic Chemistry, Karlsruhe Institute of Technology, Prof. Huber Weg 6, 76131, Karlsruhe, Germany.  
 E-mail address: [braese@kit.edu](mailto:braese@kit.edu) (S. Bräse), [nicole.jung@kit.edu](mailto:nicole.jung@kit.edu) (N. Jung).

---

Flexible applications beyond chemistry

**Lab**  **Motion**

# Flexible applications

Chemotion ▾

ELN Administration

ELN Admin ▾

Dashboard

User Management

Message Publish

Data Collector

Groups & Devices

Load OLS Terms

NoVNC Settings

UI features













































Text Templates

Generic Elements (BETA)

**Generic Segment (BETA)**

Generic Dataset (BETA)

New Segment +

#	Actions	Segment	Description	Active	Template	Belongs to
1	 	MOF reaction	MOF reaction details	✓	 	Reaction ▾
2	 	fava_pharmatwo	Pharmacokinetics	✗	 	Sample ○
3	 	plants	Plants	✗	 	Research Plan 📄
4	 	fava_pharmaone	Pharmaceutical details	✗	 	Sample ○
5	 	spectro_pc	Fluorescence	✓	 	Sample ○
6	 	photophysics	Photophysics	✓	 	Sample ○
7	 	test_ta2	test_ta2	✓	 	Sample ○
8	 	Polymer details	Polymer Details	✓	 	Sample ○
9	 	tst	Test	✗	 	Test 🗑️
10	 	MOF Segment	MOF	✓	 	Sample ○
11	 	MOF Properties	MOF Properties	✓	 	Sample ○

# Chemotion and LabIMotion: one system – diverse options

The screenshot displays the Chemotion software interface. The top bar shows the application name "Chemotion(Testing)", a dropdown menu with "All", "IUPAC, InChI, SMILES, RIn", and various icons for navigation and editing. The main window is titled "NJ-W1 Workflow" and contains a flowchart of a sample preparation process. The steps are: Start, Basic sample (bsample), Cutting (cutting), Embedding (embedding), Mounting (mounting), Sputtering (sputtering), Step 1 (step\_one), Step 2 (step\_two), and Step 3 (step\_three). The flowchart shows a sequence of steps with dashed lines indicating dependencies. A right-hand panel titled "NJ-W1" provides configuration options for the workflow. It includes tabs for "Properties", "Analyses", and "Attachments". The "Workflow" tab is active, showing a list of workflow instances. Below this, there are configuration sections for "Basic sample (bsample)", "Cutting (cutting)", "Embedding (embedding)", and "Polishing Parameters (polishing)". Each section contains input fields for various parameters and a "Next" dropdown menu. The "Basic sample (bsample)" section has fields for "Material", "Dimensions", "details", and "Next" (set to "Cutting(cutting)"). The "Cutting (cutting)" section has fields for "input\_sample", "output\_sample", "instrument", "temperature", "humidity", "cutting\_wheel", "cutting\_velocity", "cooling\_agent", "feed", and "Next" (set to "Embedding(emb...)"). The "Embedding (embedding)" section has a "Next" dropdown set to "Polishing Parameters(polishing)". The "Polishing Parameters (polishing)" section has fields for "temperature" (set to "°C"), "humidity" (set to "%"), "instrument", and "Next" (set to "Select..."). At the bottom of the right panel are "Close" and "Create" buttons.

Chemotion(Testing) - All - IUPAC, InChI, SMILES, RIn [Icons] (Testing Instance) Nicole Jung

NJ-W1 Workflow

NJ-W1 New Workflow-Test

NJ-W1

Properties Analyses Attachments

Workflow

New Workflow-Test

Basic sample (bsample)

Sample Material Dimensions details Next

input\_sample output\_sample instrument temperature humidity

cutting\_wheel cutting\_velocity cooling\_agent feed Next

Embedding (embedding)

Next

Polishing Parameters (polishing)

temperature humidity instrument Next

Close Create



---

## Chemotion Repository: Publish data



# Chemotion repository

Chemotion-Repository My DB Publications Review Embargoed Publications Newsroom How-To

Nicole Jung



Published a week ago by Yichuan Wang



Repository for molecules,  
reactions and research  
data

## Visibility and Impact

- Publish your structures, attach your characterization data, and make them citable via DOI
- Automated registration at various scientific data providers
- Long-term archival - from scientists for scientists
- **NEW** Now, publish your Reactions



Samples

published

**2818**

426 under review  
778 under embargo



Reactions

published

**1098**

460 under review  
793 under embargo



Analyses

**9320** published

Top 3: 3797 NMR  
2314 IR  
1847 Mass

## Features



Storage

Hosted by an experienced data center, Chemotion repository can store your research data reliably and securely.



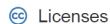
Metadata

Keep your research data findable and accessible and collect descriptions about the data. Based on DataCite Metadata Scheme.



DOI

Tie Digital Object Identifier (DOI) to your research data. This registers your data in DataCite and makes it identifiable, searchable and citable.



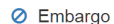
Licenses

Choose which license is suitable for your research data to allow others to re-use your data.



Data Quality

Release your research data and pass an internal review that ensures data quality.



Embargo

Put an embargo on your data. This allows you to delay the publication of your research data. You can release to make

# Chemotion repository

Chemotion-Repository

My DB

Publications

Review

Embargoed Publications

Newsroom

How-To

Chemotion Initiative

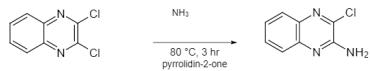


Q

Reactions

Samples

Scheme-only reactions



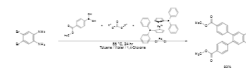
ID  
CRR-20478

Embargo

Author  
Laura Holzhauser

Analyses  
4

X-Vial



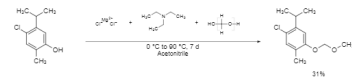
ID  
CRR-19983

Embargo

Author  
Jérôme Klein

Analyses  
8

X-Vial



ID  
CRR-19765

Embargo

Author  
Simone Gräßle

Analyses  
8

X-Vial  
-



ID  
CRR-19920

Embargo

Author  
Christoph Zippel

Analyses  
4

X-Vial



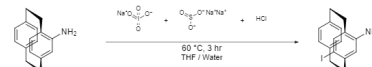
ID  
CRR-19926

Embargo

Author  
Christoph Zippel

Analyses  
4

X-Vial  
-



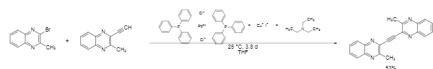
ID  
CRR-17667

Embargo

Author  
Christoph Zippel

Analyses  
4

X-Vial



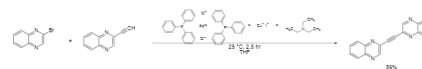
ID  
CRR-19873

Embargo

Author  
Victor Larignon

Analyses  
6

X-Vial



ID  
CRR-17106

Embargo

Author  
Victor Larignon

Analyses  
6

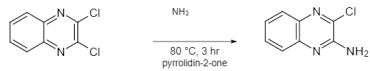
X-Vial



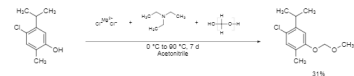
# Chemotion repository

Chemotion-Repository My DB Publications Review Embargoed Publications

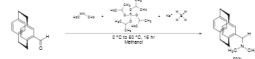
Q



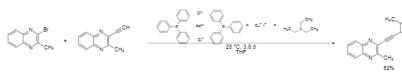
ID: CRR-20478 Embargo Author: Laura Holzhauser Anal: 4



ID: CRR-19765 Embargo Author: Simone Gräßle Anal: 8



ID: CRR-19926 Embargo Author: Christoph Zippel Anal: 4



ID: CRR-19873 Embargo Author: Victor Larignon Anal: 6



## 1H nuclear magnetic resonance spectroscopy (1H NMR)

DOI: 10.14272/BMIMNRP AEPIYDN-UHFFFAOYSA-N/CHMO0000593

<sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>, ppm)  $\delta$  = 2.40 (s, 3H, CH<sub>3</sub>), 7.23–7.28 (m, 2H, C<sub>6</sub>H<sub>4</sub>), 7.44–7.48 (m, 1H, C<sub>6</sub>H<sub>4</sub>), 7.67–7.69 (m, 1H, C<sub>6</sub>H<sub>4</sub>), 12.29 (s, 1H, NH).

### Datasets

R53A\_1H



## 13C nuclear magnetic resonance spectroscopy (13C NMR)

DOI: 10.14272/BMIMNRP AEPIYDN-UHFFFAOYSA-N/CHMO0000595

<sup>13</sup>C NMR (400 MHz, DMSO-*d*<sub>6</sub>, ppm)  $\delta$  = 20.5 (1C, CH<sub>3</sub>), 115.2 (1C, C<sub>6</sub>H<sub>4</sub>), 123.0 (1C, C<sub>6</sub>H<sub>4</sub>), 127.8 (1C, C<sub>6</sub>H<sub>4</sub>), 129.2 (1C, C<sub>6</sub>H<sub>4</sub>), 131.6 (1C, C<sub>6</sub>H<sub>4</sub>), 131.9 (1C, C<sub>6</sub>H<sub>4</sub>), 154.9 (1C, CNCH<sub>3</sub>/CONH), 159.1 (1C, CNCH<sub>3</sub>/CONH).

### Datasets

R53A\_13C



## mass spectrometry (MS)

DOI: 10.14272/BMIMNRP AEPIYDN-UHFFFAOYSA-N/CHMO0000470

EI (m/z, 70 eV, 100 °C): 161 (11) [M+H]<sup>+</sup>, 160 (94) [M]<sup>+</sup>, 132 (100), 131 (70).

### Datasets

R53A\_EI-MS



## Infrared absorption spectroscopy (IR)

DOI: 10.14272/BMIMNRP AEPIYDN-UHFFFAOYSA-N/CHMO0000630

IR (ATR,  $\tilde{\nu}$ ) = 418, 453, 469, 476, 561, 584, 599, 691, 725, 751, 779, 853, 888, 928, 945, 1007, 1122, 1156, 1188, 1208, 1276, 1285, 1344, 1380, 1422, 1432, 1485, 1502, 1567, 1601, 1659, 2707, 2769, 2836, 2881, 2958, 3003, 3098 cm<sup>-1</sup>.

### Datasets

R53\_IR

# Research data and publications

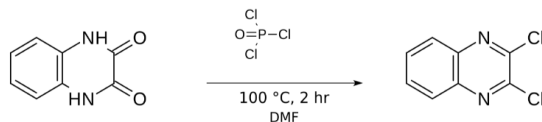
1

Supplemental  
Information Part 1

2

Research Data in  
Repository

## 2,3-Dichloroquinoxaline (10f)



Name {P1|**10f**}: 2,3-dichloroquinoxaline; Formula: C<sub>8</sub>H<sub>4</sub>Cl<sub>2</sub>N<sub>2</sub>; Molecular Mass: 199.0368; Exact Mass: 197.9752; Smiles: Clc1nc2ccccc2nc1Cl; InChIKey: SPSSDDOTEZKOOV-UHFFFAOYSA-N

Phosphoryl chloride (21.6 g, 13.1 mL, 141 mmol, 20.0 equiv) and 5 mL of DMF were added to the quinoxalinone (1.14 g, 7.0 mmol, 1.00 equiv) and heated to 100 °C for 2 h. The reaction was cooled to 21 °C, poured on ice and rested overnight. The organic phase was separated and the aqueous phase was extracted 3x with ethyl acetate; the combined organic layers were combined were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and the solvent was removed under reduced pressure. The remaining solid was purified by column chromatography (cHex/ethyl acetate 10:1). 1.32 g (6.65 mmol, 95%) of a colorless solid were obtained.

*R*<sub>f</sub> = 0.59 (cyclohexane/ethyl acetate 4:1). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>, ppm) δ = 8.02–8.07 (m, 2H, CH<sub>ar</sub>), 7.80–7.84 (m, 2H, CH<sub>ar</sub>); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>, ppm) δ = 145.4 (2C, C2N2Cl2), 140.6 (2C, C<sub>ar</sub>), 131.3 (2C, C<sub>ar</sub>), 128.3 (2C, C<sub>ar</sub>); EI (m/z, 70 eV, 20 °C): 200/198 (66/100) [M]<sup>+</sup>, 165 (21), 163 (65), 102 (46); HRMS (EI, C<sub>8</sub>H<sub>4</sub>N<sub>2</sub><sup>35</sup>Cl<sub>2</sub>): calcd 197.9752 found 197.9752; IR (ATR  $\tilde{\nu}$ ) = 3104 3063 3041 3002 2944 1955

<https://doi.org/10.14272/reaction/SA-FUHFF-UHFFFADPSC-SPSSDDOTEZ-UHFFFADPSC-NUHFF-NUHFF-NUHFF-ZZZ.1>

Additional information on the chemical synthesis is available via Chemotion repository:  
<https://doi.org/10.14272/reaction/SA-FUHFF-UHFFFADPSC-SPSSDDOTEZ-UHFFFADPSC-NUHFF-NUHFF-NUHFF-ZZZ.1>



**Thank you for listening**

# Thank you for your attention

**DFG**

**HELMHOLTZ**

RESEARCH FOR GRAND CHALLENGES



Baden-Württemberg

MINISTERIUM FÜR WISSENSCHAFT,  
FORSCHUNG UND KUNST



VIRTUAL  
MATERIALS  
DESIGN

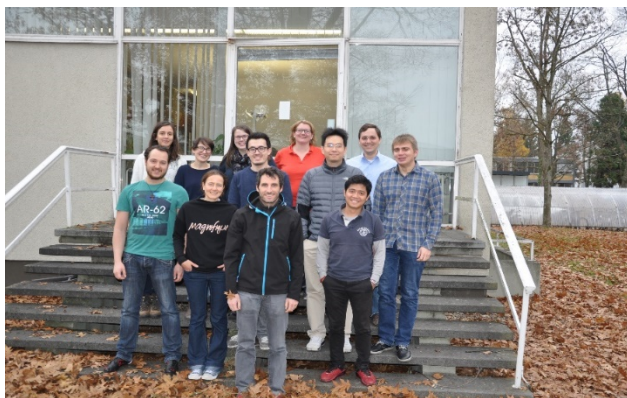


Tools for NMR spectroscopists



BEILSTEIN INSTITUT

## KIT-library & KIT-SCC & KIT - IBCS



Icons made by *Freepik*, *Madebyoliver*, *Anatoli*, *Dinosoft*, *Vectors market*, *Zlatko Najdenovski*, and *Gregor Cresnar* from [www.flaticon.com](http://www.flaticon.com). Gif-Source: <https://gifer.com/en/PsKr>

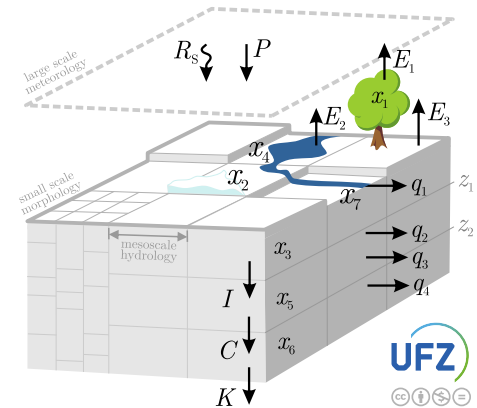
## 2. Helmholtz Open Science Forum

### The mesoscale Hydrological Model - mHM

mHM Developers team

07.04.2022

### The mesoscale Hydrologic Model **mHM**



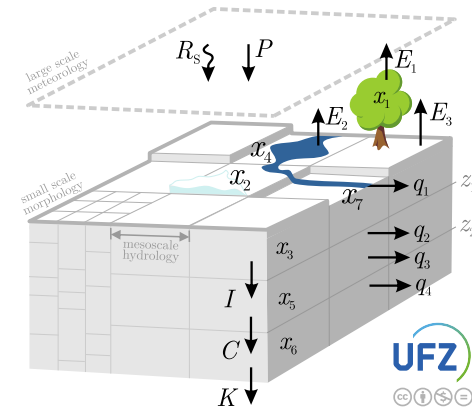
## 2. Helmholtz Open Science Forum – mHM

7. April 2022

### Inhalt

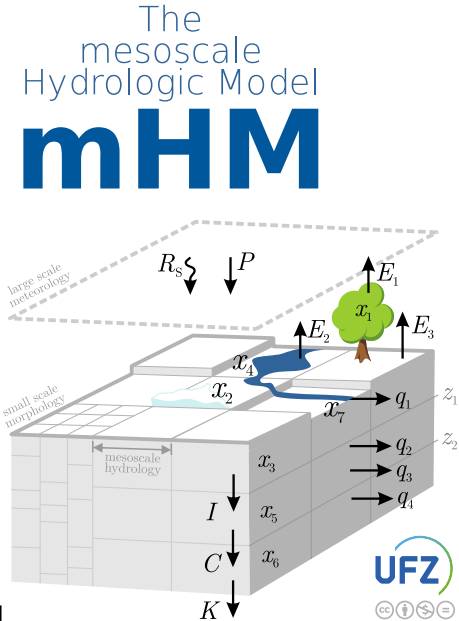
- Was ist mHM?
- Genutzte Infrastruktur
  - Entwicklung
  - Öffentlichkeit
  - Kommunikation
  - Installation
- Ausblick

### The mesoscale Hydrologic Model **mHM**



### Was ist mHM?

- räumlich explizit verteiltes hydrologisches Modell
- Gitterzellen als primäre hydrologische Einheit
- Berücksichtigte Prozesse:
  - Interzeption durch die Baumkronen
  - Schneeakkumulation und -schmelze
  - Bodenfeuchtedynamik
  - Infiltration und Oberflächenabfluss
  - Evapotranspiration
  - unterirdische Speicherung und Abflussbildung
  - tiefe Perkolations und Basisabfluss
  - Abflusdämpfung und Hochwasserablaufberechnung

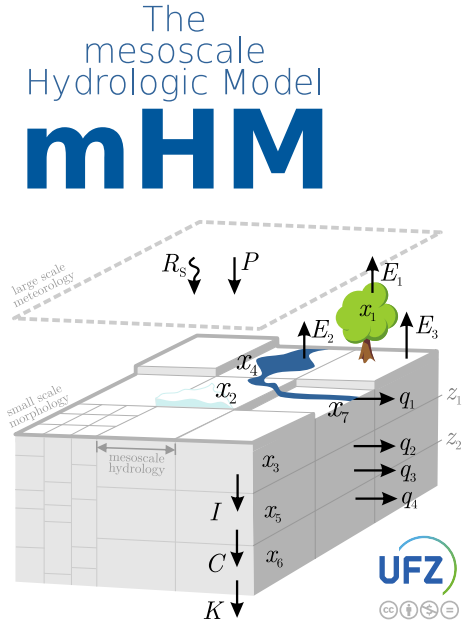


## 2. Helmholtz Open Science Forum – mHM

7. April 2022

### Was ist mHM?

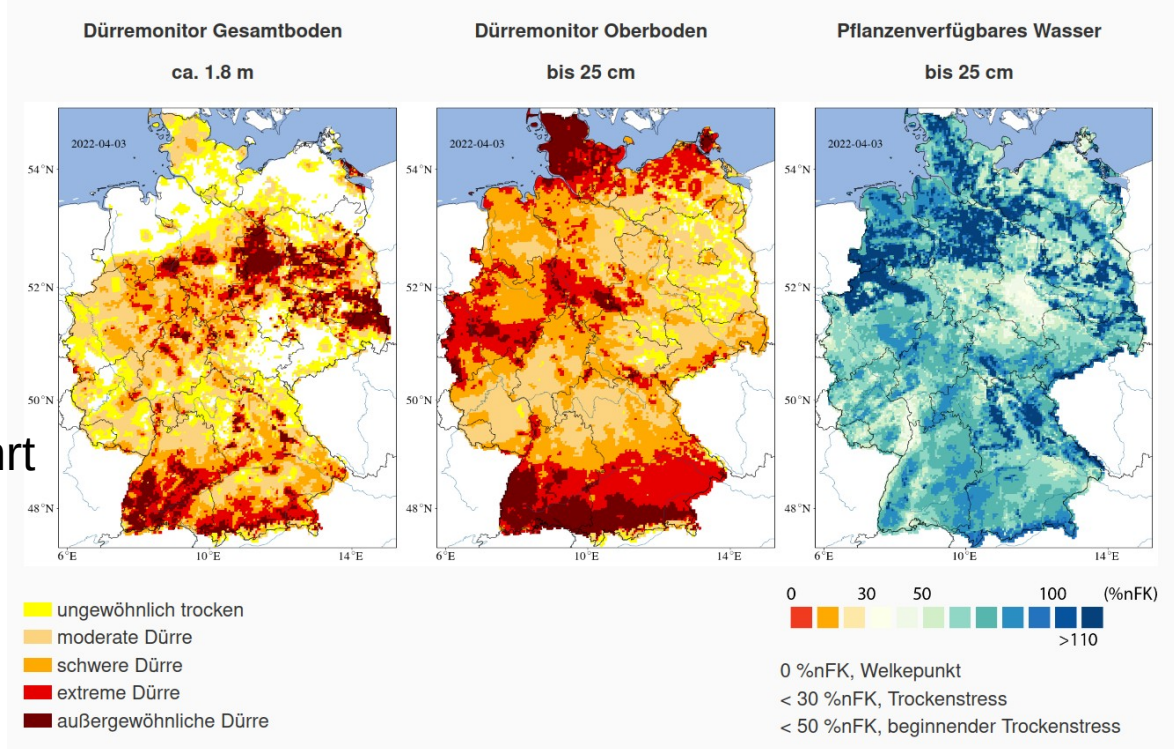
- Modern Fortran + Cmake
- Modulare Entwicklung: Cmake Package Manager
- Entwicklung seit über 18 Jahren am UFZ
- Versionskontrolliert (git): 1600+ Commits
- 25+ Entwickler\*innen
- Wissenschaftlich fundiert:
  - Samaniego et al. (2010)
  - Kumar et al. (2013)
  - Thober et al. (2019)
- bewährtes Modell in der Wissenschaftsgemeinschaft
- Wissenschaftliche Basis des deutschen Dürremonitors





### Was ist mHM? Dürremonitor

- Tägliche Berechnung der Bodenfeuchte mit mHM
- Angetrieben durch DWD Messdaten
- operationelle Modellkette wird automatisiert am UFZ Rechencluster ausgeführt

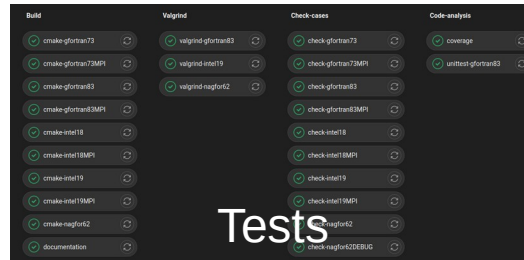


## 2. Helmholtz Open Science Forum – mHM

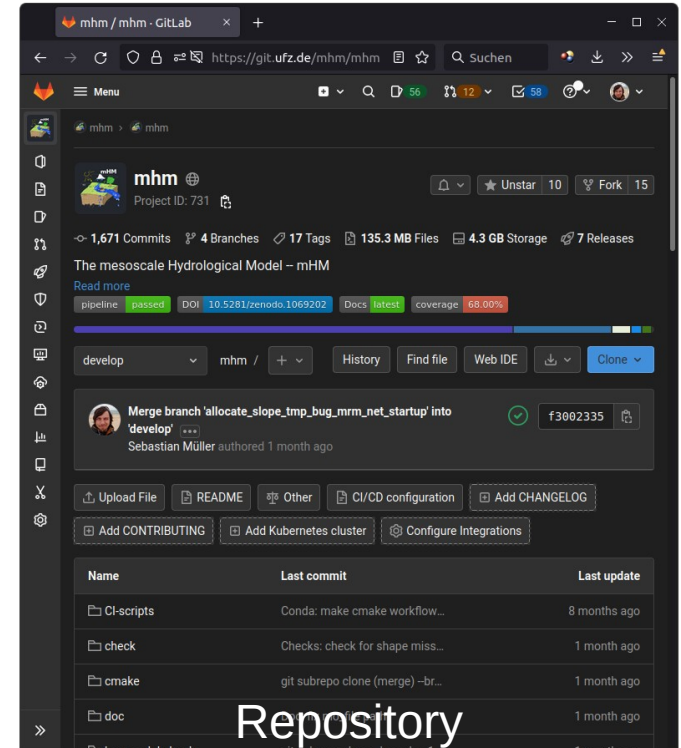
7. April 2022

### Genutzte Infrastruktur: Agile Entwicklung

- git.ufz.de/mhm/mhm – GitLab repository
  - Issues – interne Diskussion
  - Merge Requests: Entwicklung + Reviews
  - GitLab-Runner: Test nach jeder Änderung
  - GitLab-Pages: Code-Dokumentation



Tests



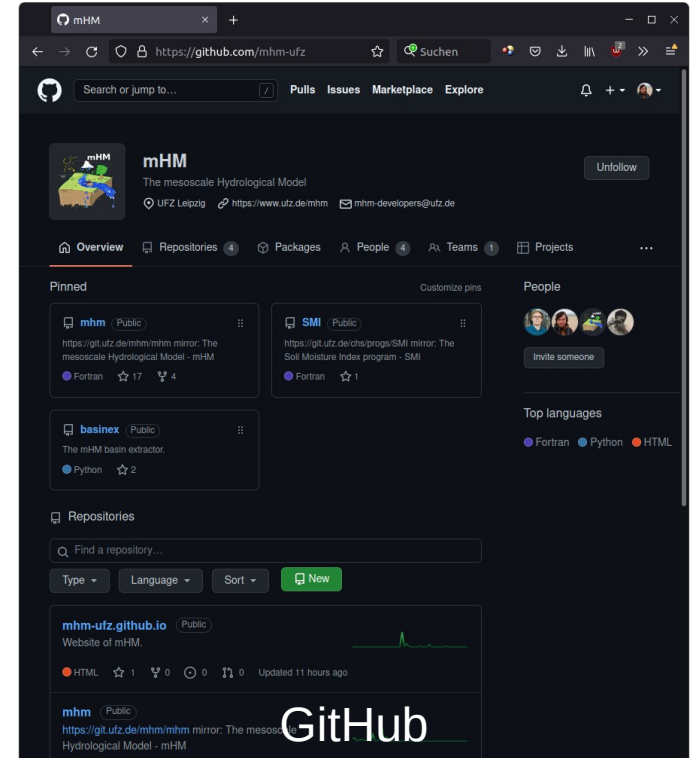
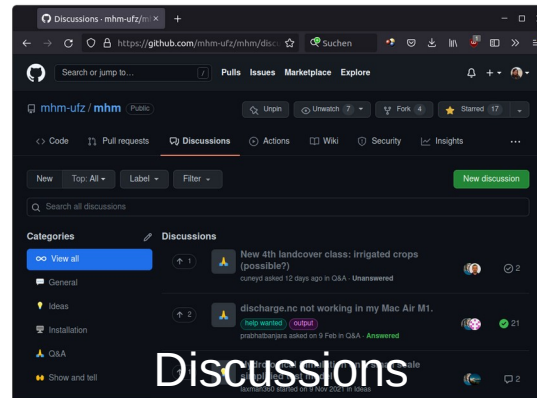
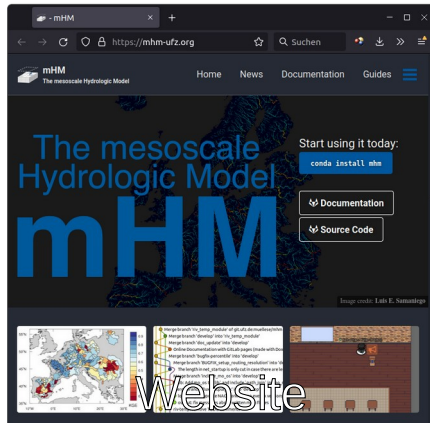
Repository

## 2. Helmholtz Open Science Forum – mHM

7. April 2022

### Genutzte Infrastruktur: Öffentlichkeit

- github.com/mhm-ufz/mhm – GitHub mirror
  - Sichtbarkeit
  - „öffentliche“ Diskussion
  - Ort für assoziierte Software (basinex)
  - Hosten der Website (mhm-ufz.org)

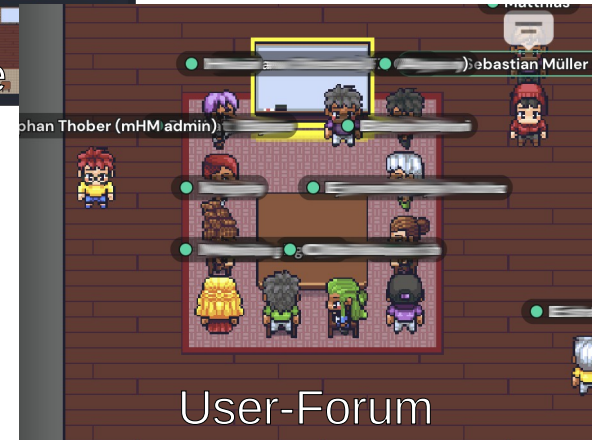
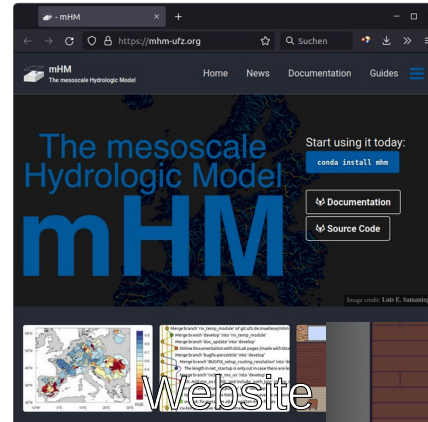


## 2. Helmholtz Open Science Forum – mHM

7. April 2022

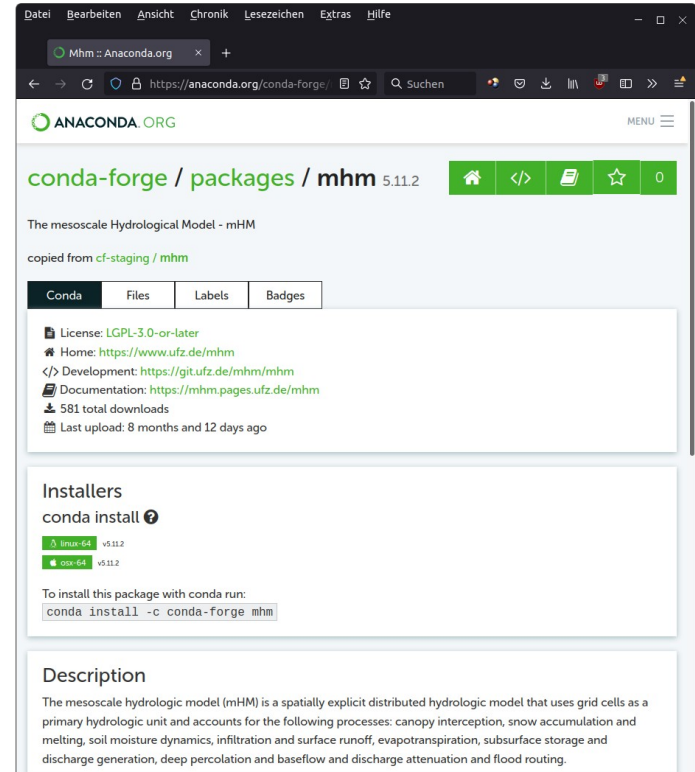
### Genutzte Infrastruktur: Kommunikation

- **Website:** mhm-ufz.org
  - Jekyll Website auf GitHub
  - Alles an einem Ort
  - Dokumentation
  - Diskussion
  - Anleitungen (Installation, Kompilation)
  - News mit Kommentarfunktion
- **User-Forum:** Gather-Town
  - Alle 2 Monate
  - Offenes Forum zur Diskussion
  - Hilfe und Kurze Wege



## Genutzte Infrastruktur: Installation

- **conda**: etablierter Paketmanager
  - Paketerstellung über conda-forge (GitHub) (Community projekt)
  - `conda install mhm`
  - Sehr niederschwellig für Neulinge
- **Anleitungen**: Website
  - Für alle Systeme
  - Für mehrere Compiler
- **mhm\_pybind**: experimentelle Python bindings
  - Automatisiertes kompilieren der `mhm_lib`
  - Scikit-build: cmake Projekt mit pip installieren



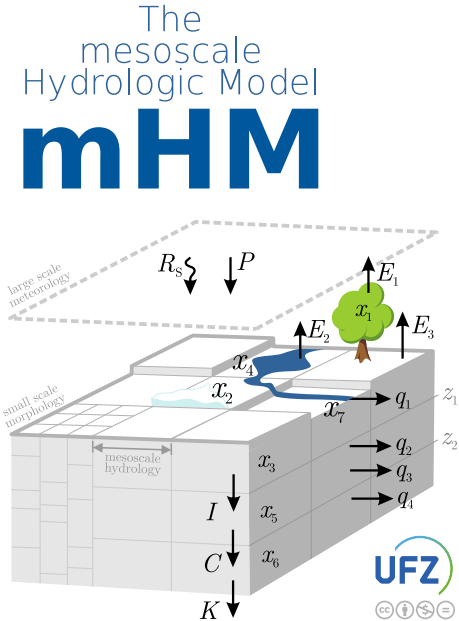
The screenshot shows a web browser window displaying the Anaconda.org website. The page is for the conda-forge package 'mhm' version 5.11.2. The browser's address bar shows the URL 'https://anaconda.org/conda-forge/mhm'. The page content includes the Anaconda logo, navigation links, and a description of the package: 'The mesoscale Hydrological Model - mHM'. It also lists the license (LGPL-3.0-or-later), home page, development repository, documentation, and download statistics. Under the 'Installers' section, there are two options: 'conda install' and 'conda install -c conda-forge mhm'. The 'Description' section provides a detailed overview of the model's capabilities.

## 2. Helmholtz Open Science Forum – mHM

7. April 2022

### Ausblick

- mHM v6
- mhmpy: mHM Läufe bequem als Python Skripte aufsetzen
- Kopplung mit anderen Modellen über Python





## Helmholtz Digital Services for Science — Collaboration made easy.

April 2022

Tobias Huste



**Free of charge!**  
(in Helmholtz)

## I. Helmholtz Cloud + Backbone

- Provide high performance **collaboration and community services**
- **Connect all centres** and their world-wide collaboration partners!
- **Secure, simple access** and **easy-to-use**

## II. Establish **best-practices** for **Research Software Engineering**:

- High level of knowledge, quality, visibility and sustainability
- State-of-the-art SW development infrastructures

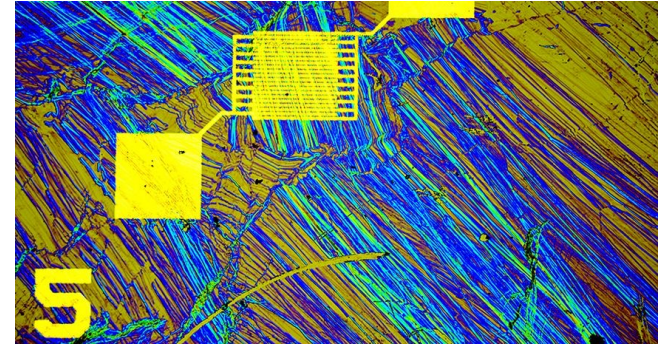
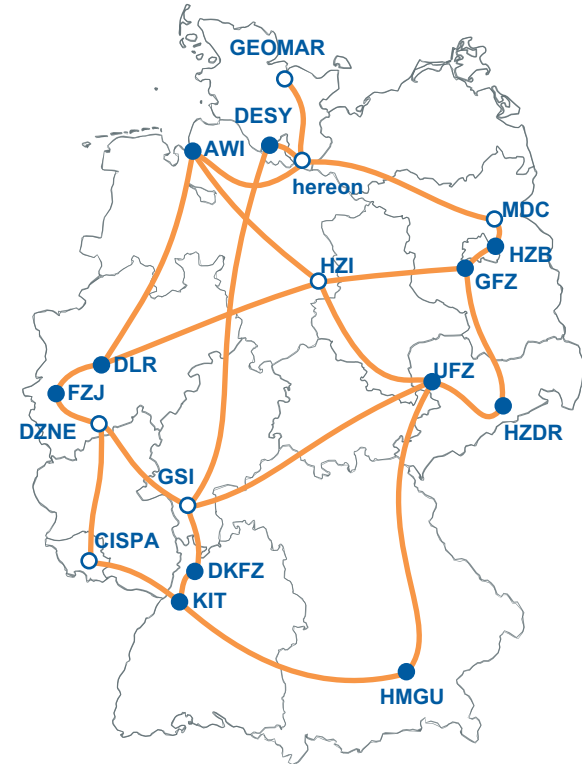


Bild 1: Strukturen aus organischen Halbleiter-Polymeren/ Himani Arora/HZDR  
Bild 2: Mosaic Expedition/Stefan Hendricks/AWI



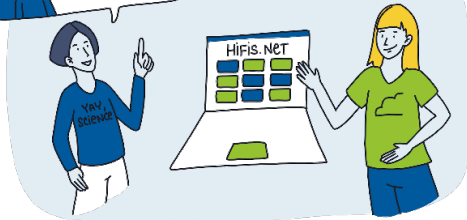
# Who is HIFIS?

- **Team: 11 (out of 18) Helmholtz centres**
- Main Coordination: DESY
- Three Clusters:
  - **Backbone Cluster** (head: DESY)  
Core Services for Authentication, Authorisation, Large Data Transfer, etc.
  - **Cloud Cluster** (head: HZB)  
Helmholtz Cloud Platform
  - **Software Cluster** (head: HZDR)  
Platform, Training, Support for high quality, sustainable Research Software Engineering

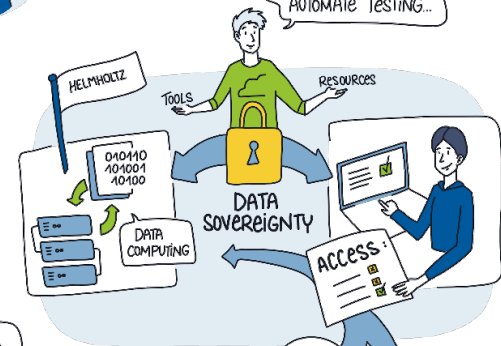


WE ARE STARTING A  
**Scientific Project!**  
CAN YOU  
HELP US?

SURE! HERE ARE TOOLS YOU CAN USE,  
FOR EXAMPLE GIT LAB, SYNC & SHARE, ...



THAT'S HOW YOU  
AUTOMATE TESTING..



**HIFIS** | HELMHOLTZ  
FEDERATED  
IT SERVICES

WE HELP YOU TO **TRANSFER**  
**YOUR RESULTS:**  
DOI, LICENSING & STORAGE!



**SUSTAINABILITY**  
10+ YEARS

WE SUPPORT YOU WITH  
COURSES, CONSULTING &  
COLLABORATION TOOLS!



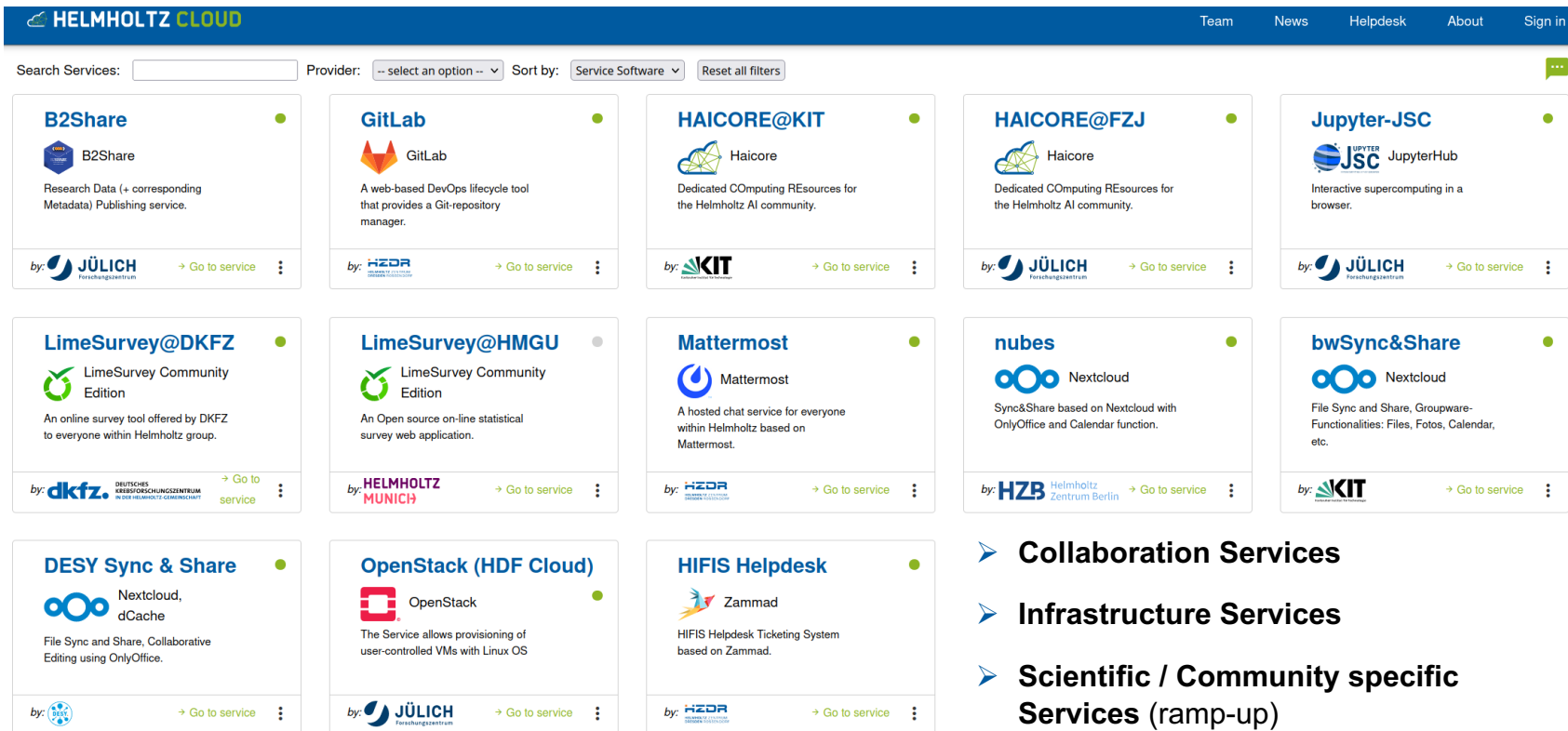
ILLUSTRATION: LORNA SCHÜTTE



➤ [Link to this poster](#)

# I. Cloud Services + Portal

## ➤ [cloud.helmholtz.de](https://cloud.helmholtz.de)



The screenshot displays the HELMHOLTZ CLOUD portal interface. At the top, there is a navigation bar with links for Team, News, Helpdesk, About, and Sign in. Below the navigation bar, a search bar is present with the text "Search Services:" followed by an input field. To the right of the search bar, there are filters for "Provider:" (set to "-- select an option --"), "Sort by:" (set to "Service Software"), and a "Reset all filters" button. The main content area is a grid of service cards, each representing a different cloud service. Each card includes the service name, a logo, a brief description, and the provider information. The services shown are:

- B2Share**: Research Data (+ corresponding Metadata) Publishing service. Provider: JÜLICH Forschungszentrum.
- GitLab**: A web-based DevOps lifecycle tool that provides a Git-repository manager. Provider: HZDR Helmholtz-Zentrum für Materialforschung und -technik.
- HAICORE@KIT**: Dedicated Computing REsources for the Helmholtz AI community. Provider: KIT Karlsruhe Institute of Technology.
- HAICORE@FZJ**: Dedicated Computing REsources for the Helmholtz AI community. Provider: JÜLICH Forschungszentrum.
- Jupyter-JSC**: Interactive supercomputing in a browser. Provider: JÜLICH Forschungszentrum.
- LimeSurvey@DKFZ**: LimeSurvey Community Edition. An online survey tool offered by DKFZ to everyone within Helmholtz group. Provider: dkfz DEUTSCHES KREBSFORSCHUNGSZENTRUM HEIDELBERGHELMHOLTZ-GEMEINSCHAFT.
- LimeSurvey@HMGU**: LimeSurvey Community Edition. An Open source on-line statistical survey web application. Provider: HELMHOLTZ MUNICH.
- Mattermost**: A hosted chat service for everyone within Helmholtz based on Mattermost. Provider: HZDR Helmholtz-Zentrum für Materialforschung und -technik.
- nubes**: Sync&Share based on Nextcloud with OnlyOffice and Calendar function. Provider: HZB Helmholtz Zentrum Berlin.
- bwSync&Share**: File Sync and Share, Groupware-Functionalities: Files, Fotos, Calendar, etc. Provider: KIT Karlsruhe Institute of Technology.
- DESY Sync & Share**: File Sync and Share, Collaborative Editing using OnlyOffice. Provider: DESY Deutsches Elektronen-Synchrotron.
- OpenStack (HDF Cloud)**: The Service allows provisioning of user-controlled VMs with Linux OS. Provider: JÜLICH Forschungszentrum.
- HIFIS Helpdesk**: HIFIS Helpdesk Ticketing System based on Zammad. Provider: HZDR Helmholtz-Zentrum für Materialforschung und -technik.

- Collaboration Services
- Infrastructure Services
- Scientific / Community specific Services (ramp-up)

[No account? Sign up.](#)

## HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

Login to Helmholtz AAI user's account

Search

	29 Mayis University
	A'SHARQIYAH UNIVERSITY
	A*STAR - Agency for Science, Technology and Research
	A. T. Still University
	AAF Virtual Home
	aai.lab.maaen.sa
	AAI@EduHr Single Sign-On Service
	Aalborg University
	Aalto University
	Aarhus School of Architecture

[Contact](#) [Documentation](#)

- <https://login.helmholtz.de>
- <https://aai.helmholtz.de/howto>
- <https://aai.helmholtz.de/concept>

### Best-practices for sustainable Research Software Engineering on multiple levels:

#### Education & Training

- **Courses, material and workshops** for getting you started or boosting your software engineering practice.

#### Community

- Build and foster communities to support the necessary **cultural change** when dealing with research software.

#### Consulting

- **Contact points** for researchers for questions and problems in the context of RSE.

#### Technology

- Provide a sustainable, well integrated and easy to use **technology infrastructure** for research software development.

### Education & Training

#### Basic

- First Steps in Python-Programming
- Project Management with GitLab
- Version Control using Git

#### Intermediate

- GitLab for Software Development in Teams
- Let us Make Your Script Ready for Publication
- OOP Programming with Python (available in 2022)
- Continuous Integration (CI) using GitLab CI (available in 2022)

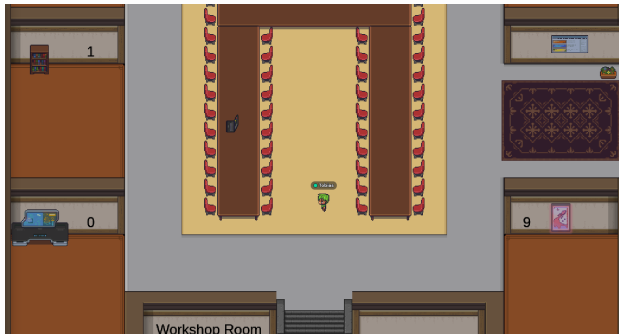
#### Advanced

- Using Containers in Science
- Test Automation with Python

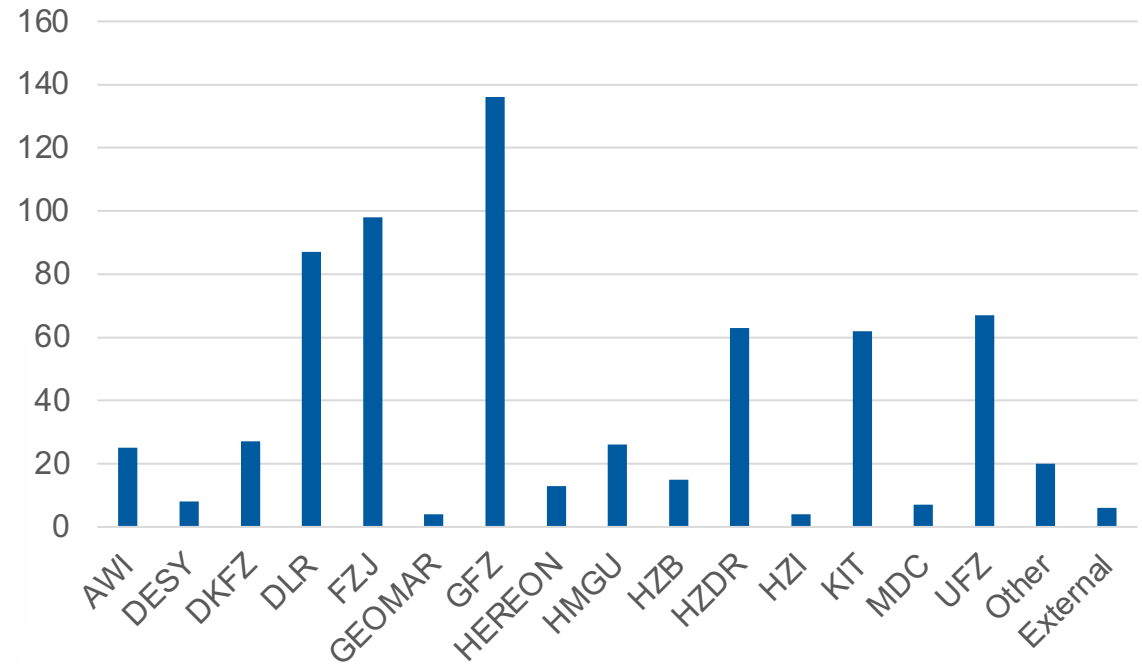
## II. Software Services

### Education & Training

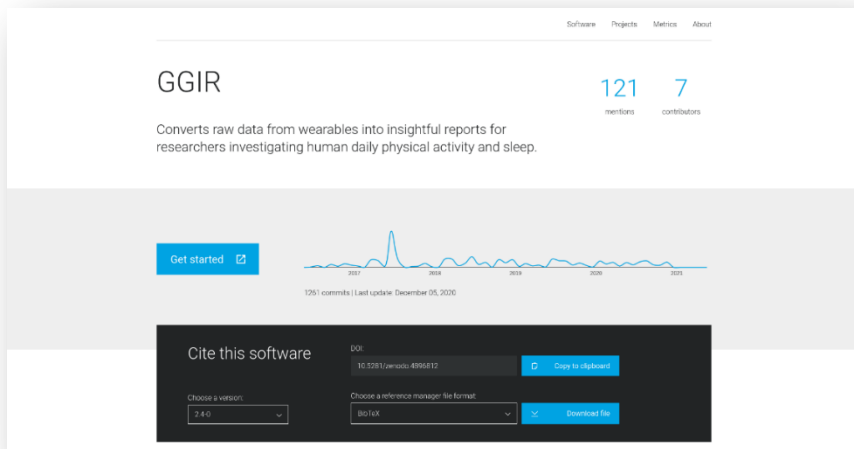
- **36** online workshops in 2021
- About **668** attendees from **16** Helmholtz centers



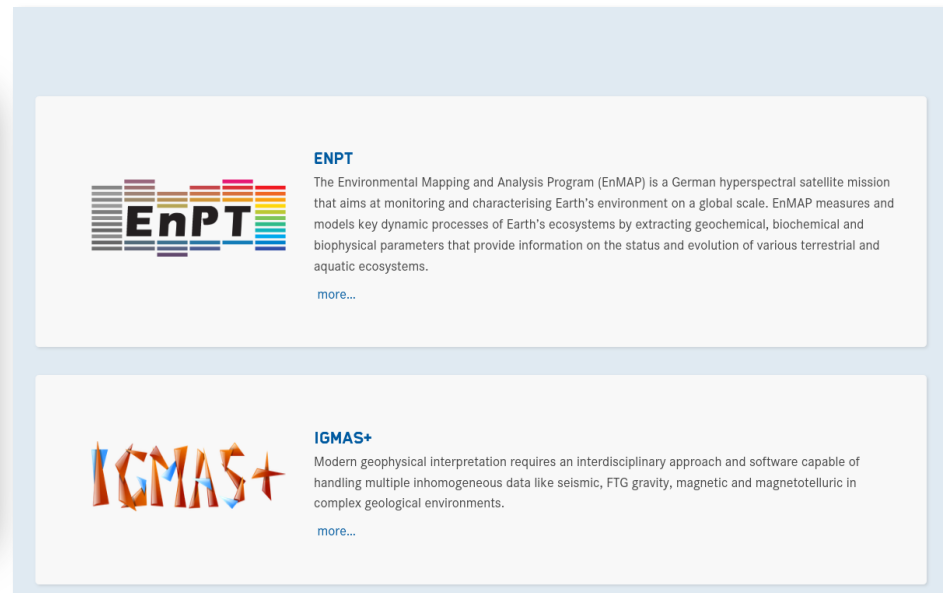
Number of Attendees per Center



## Community



- Started development of a Research Software Directory for Helmholtz
- To be built on top of and in cooperation with the Netherlands eScience center [solution](#)



- Present & Promote top success stories of Research Software Engineering in Helmholtz
- approx. **90 Proposals** so far, to be presented in: <https://hifis.net/spotlights>



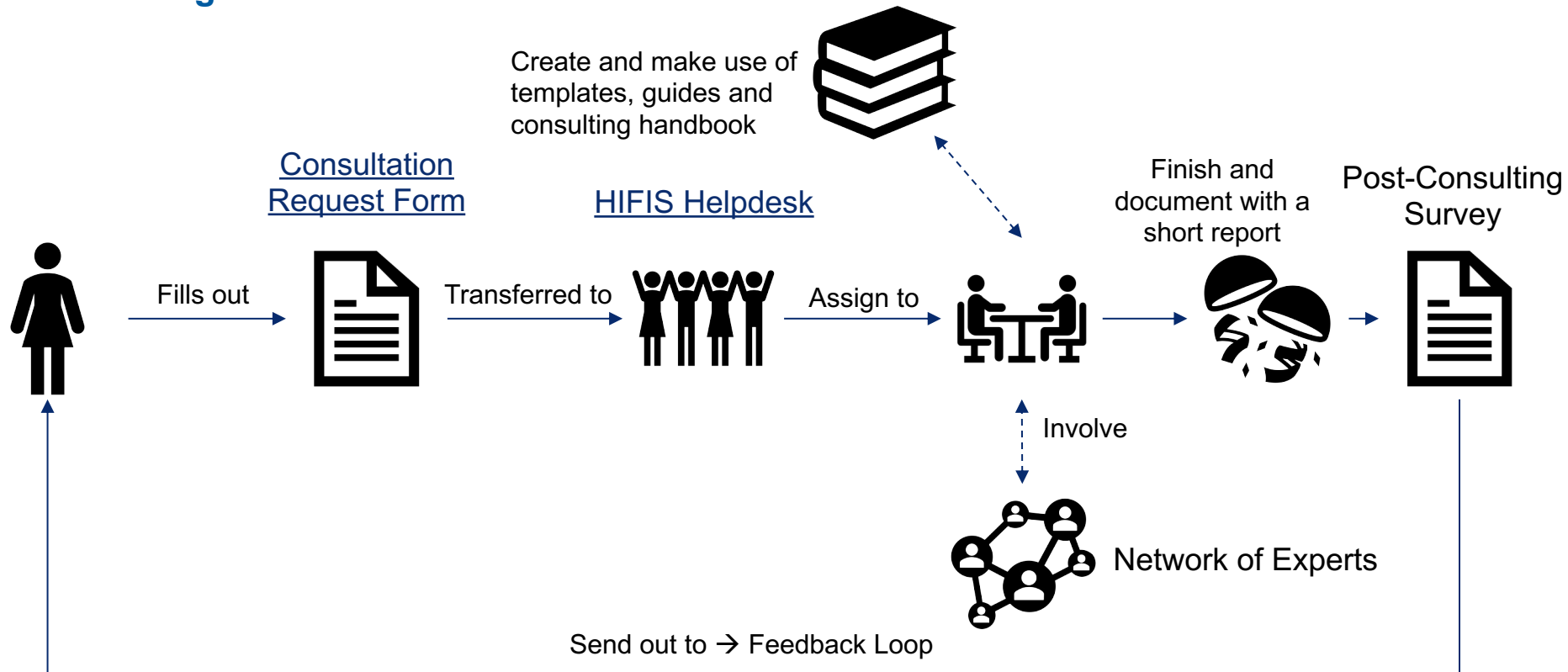
### Consulting

- Free-of-charge software consulting for research groups within the Helmholtz umbrella
- Possible topics include, but are not limited to Licensing and Open Source, setting up new projects, code migrations etc.
- **Initial References:**
  - Making a project HPC ready
  - Setting up and introducing Continuous Integration
  - Frequent requests about licensing and Open Source



Katerina Limpitsouni via  
undraw.co

## Consulting



## II. Software Services

### Consulting

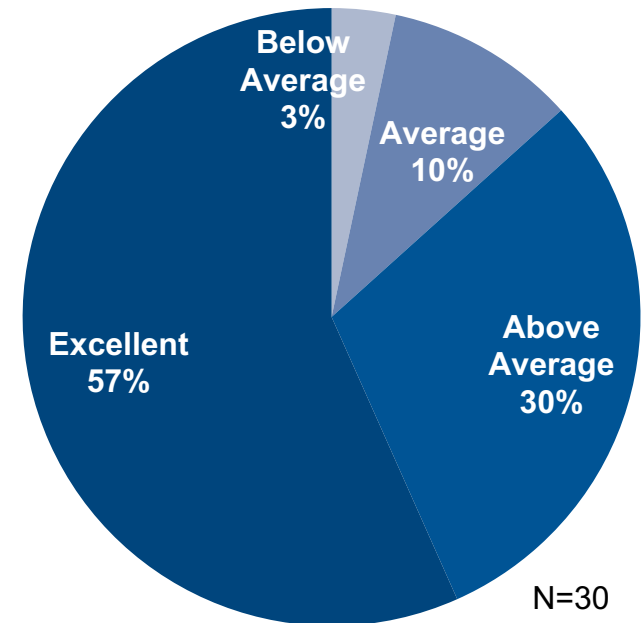
#### Net Promoter Score



*This was a great experience. Having worked at a Helmholtz institution for many years, it is a bit difficult for senior staff to understand that this is really a free and open service meant to support research groups in different fields. This is quite uncommon in “old Helmholtz” and it may take a while before HIFIS becomes widely known and accepted...*

”

#### Impact of the consultation on your project or work



### Technology

*Provide a sustainable, well integrated and easy to use technology infrastructure for research software development.*

Transparency

Accessibility

Reliability

Reusability

- Everything we implement is shared as Open Source software
  - <https://gitlab.com/hifis/ansible>
- Infrastructure as code → Remove manual components, Transparently visible for Helmholtz employees
  - <https://gitlab.hzdr.de/hifis-software-deployment/>

# II. Software Services

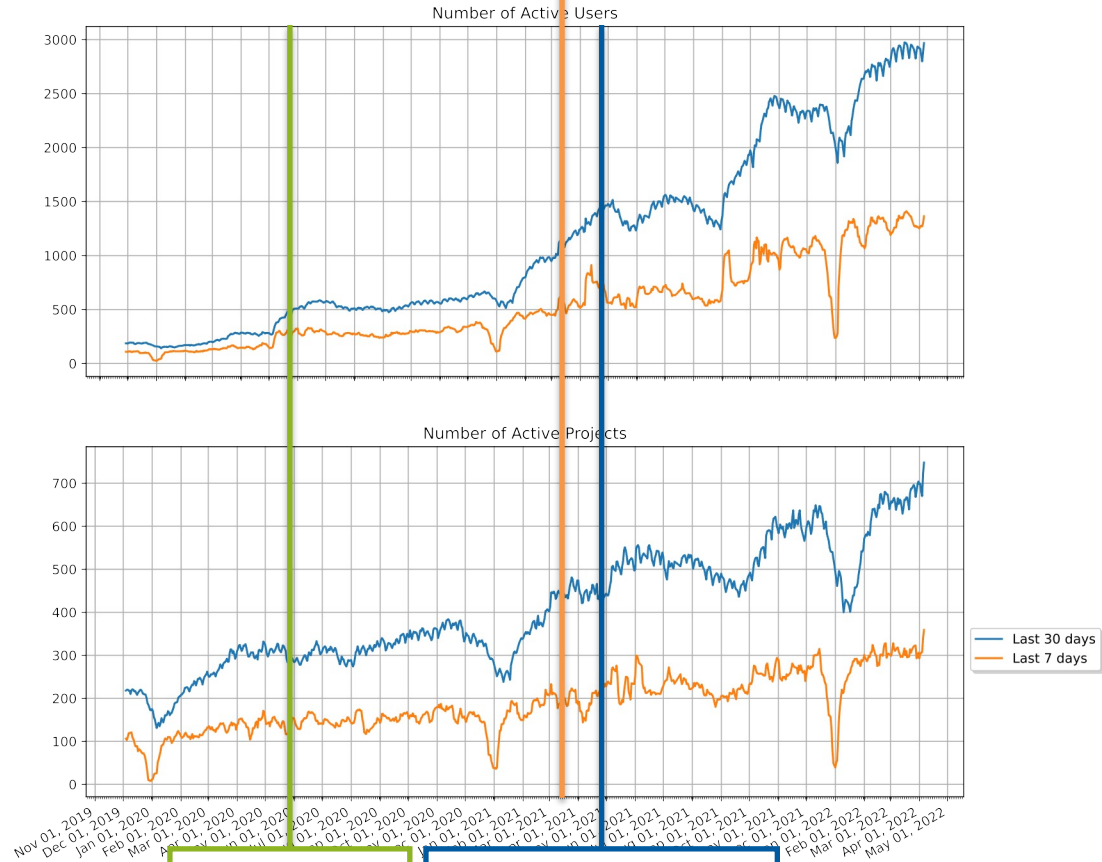
Relaunched as Helmholtz GitLab and announced on hifis.net



Technology



## Helmholtz Codebase - Usage Statistics



Connected the Helmholtz AAI.

Helmholtz Cloud portal made available.

Helmholtz Federated IT Services

## I. Helmholtz Digital Services for Science - Collaboration made easy.

- Try it! Software and Cloud Services readily available for Helmholtz + Partners.

## II. Cross-center Collaboration

## III. Services around RSE highly requested

## IV. Consult us:

- <https://hifis.net>
- [support@hifis.net](mailto:support@hifis.net)
- [Resonator Podcast](#)



[Podcast](#)



[Poster](#)

# QUESTIONS?



Katerina Limpitsouni via  
undraw.co



2. Helmholtz-Open Science Forum: Forschungssoftware

## **HERMES: Automated software publication with rich metadata**

Tobias Schlauch

*German Aerospace Center (DLR) & Humboldt-Universität zu Berlin*

with Oliver Bertuch (FZ Jülich), Oliver Knodel (HZDR), Guido Juckeland (HZDR), Stephan Druskat (DLR), Michael Meinel (DLR)

License **CC-BY-4.0 International**

# Software publication

## The good news

---

- **Software is an important research output**
- **Publishing research software supports**
  - **Sustainability**
  - **Reproducibility**
  - **Academic credit**



# Software publication

## The less good news

### To be FAIR:

1. It's a lot of work
2. It's a lot of manual work

- Identifiers (DOIs)
- Rich metadata
- Accessibility
- Machine-readable metadata
- Documenting dependencies
- Licenses
- Provenance
- ...
- Versions!



Montag 13:30 - Raum Schniecke - Workshop: Let's talk FAIR for research software

Hosts: Martin Stoffers, Tobias Schlauch, Alexander Struck, René Caspart

The session will cover practices in and handling of research software development and how they relate to the FAIR principles [0]. We will introduce current challenges we have faced and/or identified in our daily work and research.

Using the format World Cafe, we like to discuss such challenges with you to gather your perspectives and ideas. Existing material, e.g. [1] software publication is discussed and (ideally) improved. The Findability aspect is further investigated when we discuss how/when/where publication may be touched upon.

The situation in HPC environments may differ from common research labs and we will look at user-oriented and interoperable. The audience is free to focus on any aspect or cover all.

[0] <https://www.rtd-alliance.org/groups/fair-research-software-fair4rs-wg>

[1] <https://gitlab.com/hifs/hifs-workshops>

2020-11-20 13:50  
DOI: 10.5281/zenodo.50628

### Towards FAIR principles for research software

Anna-Lena Lamprecht<sup>1\*</sup>, Leyla Garcia<sup>2\*</sup>, Mateusz Kuzak<sup>3,4</sup>, Carlos Martinez<sup>5</sup>, Ricardo Arella<sup>6</sup>, Eva Martin Del Pico<sup>7</sup>, Victoria Dominguez Del Angel<sup>8</sup>, Stephanie van de Sandt<sup>9</sup>, Jan Jost<sup>10</sup>, Paula Andrea Martinez<sup>11</sup>, Peter McQuillen<sup>12</sup>, Alfonso Valencia<sup>13,14</sup>, Jennifer Harrow<sup>15</sup>, Fotis Psomopoulos<sup>16</sup>, Josep L.L. Gelji<sup>17</sup>, Neil Chiu Hong<sup>18</sup>, Carole Goblet<sup>19</sup> and Salvador Capella-Gutierrez<sup>1,24</sup>

<sup>1</sup> Utrecht University, The Netherlands  
E-mail: a.l.lamprecht@uu.nl, ORCID: <https://orcid.org/0000-0001-1955-5606>

<sup>2</sup> ZIMED Information Centre for Life Sciences, Germany  
E-mail: [lgarcia@zimed.de](mailto:lgarcia@zimed.de), ORCID: <https://orcid.org/0000-0001-3986-0510>

<sup>3</sup> Netherlands eScience Center, The Netherlands  
E-mail: [m.kuzak@esciencecenter.nl](mailto:m.kuzak@esciencecenter.nl), ORCID: <https://orcid.org/0000-0001-0087-6021>

<sup>4</sup> Netherlands eScience Center, The Netherlands  
E-mail: [c.martinez@esciencecenter.nl](mailto:c.martinez@esciencecenter.nl), ORCID: <https://orcid.org/0000-0001-5565-2577>

<sup>5</sup> EMBL-EBI, UK  
E-mail: [carlos@ebi.ac.uk](mailto:carlos@ebi.ac.uk), ORCID: <https://orcid.org/0000-0002-8253-7375>

<sup>6</sup> Barcelona Supercomputing Center (BSC), Spain  
E-mail: [eva.martin@bsc.es](mailto:eva.martin@bsc.es), ORCID: <https://orcid.org/0000-0001-1824-2849>

<sup>7</sup> Institut Français de Bioinformatique (IFB), France  
E-mail: [victoria.dominguez@france-bioinformatique.fr](mailto:victoria.dominguez@france-bioinformatique.fr), ORCID: <https://orcid.org/0000-0002-5314-6651>

<sup>8</sup> ICBN, Switzerland  
E-mail: [stephanie.van.de.sandt@cern.ch](mailto:stephanie.van.de.sandt@cern.ch), ORCID: <https://orcid.org/0000-0002-9576-1974>

<sup>9</sup> National Life Science Supercomputing Centre, Technical University of Denmark, Denmark  
E-mail: [jost@life-science.dtu.dk](mailto:jost@life-science.dtu.dk), ORCID: <https://orcid.org/0000-0001-6666-1520>

<sup>10</sup> National Imaging Facility, Australia  
E-mail: [peter.mcmquillen@nif.scrib.nsw.gov.au](mailto:peter.mcmquillen@nif.scrib.nsw.gov.au), ORCID: <https://orcid.org/0000-0002-8990-1985>

<sup>11</sup> Oxford e-Research Centre, UK  
E-mail: [peter.mquillen@nif.scrib.nsw.gov.au](mailto:peter.mquillen@nif.scrib.nsw.gov.au), ORCID: <https://orcid.org/0000-0001-2087-1982>

<sup>12</sup> Barcelona Supercomputing Center (BSC), Spain  
E-mail: [jost@life-science.dtu.dk](mailto:jost@life-science.dtu.dk), ORCID: <https://orcid.org/0000-0002-8937-6789>

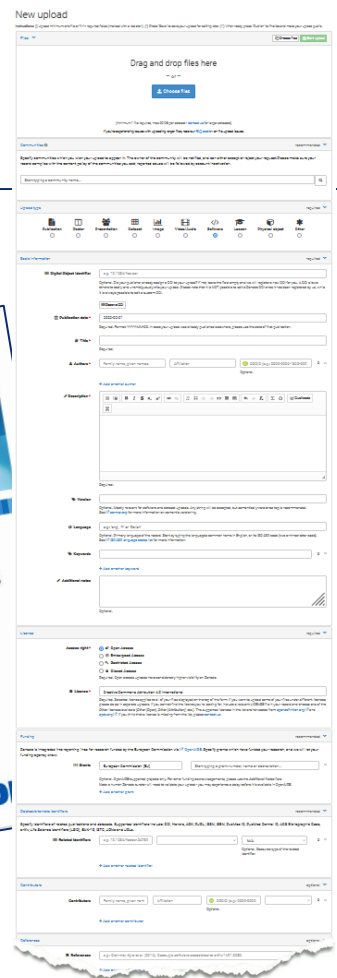
<sup>13</sup> ELIXIR Hub, UK  
E-mail: [jharrow@ebi.ac.uk](mailto:jharrow@ebi.ac.uk), ORCID: <https://orcid.org/0000-0001-0338-3070>

<sup>14</sup> Institute of Applied Bioinformatics, CERIT, Greece

3451-3484 © 2020 - ICBN Press and the authors.  
This article is published online with Open Access and distributed under the terms of the Creative Commons Attribution License (CC BY 4.0).

Leitlinien zur Sicherung guter wissenschaftlicher Praxis

Kodex



# Software publication

## What does it look like?

- **State-of-the-art satisfaction of FAIR principles:** metadata (+ software) in a publication repository
- **We have:** metadata, publication repositories

```
model = getattr(spectra, spectrum_dict["type"])

if norm.unit in (u.Unit("erg"), u.Unit("erg cm-3")) and norm_type != "integral":
    raise NameError(
        "Normalisation different than 'in"
    )

# check the units of the normalisation
# cm-3 is the only one allowing more than
if norm.unit == u.Unit("cm-3"):
    if norm_type == "differential":
        final_model = model(norm, **spect
    elif norm_type == "gamma-1":
        final_model = model.from_norm_at_
            norm, **spectrum_dict["parame
    elif norm_type == "integral":
        final_model = model.from_normalis
            norm, **spectrum_dict["parame
    )

{
  "@context": "https://doi.org/10.5063/schema/codemeta-2.0",
  "@type": "SoftwareSourceCode",
  "license": "https://spdx.org/licenses/BSD-3-Clause",
  "codeRepository": "https://github.com/cosimoNigro/agnpy",
  "contIntegration": "https://github.com/cosimoNigro/agnpy/actions",
  "dateCreated": "2019-12-17",
  "datePublished": "2022-01-31",
  "dateModified": "2021-08-02",
  "downloadUrl": "https://github.com/cosimoNigro/agnpy/releases/tag/v0.1.6",
  "issueTracker": "https://github.com/cosimoNigro/agnpy/issues",
  "name": "agnpy",
  "version": "0.1.8",
  "identifier": "10.5281/zenodo.4955175",
  "description": "agnpy is a python package focusing on the computation of the
  "applicationCategory": "astrophysics",
  "funding": "ESCAPE EU H2020 B24064",
  "developmentStatus": "active",
  "isPartOf": "https://www.astropy.org/affiliated/#affiliated-packages",
```



January 31, 2022

agnpy

agnpy is a python package focusing on the computation of the radiative processes of relativistic particles accelerated in the jets of Active Galactic Nuclei (AGN). It includes classes describing the galaxy components responsible for line and thermal emission and calculates the absorption due to gamma-gamma pair production on soft (IR-UV) photon fields.

Files (3.3 MB)	Size
cosimoNigro/agnpy-v0.1.8.zip	5.3 MB

Release	Date	Downloads
agnpy v0.1.8	2021	157
agnpy an open-source python package modelling the radiat...	2021	0
Flaremodel: An open-source Python package for one-zone num...	2021	0
ADS: 2022arXiv121112925M		0

# Software publication

## Where can HERMES help?

---

- **We have:**

- Publication repositories with metadata input forms/pull-based workflows
- Metadata from different sources, in different formats/modes

- **We want:**

- As little to do with forms as possible
- Control over what goes into the publication (push-based workflows)
- A collated set of rich metadata for publication

# HERMES

Automated software publication with rich metadata

---

- **HE**lmholtz **RI**ch **ME**tadata **S**oftware publication
- **Aim: Support RSEs in automatically publishing their software with rich metadata**
- **Helmholtz Metadata Collaboration Project:** 07/2021-06/2023 – ZT-I-PF-3-006
- **Three Helmholtz centres:**
  - German Aerospace Center (DLR)
  - Forschungszentrum Jülich
  - Helmholtz-Zentrum Dresden-Rossendorf
- **Concept paper:** *Software publications with rich metadata: state of the art, automated workflows and HERMES concept* [[arXiv:2201.09015](https://arxiv.org/abs/2201.09015) | [PubPeer](#)]
- **Website:** [software-metadata.pub](https://software-metadata.pub)

# Software publication with HERMES

## Scope

---

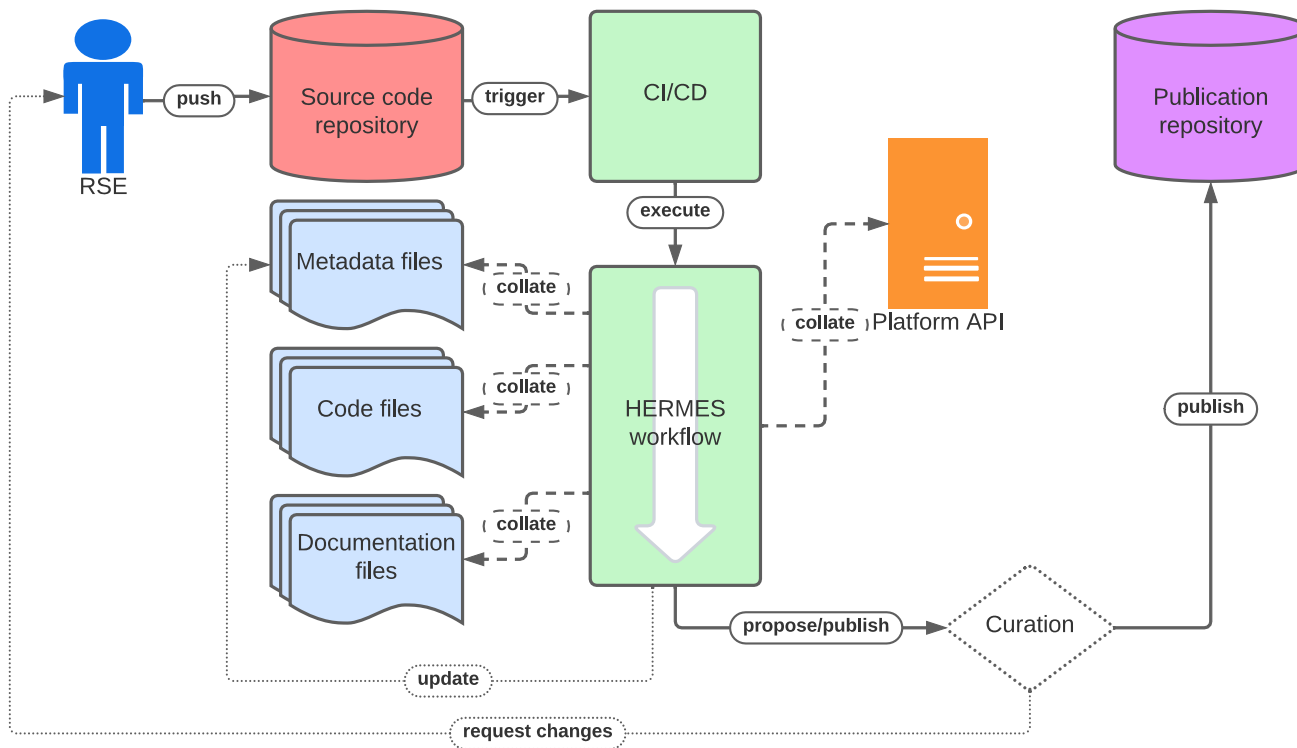
The user receives assistance in depositing software in an automated fashion. This may be used to create publications purely with rich metadata (to be at least FAIR [5], even for closed source software) or with attached artifacts like source code, executables, etc. (to be more easily reusable). To achieve this, HERMES provides

- an extensible, configurable and automatable toolchain with capability to be executed for<sup>15</sup>
  - N software publications in
  - M target publication repositories
  - from the same origin
  - as configured by the user,
- initially harvesting and collating **statically available metadata** from formerly described **metadata sources** and
- initially targeting
  - **InvenioRDM** and
  - **Dataverse project**
- for deposits of metadata and artifacts according to curator-defined requirements
- and output of the respective metadata in a structured format (e.g., **CodeMeta files**) for further reuse.

Druskat, S., Bertuch, O., Juckeland, G., Knodel, O., & Schlauch, T. (2022). *Software publications with rich metadata: state of the art, automated workflows and HERMES concept*. ArXiv, [abs/2201.09015](https://arxiv.org/abs/2201.09015).

# Software publication with HERMES

## Overview (simplified)



# What will HERMES give you?

## Project outputs

---

- **Software**
  - Software for software publication workflow automation
- **CI templates**
  - GitLab CI, GitHub Action, Jenkins, Travis CI
- **Training materials**
  - Adaption of HIFIS training materials to include workflow usage
- **Project website**
  - One-stop shop for information and documentation
- **Policy proposals**
  - Proposals for updates to policies/ guidelines at Helmholtz and Allianzinitiative

# Thank you!

## Where to learn more about project HERMES?

---



Stephan Druskat, DLR, PI, [@stdruskat](#)



Oliver Bertuch, FZJ, PI, [@poi\\_ki\\_lo\\_therm](#)



Guido Juckeland, HZDR, PI, [@GuidoJuckeland](#)



Oliver Knodel, HZDR, [@olikno1](#)



Tobias Schlauch, DLR, [@TobiasSchlauch](#)



Michael Meinel, DLR, [@led02](#)

- Find us on [Twitter](#)
- Write an email to [team@software-metadata.pub](mailto:team@software-metadata.pub)
- Go to [software-metadata.pub](http://software-metadata.pub)



**HELMHOLTZ**

Open Science