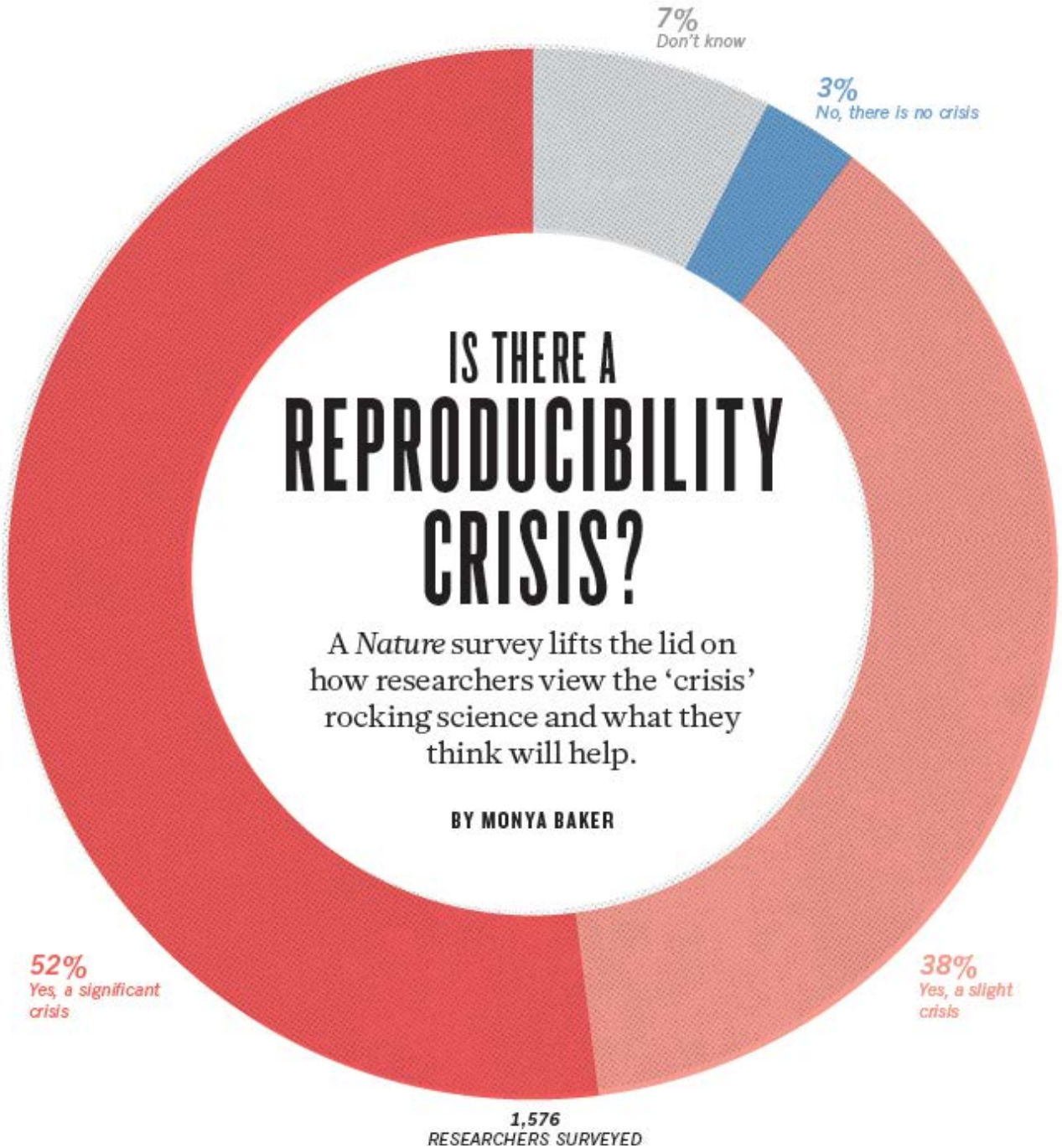
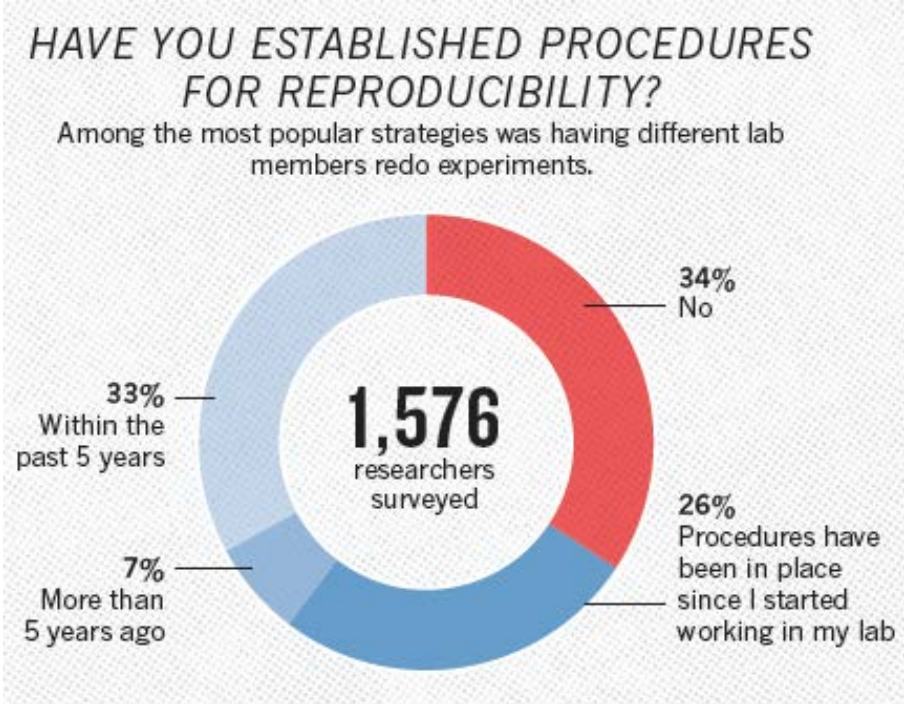
AGMB Arbeitsgemeinschaft für medizinisches Bibliothekswesen e.V.

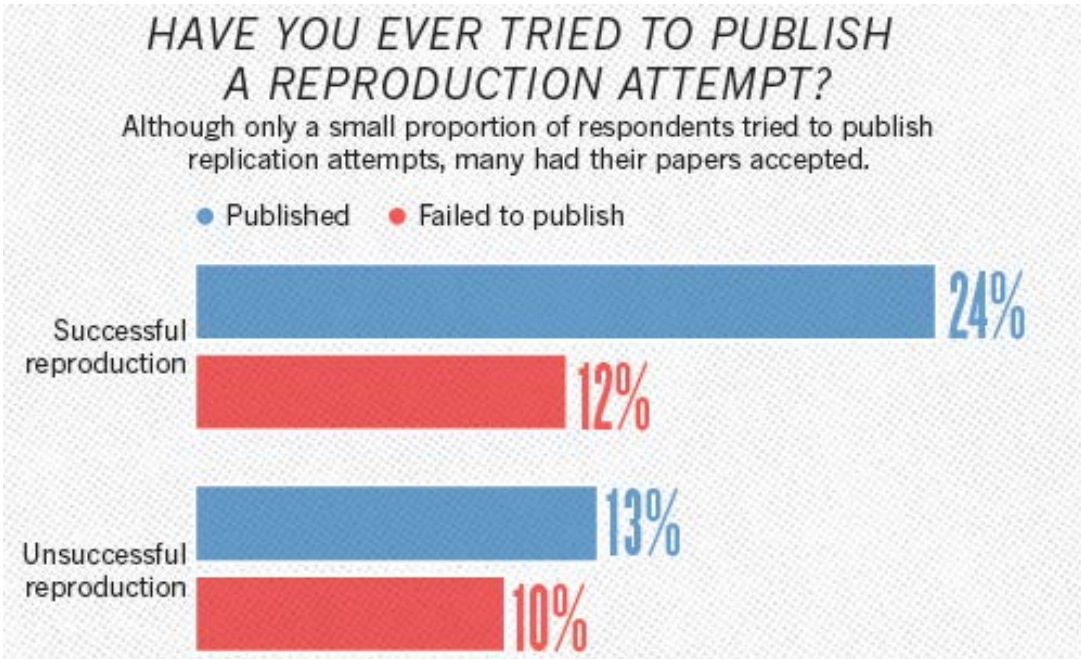
Jahrestagung 2019 der AG für Medizinisches Bibliothekswesen

FAIR Data Management als gemeinsame Herausforderung

Prof. Dr. Ulrich Sax
Institut für Medizinische Informatik
Universitätsmedizin Göttingen
ulrich.sax@med.uni-Goettingen.de
http://orcid.org/0000-0002-8188-3495
http://www.mi.med.uni-goettingen.de





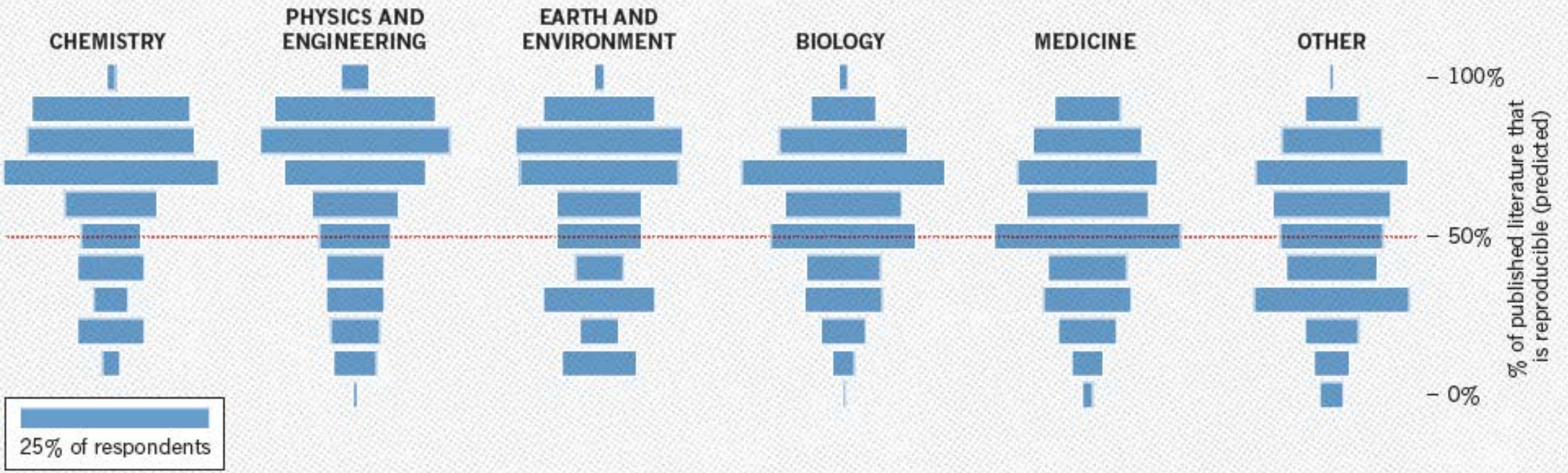


A 'CRISIS' IN NUMBERS

Nature surveyed 1,576 scientists online to get their thoughts on reproducibility in their field and in science in general. See **go.nature.com/2vjr4y** for more charts and access to the full data.

HOW MUCH PUBLISHED WORK IN YOUR FIELD IS REPRODUCIBLE?

Physicists and chemists were most confident in the literature.



452 | NATURE | VOL 533 | 26 MAY 2016

www.nature.com/scientificdata

UNIVERSITÄTSMEDIZIN IMG

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES
» Research data
» Publication characteristics

Comment: T Principles fo managemen

Mark D. Wilkinson *et al.*#

There is an urgent need to improve
set of stakeholders—representing a
come together to design and jointly
to as the FAIR Data Principles. The
enhance the reusability of their dat
scholar, the FAIR Principles put spec
find and use the data, in addition t
formal publication of the FAIR Princ
implementations in the community

Received: 10 December 2015
Accepted: 12 February 2016
Published: 15 March 2016

Box 2 | The FAIR Guiding Principles

To be Findable:
F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:
A1. (meta)data are retrievable by their 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 data are no longer available

To be Interoperable:
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

To be Reusable:
R1. meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1. (meta)data are released with a clear and accessible data usage license
R1.2. (meta)data are associated with detailed provenance
R1.3. (meta)data meet domain-relevant community standards

Wilkinson et al. 2016
SCIENTIFIC DATA
3:160018 | DOI:
10.1038/sdata.2016.18

SHARE

DAVIDE BONAZZI/SALZMAN ART

A solution to psychology's reproducibility problem just failed its first test

By David Adam | May. 23, 2019 , 10:00 AM

Behavior change is difficult—just ask any psychologist. A new study shows behavior change among psychologists is no different. Efforts to improve the robustness of research by asking psychologists to state their methods and goals ahead of time, a process called preregistration, have stumbled at the first hurdle.

Unmet plans

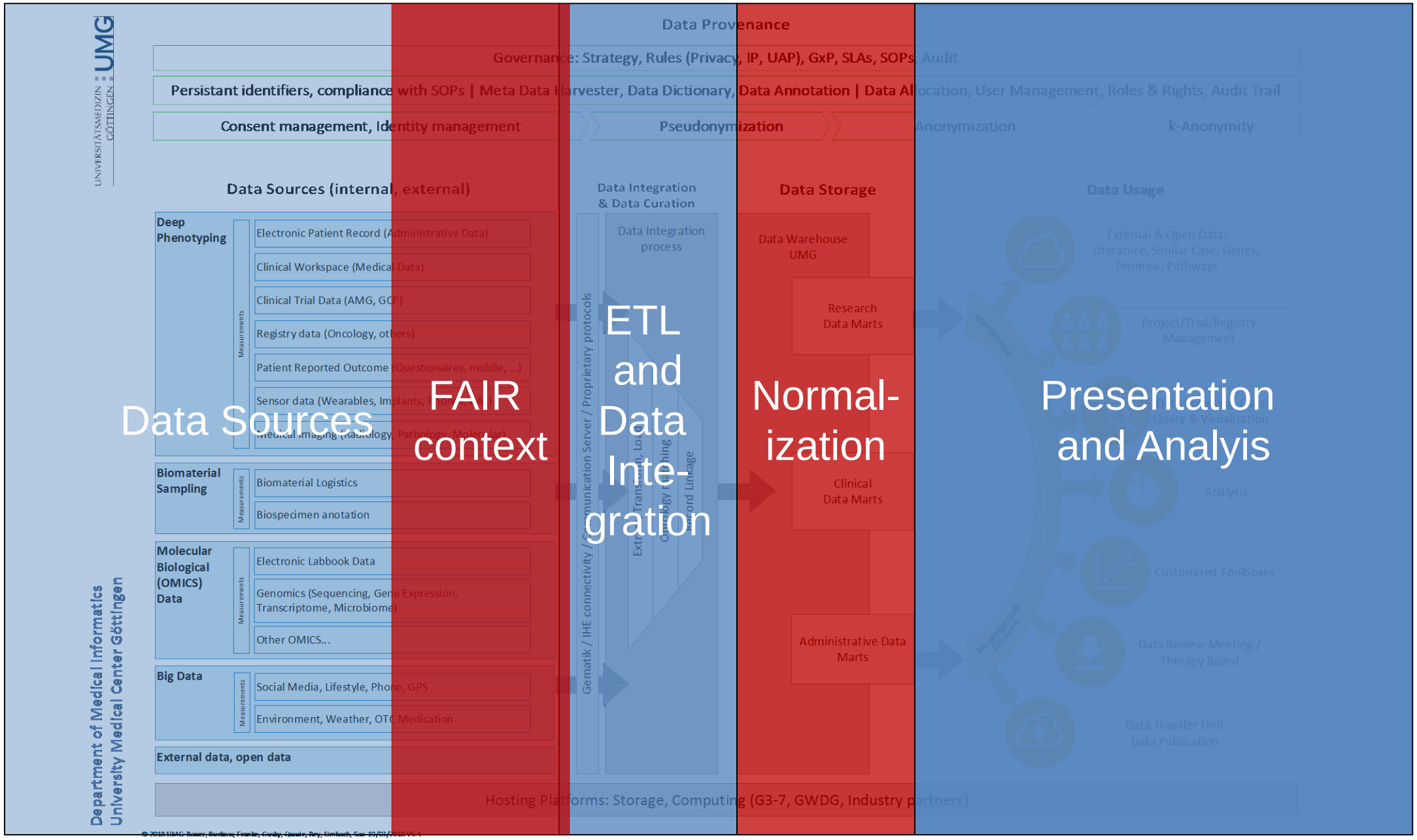
Of 27 studies in a psychology journal, just one followed its preregistered plan to a T. Researchers identified deviations from plans in eight categories.

	All deviations disclosed	Undisclosed deviations	No deviations
Hypothesis/ research question	3	5	19
Variables	0	4	23
Direction of effect	0	6	21
Operational-ization of variables	4	3	20
Sample size	5	10	12
Exclusion criteria	3	15	9
Procedure	2	1	24
Statistical model	6	13	8

The lack of transparency is troubling, but understandable, Vanpaemel says: Some researchers might fear their paper won't be published if they admit to not having entirely followed their preregistration. "As soon as we see more papers being published [with] transparent changes, these concerns will be hopefully lessened."

doi:10.1126/science.aa1207

Biomedical Informatics Pipeline



Medizinformatik-Initiative.de

MEDIZIN INFORMATIK INITIATIVE

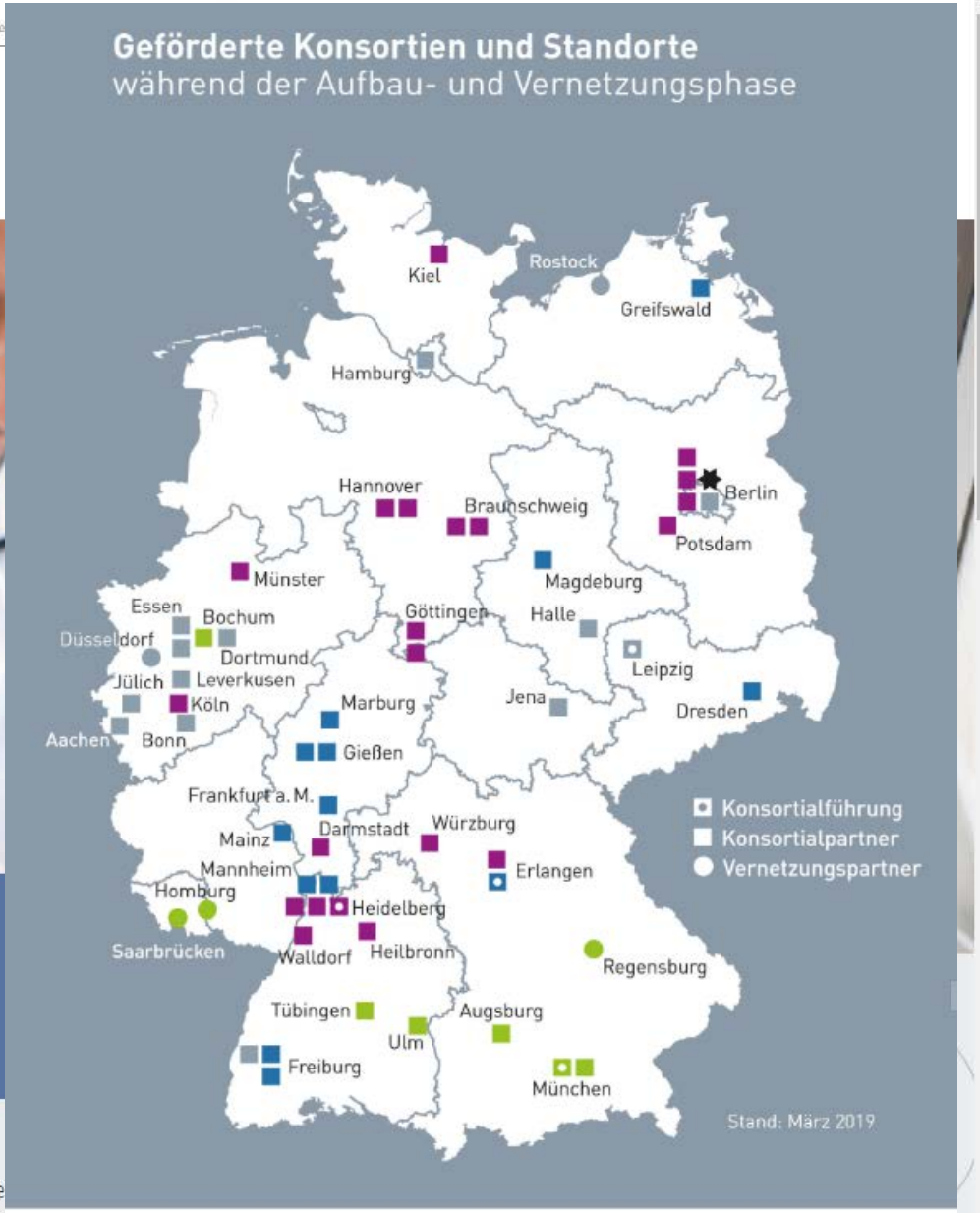
Suchbegriff eingebe

ÜBER DIE INITIATIVE KONSORTIEN ZUSAMMENARBEIT

Vernetzen. Forschen. Heilen.

Forschung stärken, Versorgung verbessern.
Medizininformatik.

Um Daten aus Krankenversorgung und Forschung besser nutzbar zu mache
Bundesministerium für Bildung und Forschung die Medizininformatik-Initiat
Millionen Euro. Die Fördermaßnahme soll die medizinische Forschung stärke
Patientenversorgung verbessern.



DIFUTURE
HiGHmed
MIRACUM
SMITH
Koordinationsstelle

MEDIZIN INFORMATIK INITIATIVE

Vorträge aus den Arbeitsgruppen des Nationalen Steuerungsgremiums

AG Interoperabilität: erste Erfolge und künftige Herausforderungen

Thomas Ganslandt[1], Martin Boeker[1], Björn Schreiweis[2], Birger Haarbrand[2], Danny Ammon[3], Silke Haferkamp[3], Holger Stenzhorn[4], Karoline Buckow[5], Ulrich Sax[2]

[1] miracum, [2] HiGHmed, [3] SMITH, [4] DIFUTURE, [5] Koordinierungsstelle des NSG (TMF)

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

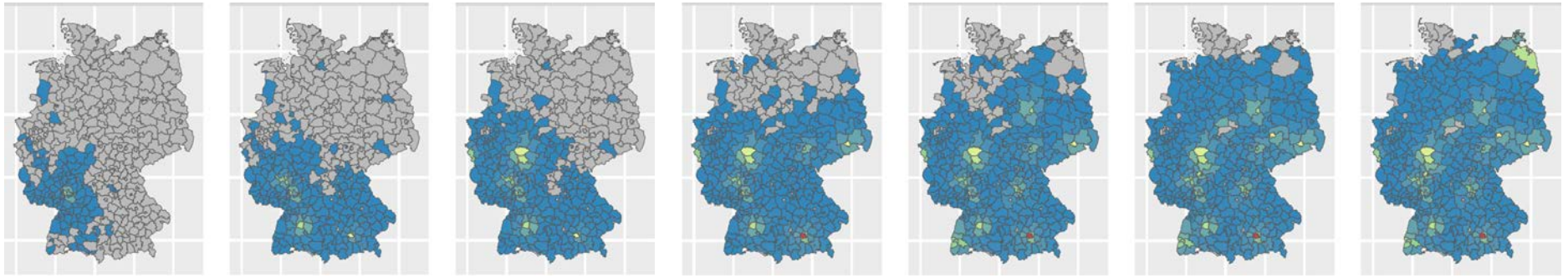
Demonstrator-Studie

20 Standorte

19 Freigaben

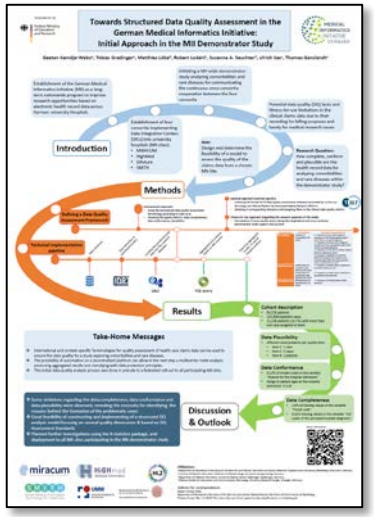
1,8 Mill. Patienten

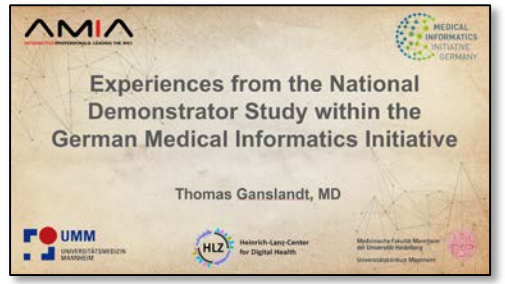
3,2 Mill. Fälle



Einzugsgebiet im Verlauf der Datenbereitstellung (09/2018 - 03/2019)

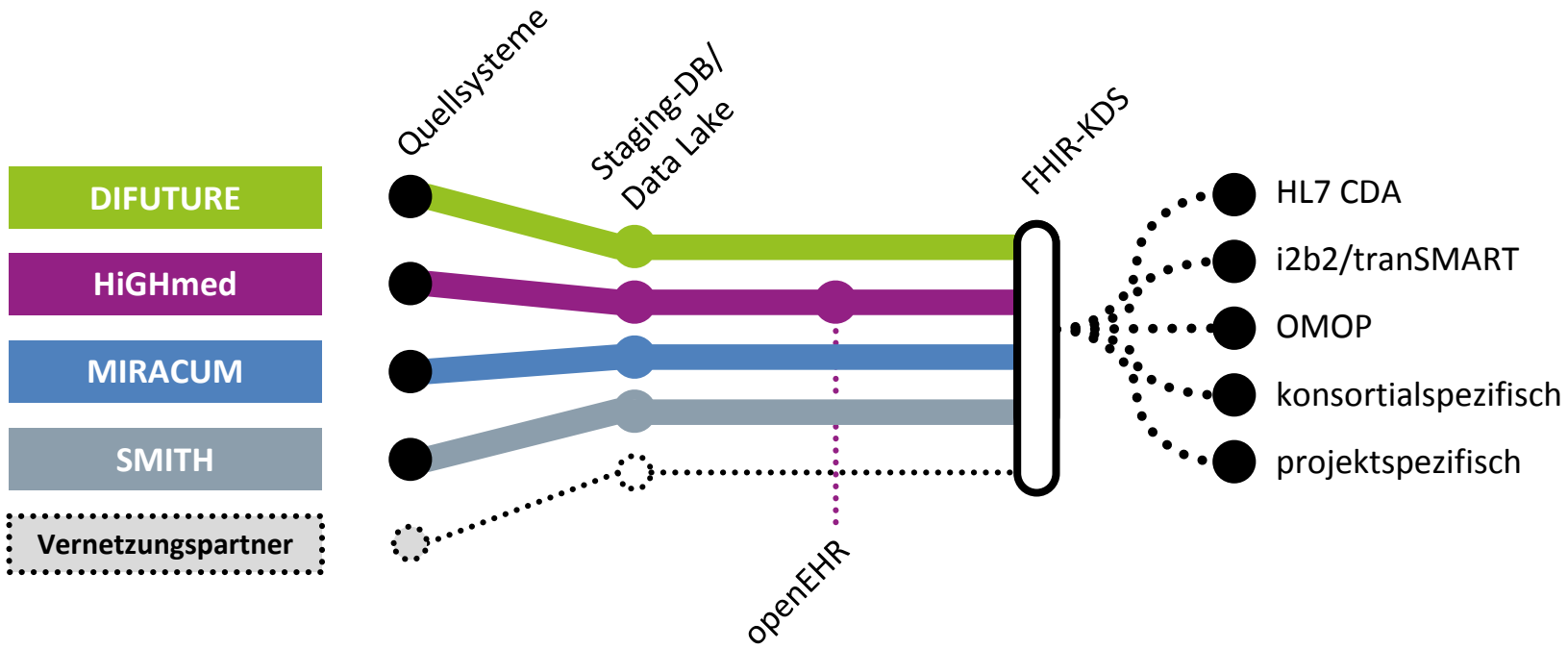
MEDINFO 2019 Lyon (Poster)





AMIA 2019 Washington (Vortrag)

Konvergenz durch Nutzung des MII-Kerndatensatzes



- Use Cases der MII
- Versorgungspartner außerhalb der MII
- Patientenportale
- Observational Health Data Sciences & Informatics (OHDSI)
- European Health Data & Evidence Network (EHDEN)
- Nationale Forschungs-Daten-Infrastruktur (NFDI)
- KI-Projekte

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Nationale Forschungsdaten-infrastruktur (NFDI) DFG

Die nationale Forschungsdateninfrastruktur (NFDI) soll die Datenbestände von Wissenschaft und Forschung systematisch erschließen, nachhaltig sichern und zugänglich machen sowie (inter-)national vernetzen. Sie wird in einem aus der Wissenschaft getriebenen Prozess als vernetzte Struktur eigeninitiativ agierender Konsortien aufgebaut werden.

Absichtserklärungen zum Aufbau der NFDI

Bis zum 4. Juli 2019 sind die Absichtserklärungen für eine Antragstellung im Jahr 2019 in Form von verbindlichen Voranmeldungen sowie für eine Antragstellung in den Jahren 2020 und 2021 in Form von unverbindlichen Absichtserklärungen in der DFG-Geschäftsstelle eingegangen.

Verbindliche Voranmeldungen für eine Antragstellung im Jahr 2019:

- NFDI letters of intent - National Research Data Infrastructure for Personal Health Data (NFDI4Health) (PDF | 96 KB)
- NFDI letters of intent - Nationale Forschungsdateninfrastruktur für die Ingenieurwissenschaften (NFDI4Ing) (PDF | 350 KB)
- NFDI letters of intent - Consortium of the Medical Informatics Initiative (MII) and the German Centers for Health Research (DZG) (NFDI4Medicine) (PDF | 185 KB)
- NFDI letters of intent - National Research Data Infrastructure for Mobility Technology (NFDI4MobilTech) (PDF | 170 KB)
- NFDI letters of intent - National Research Data Infrastructure for Materials Science and Engineering (NFDI4MSE) (PDF | 414 KB)
- NFDI letters of intent - Particle, Astroparticle, Hadron and Nuclear Physics accelerates the NFDI (PAHN-PaN) (PDF | 259 KB)
- NFDI letters of intent - Text+: Language- and Text-Based Research Data Infrastructure (Text+) (PDF | 157 KB)
- NFDI letters of intent - Mathematical Research Data Initiative (MaRDI) (PDF | 169 KB)
- NFDI letters of intent - NFDI for Agricultural Sciences (NFDI4Agri) (PDF | 270 KB)
- NFDI letters of intent - Biodiversity, Ecology & Environmental Data (NFDI4BioDiversity) (PDF | 249 KB)
- NFDI letters of intent - NFDI for Catalysis-Related Sciences (NFDI4cat) (PDF | 136 KB)
- NFDI letters of intent - Fachkonsortium Chemie für die Nationale Forschungsdateninfrastruktur (NFDI4Chem) (PDF | 240 KB)
- NFDI letters of intent - Consortium for research data on material and immaterial ulture) (PDF | 338 KB)
- nsortium Earth System Science / NFDI Konsortium Erdsystemforschung (NFDI4Earth) (PDF | 312 KB)

https://www.dfg.de/foerderung/programme/nfdi/absichtserklaerungen/index.html

Gene	Patient's Variant	Level of Evidence	Seg. Mean	Size (Mb)
AURKA	ampl.	B3	0.58	39.2
BIRC7	ampl.	B3	0.58	39.2
BRCA1	del.	A2 ,B2 ,A3 ,B3	-0.84	2.6
CCNE1	ampl.	A3 ,B3	0.73	15.2
CDK12	del.	A3	-0.87	1.9
FANCA	del.	B2	-0.71	59.8
FRS2	ampl.	B3	0.73	6.5
MDM2	ampl.	B2	0.73	6.5
NF1	del.	B2 ,B3	-0.85	2.6
PALB2	del.	B2 ,B3	-0.70	17.5
RICTOR	ampl.	B2	1.56	12.5
SUZ12	del.	B3	-0.87	4.5
TOP1	ampl.	B2	0.58	39.2
TSC2	R505X,del.	B2 ,B3	-0.78	11.2

Levels of Evidence: Findings are classif ed into 6 levels of evidence combining the axis A-B and the axis 1-2-3. Level A means evidence in the same cancer type. Level B means evidence in any other cancer type. On the 1-2-3 axis, level 1 means evidence supported by drug approval organizations or clinical guidelines, level 2 contains clinical evidence (clinical trials, case reports) and level 3 consists of preclinical evidence. The distribution of f ndings into levels is summarized in the right f gure

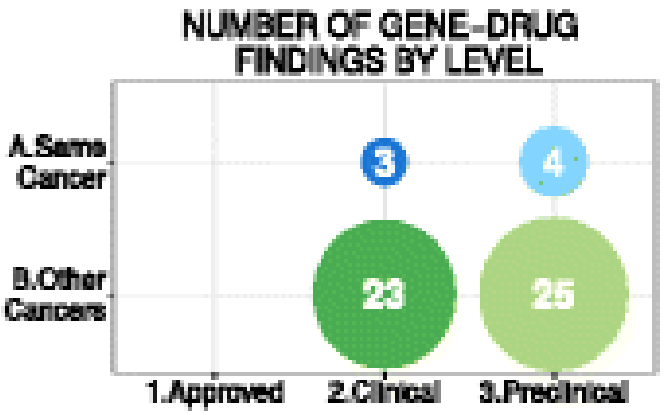
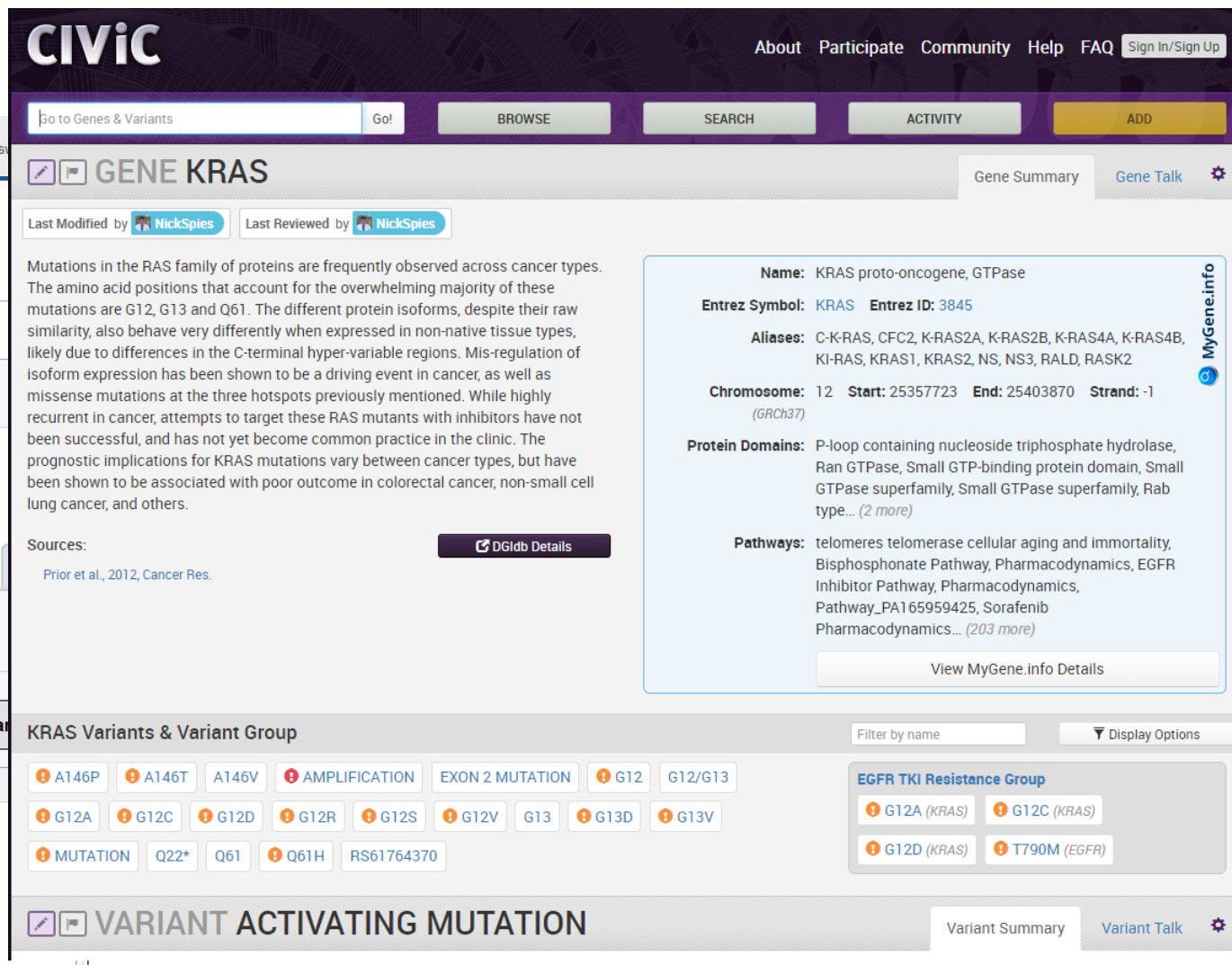


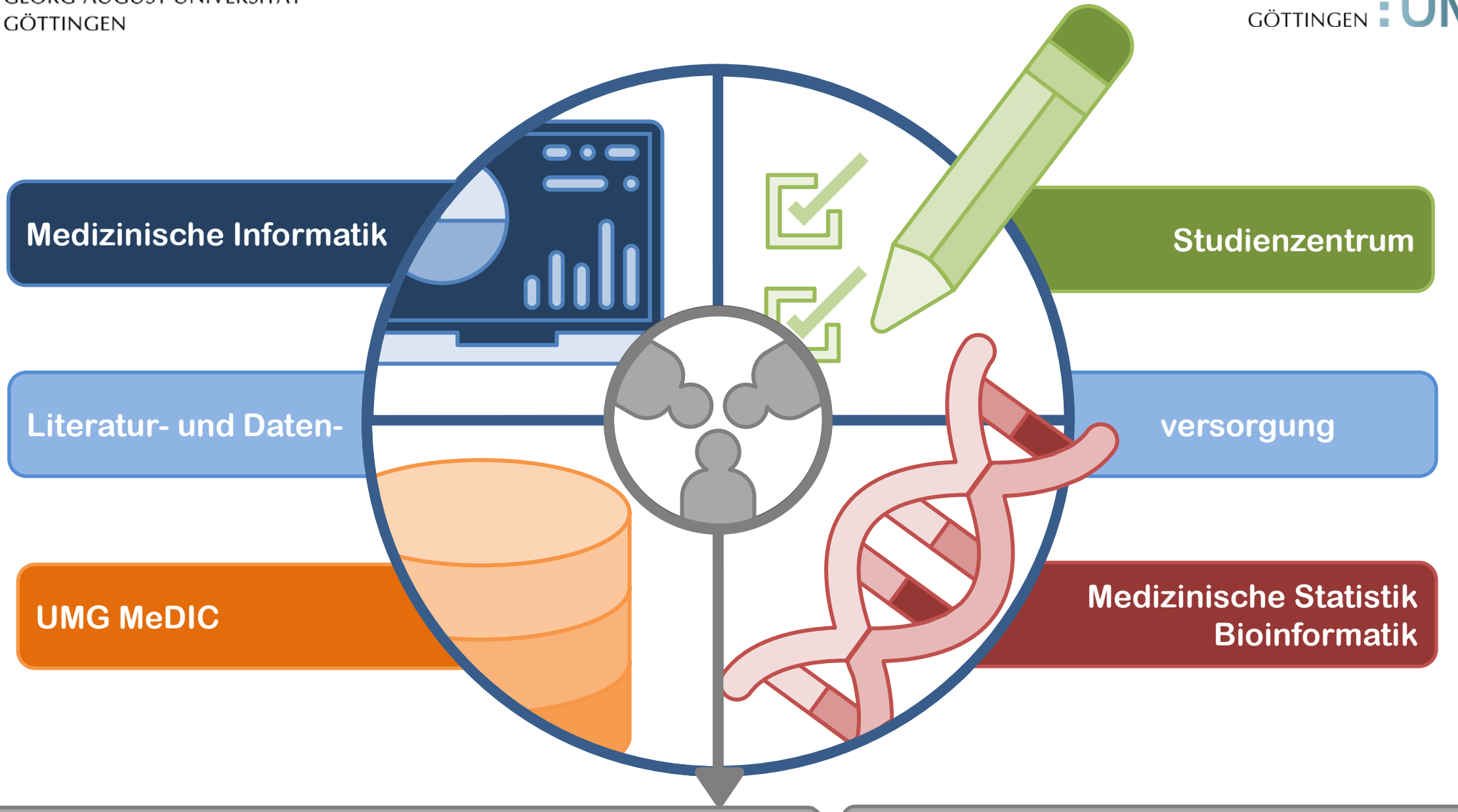
Table of Results: All the predicitve associations are detailed in this table. The results are sorted by 1) drug frequency, 2) levels of evidence (A1-B1-A2-B2-A3-B3). To allow a quick interpretation, the type of association (response, resistance) is colored (green, red) and new variants are gray and underlined.



3	☐	Recruiting	Systematic Hereditary **Pancreatic Cancer** Risk Assessment and Implications for Personalized Therapy
4	☐	Recruiting	Losartan and Nivolumab in Combination With FOLFIRINOX and SBRT in Localized **Pancreatic Cancer**
5	☐	Recruiting	Study of Gemcitabine, Nab-paclitaxel, and Ficlatuzumab (AV-299) in Patients With Advanced **Pancreatic**

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

UNIVERSITÄTSMEDIZIN
GÖTTINGEN UMG

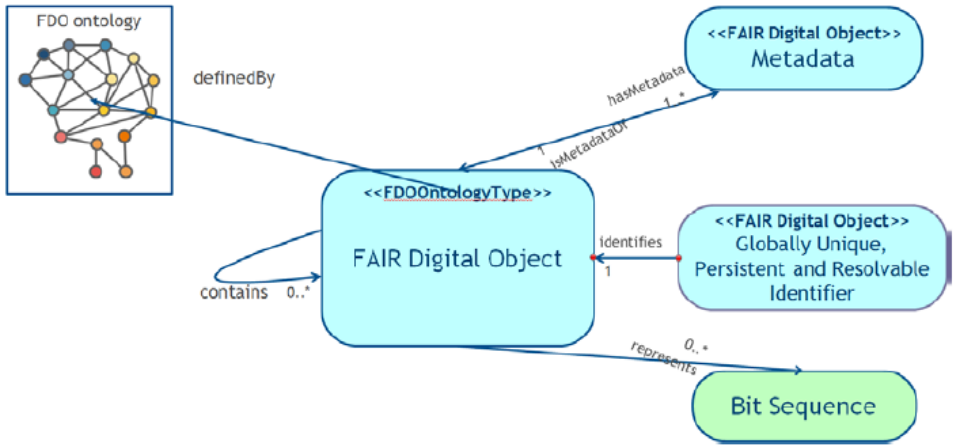


Netz, Server, Speicher, Nutzer, Management-Sicht, Applikationen

secuTrial, REDCap/OpenClinica, SEEK, tranSMART

OFFENE THEMEN

- Modell für FAIR Digital Objects



- FAIR Assessment Frameworks

Zusammenfassung

- Biomedizinische Daten sind der Motor für die klinische und biomedizinische Forschung. Entscheidend für die Nutzbarkeit der Daten ist es, den Kontext bzw. Zweck der Datenerhebung (Provenance) zusammen mit den Daten vorzuhalten.
- Für die Interpretation der Daten müssen ähnliche Fälle bzw. Erkenntnisse aus öffentlichen Datenbanken bzw. der Literatur mit herangezogen werden.
- Das ist eine gemeinsame Herausforderung der Lebenswissenschaften und der Bibliothekswissenschaften.
- Eine disziplinübergreifende Agenda zu den FAIR guiding principles würde helfen:
    - Systematische automatisierte Literatursuche
    - Übergang von wissenschaftlichen Publikationen zu verbundenen Datenpublikationen
    - Nachhaltige Findbarkeit und Zugänglichkeit von Daten und Publikationen (!)

Department of Medical Informatics

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

AGMB Arbeitsgemeinschaft für medizinisches Bibliothekswesen e.V.

Jahrestagung 2019 der AG für Medizinisches Bibliothekswesen

FAIR Data Management als gemeinsame Herausforderung

Prof. Dr. Ulrich Sax
Institut für Medizinische Informatik
Universitätsmedizin Göttingen
ulrich.sax@med.uni-Goettingen.de
http://orcid.org/0000-0002-8188-3495
http://www.mi.med.uni-goettingen.de