



From St. Gallen into the R-Universe — towards a national collaboration platform for official statistics

adminR Spring Meetup Bern-Liebefeld
06.06.2024

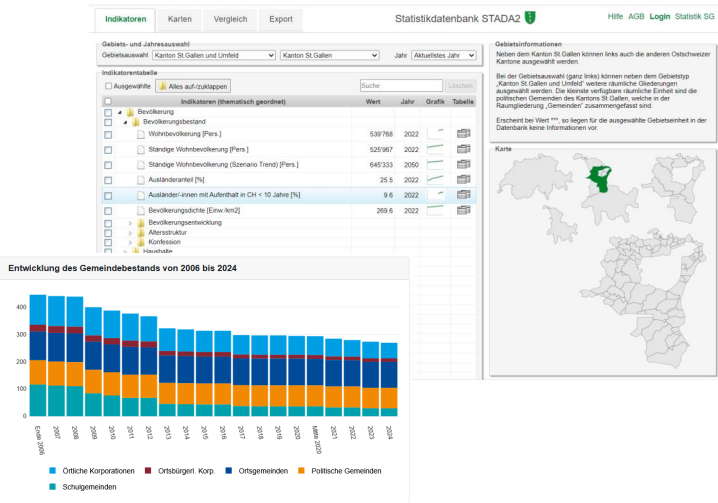
Jan Wunder, Félix Luginbühl
Fachstelle für Statistik Kanton St.Gallen

Overview

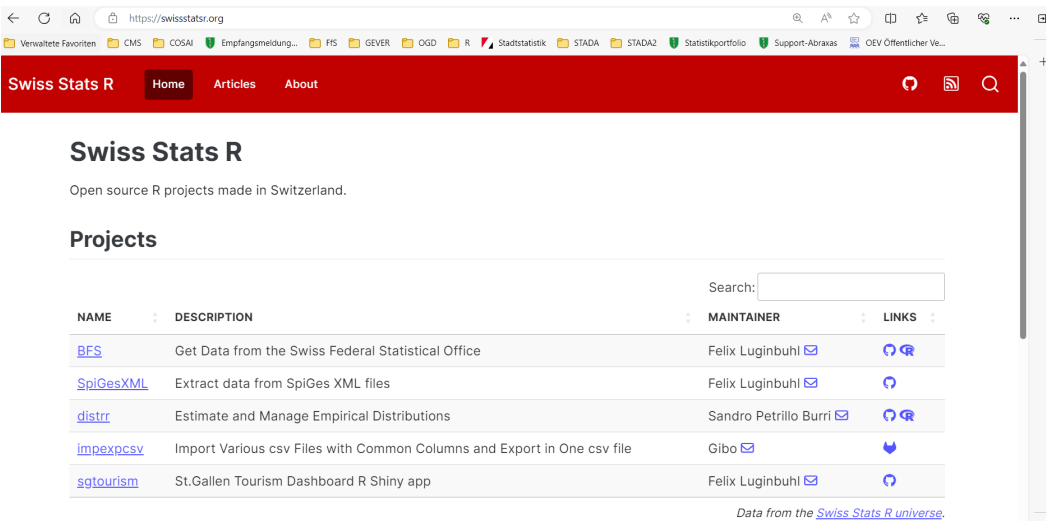
Our team



Our infrastructure



Our collaboration platform (Félix)

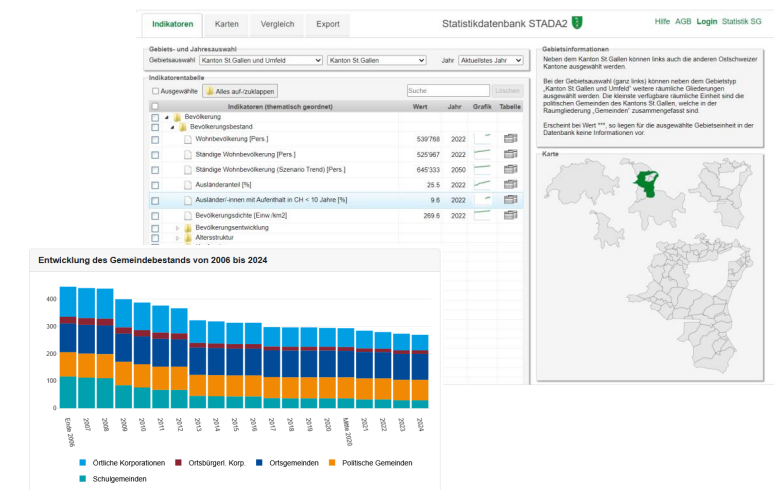


Fachstelle für Statistik Kanton St.Gallen

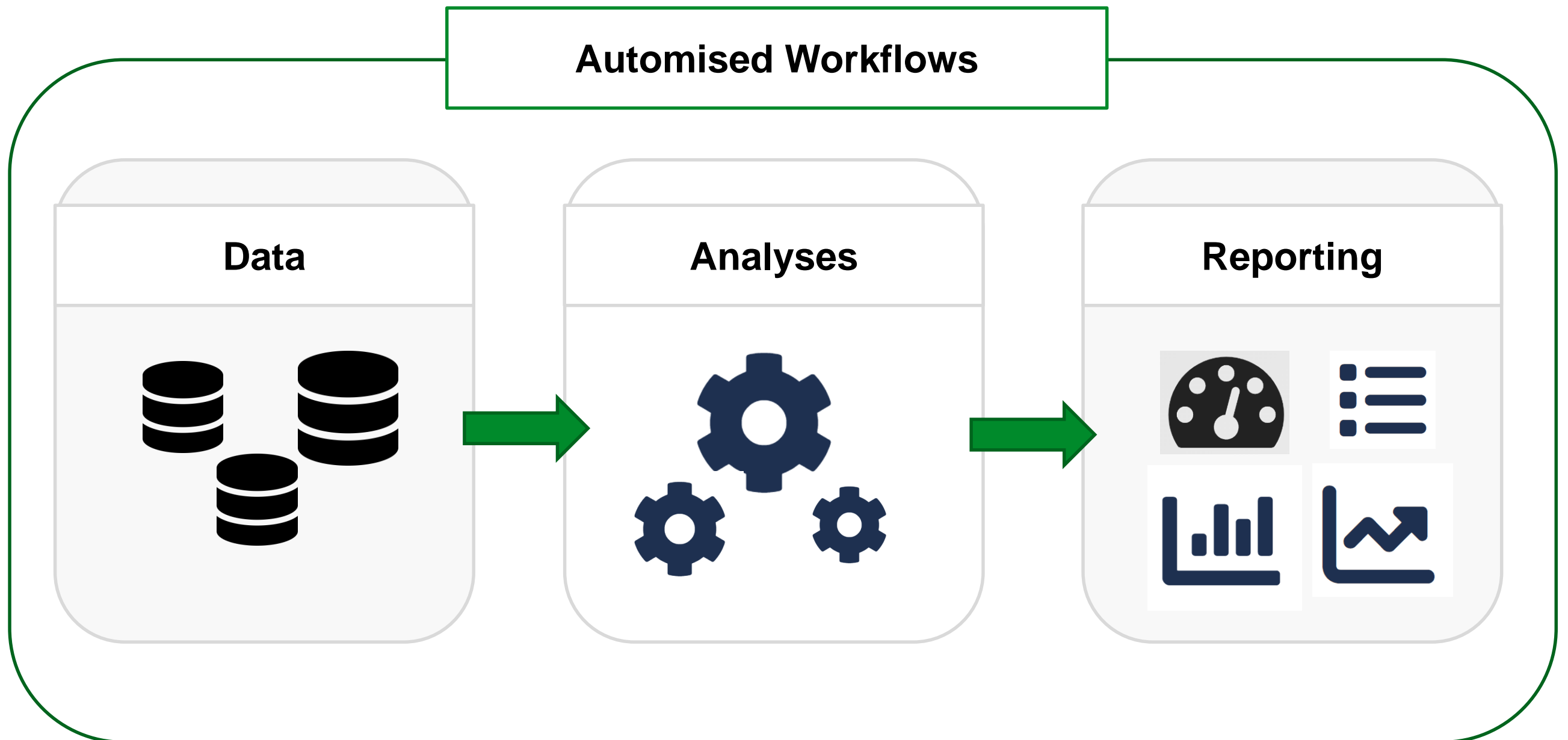
The Wunder-Team ☺

In a nutshell:

- 12 Data scientists/analysts
- Median-experience: ca. 15 years
- Representing Canton and City St.Gallen
- 400+ specific enquiries per year (administration, enterprises, media, political parties, citizens, students, scientists)
- 100'000+ annual access to our statistical data bases and the statistics portal statistik.sg.ch

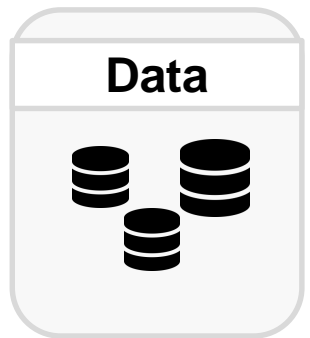


The Future of Data and Statistics

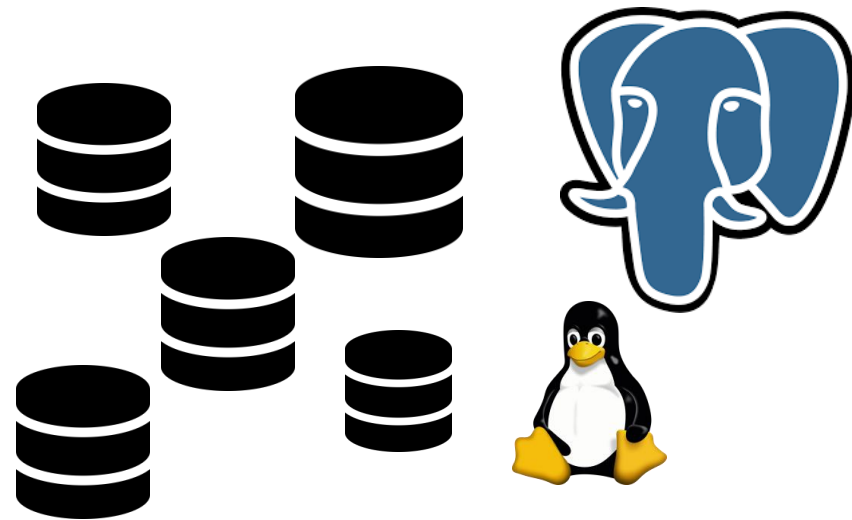


Strategic Focus Data

«Data is the new gold»



New Data Warehouse (DWH
for FfS-Data & other offices



Boost OGD, incl. Linked Open Data
(Construct a Triple Store)



5 stars of Open Data

5-Star Deployment Scheme for Open Data (source: 5stardata.info)

- Standardise metadata/definitions, increase transparency
- Goal: Statistical Office for Data and Statistics



Strategic Focus Analysis

New Methods: Machine Learning (ML) & «Experimental Statistics»



- Valorisation of large data resources!
- «Nowcasting» and «Forecasting»
- Population projections
- Braindrain analyses
- AHV/IV-Projections (COSAI)
- ...



<https://www.destatis.de/EN/About-Us/Events/Machine-Learning>

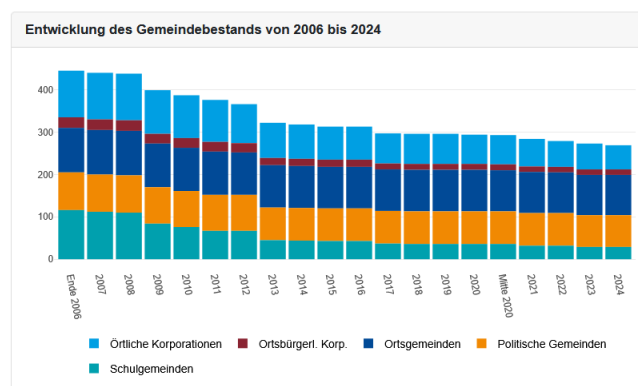
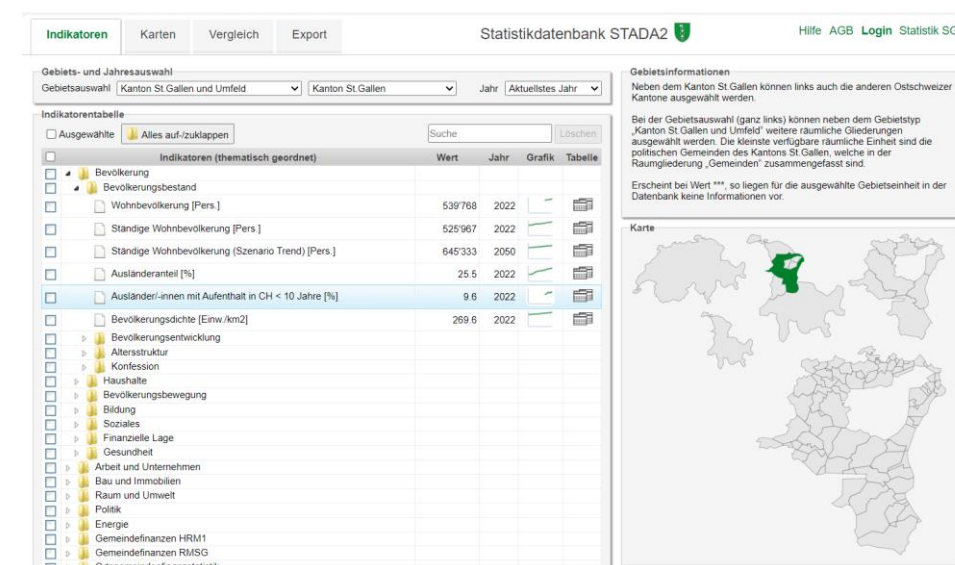


Strategic Focus Reporting

Interactivity for all!

- New web prescence
- New interactive dashboards & tools
- New municipality portrait/dashboard
- Online reports
- Data stories
- ...

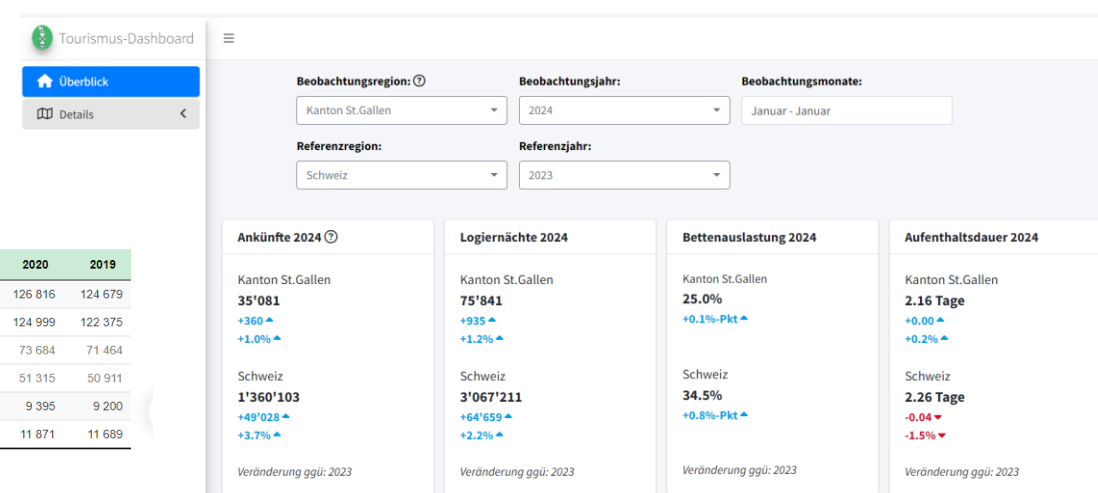
Reporting



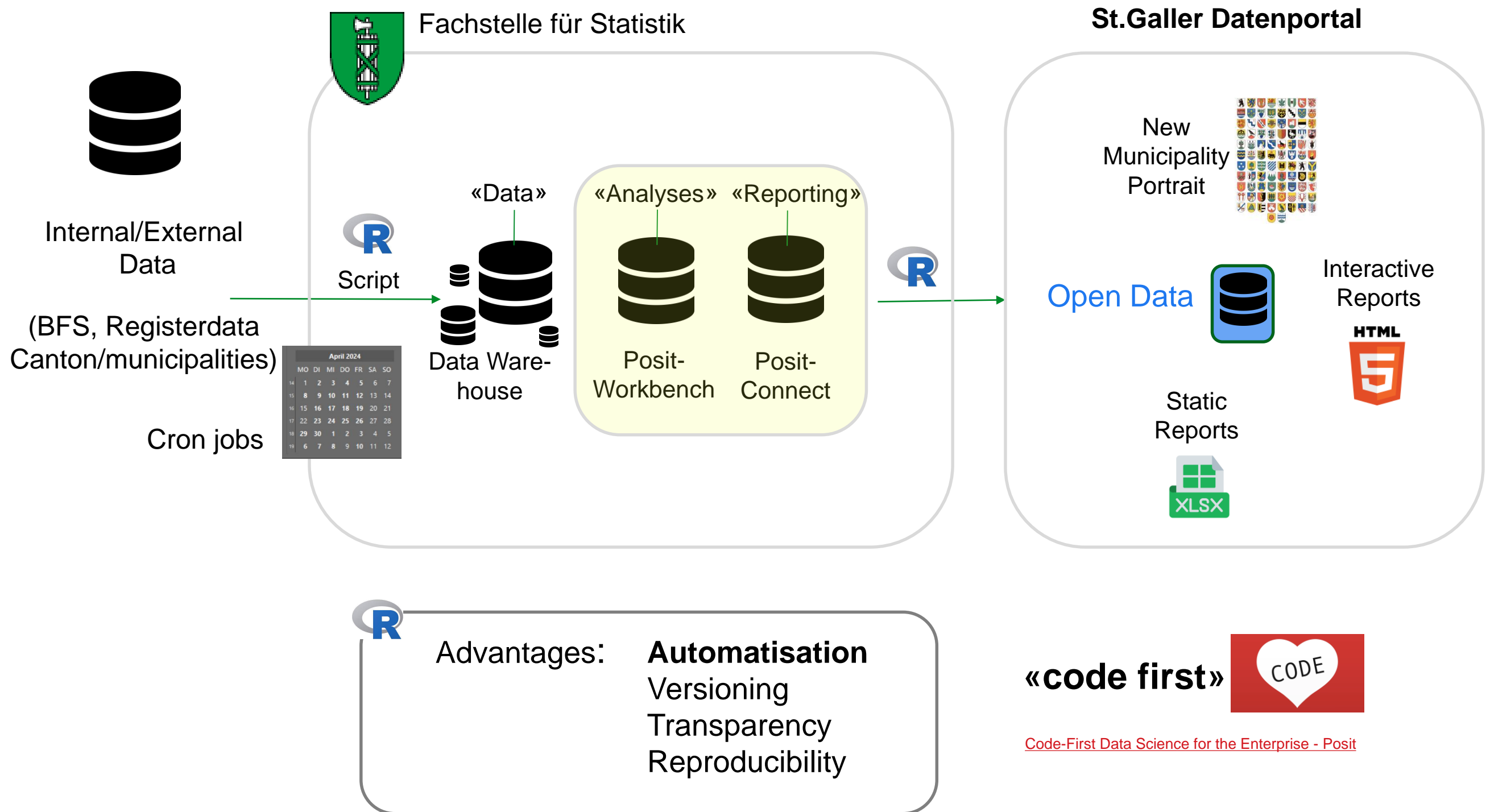
Ausländerbestand

	2022	2021	2020	2019
Bestand ausländische Wohnbevölkerung Ende Jahr total	133 924	129 319	126 816	124 679
Bestand ständige ausländische Wohnbevölkerung	132 374	127 482	124 999	122 375
- davon: EU/EFTA (FZA)	79 210	75 146	73 684	71 464
- davon: Drittstaaten	53 164	52 336	51 315	50 911
Grenzgänger (gem. GGS)	9 834	9 581	9 395	9 200
Grenzgängerbewilligungen	12 371	12 048	11 871	11 689

Quellen: SEM, ZEMIS; Grenzgänger: BFS, GGS.
© Fachstelle für Statistik SG



New infrastructure: «code first» at its core - Workflows & Automatisations



New architecture: New worlds!

Introduction of several new tools,
often first application in SG in 2023/2024

Neu @ SG



Gitea

Neu @ SG



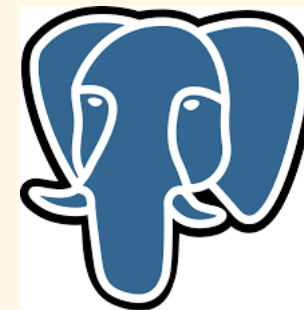
Versioning

Neu @ SG



Linux-Server

Neu @ SG



Postgres - DB

Data mgmt

Data analyses

Neu @ SG



Wiki-Systeme

Knowledge mgmt



New architecture: «All-in-one-solution»



A programming language for statistical computing and graphics

«Come for the language, stay for the community»

Colin Fay, 2019

Academia

ETH zürich



unhe
UNIVERSITÉ DE
NEUCHÂTEL

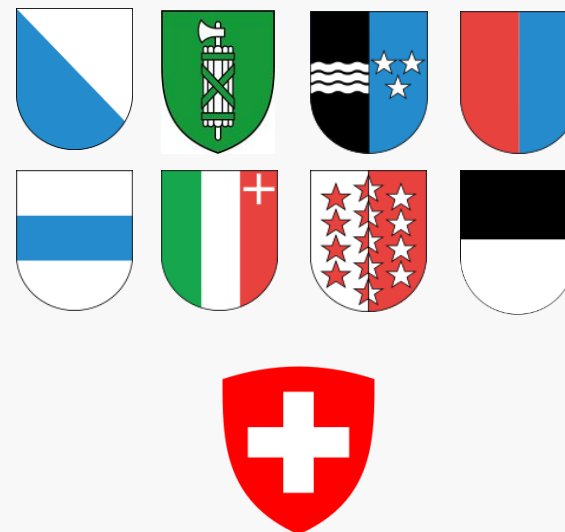


Unil
UNIL | Université de Lausanne

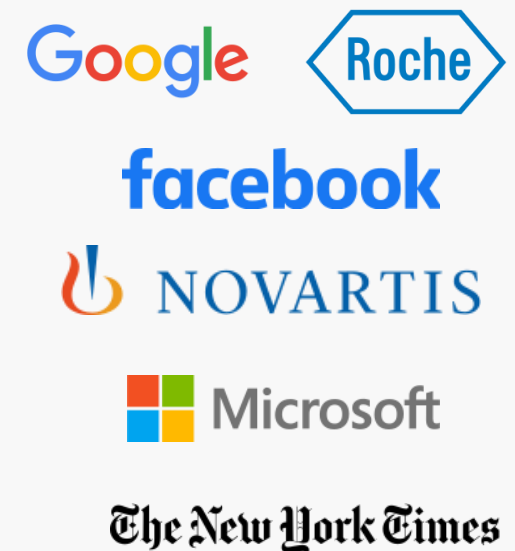
**UNI
FR**
UNIVERSITÉ DE Fribourg
UNIVERSITÄT FREIBURG



Administration



Wirtschaft



Collaboration at all levels of government & CH+

Municipalities

Statistical office of the City of St.Gallen

Union of St.Gallen municipalities

Canton St.Gallen

National Working group CORSTAT

Statistics platform Lake Constance IBK

Destatis/DACH/EUROSTAT



A brief history of national R-collaboration

Some key dates:

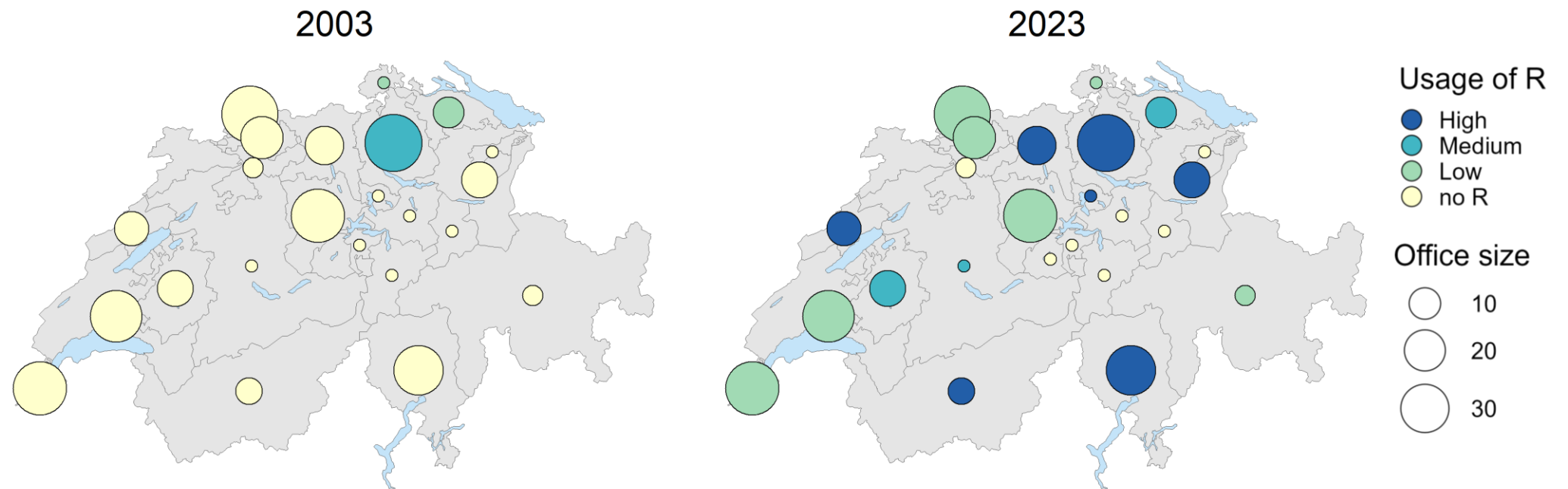
- **2018 in Zurich & Berne**
Foundation of adminR
List of curated packages on Github (2021)
- **2022 in Aarau (adminR spring meetup)**
Call for closer R collaboration at the national level
- **2022 in Lucerne (Swiss Statistics Days)**
Poster to call for a national R code base
- **2023 in Basel (Swiss Statistics Days)**
Announcement of R Working Group at CORSTAT



The situation at the national level



Nationwide
Survey



Key results

- In 2023, 7 Cantons use R as main tool (2024: 8 Cantons)
- Between 2003 and 2023, almost all larger Cantons increased importance of R
- Currently, several Cantons reflect on further increasing or introducing R

Swiss Statistics Days in Basel 2023




New R-Working Group @ CORSTAT

	<p>Founded: October 31, 2023</p> <p>Members: 17 offices (13 cantons, 4 cities)</p> <p>SG, ZH, VS, BL, AG, NE, TG, ZG, VD, TI, BS, FR, LU</p> <p>Stadt Zürich, Stadt Bern, Ville de Lausanne, Stadt St.Gallen</p> <p>31 PLANNER resp. CODER</p>
---	--

- Transversal KORSTAT group
- Collaboration of the «R-Cantons» to plan and develop R packages for nationwide use
- Coordination of R activities at Swiss statistical offices
- Working/coding/communication language: EN



Collaboration: CORSTAT «R» working group

Persons	Role
 <p>Jan Wunder jan.wunder@sg.ch</p> <p>Telefon: 058 22 98 735</p> <p>Website: http://www.statistik.sg.ch</p> <p>Position: Stellenleiter</p> <p>Abteilung: Fachstelle für Statistik Kanton St.Gallen</p> <p>Ort: St.Gallen</p>	Co-Lead
 <p>Andrea Schnell andrea.schnell@statistik.ji.zh.ch</p> <p>Telefon: 043 259 75 58</p> <p>Position: Co-Amtsleiterin</p> <p>Abteilung: Statistisches Amt Kanton Zürich</p>	Co-Lead



Collaboration: SwissStatsR.org

The screenshot shows the homepage of SwissStatsR.org. The browser's address bar displays 'https://swissstatsr.org'. The website has a red header with navigation links: 'Swiss Stats R', 'Home', 'Articles', and 'About'. On the right side of the header are icons for GitHub, RSS, and a search function. Below the header, the main heading 'Swiss Stats R' is followed by the tagline 'Open source R projects made in Switzerland.' To the right, a yellow callout box states 'Co-Lead SG-ZH 17 Cantons & Cities'. The 'Projects' section features a table with columns for NAME, DESCRIPTION, MAINTAINER, and LINKS. A search bar is located above the table. The table lists five projects: BFS, SpiGesXML, distr, impexpcsv, and sgtourism, each with its description, maintainer, and links to GitHub and other resources. At the bottom of the projects section, it says 'Data from the Swiss Stats R universe.'

NAME	DESCRIPTION	MAINTAINER	LINKS
BFS	Get Data from the Swiss Federal Statistical Office	Felix Luginbuhl	
SpiGesXML	Extract data from SpiGes XML files	Felix Luginbuhl	
distr	Estimate and Manage Empirical Distributions	Sandro Petrillo Burri	
impexpcsv	Import Various csv Files with Common Columns and Export in One csv file	Gibo	
sgtourism	St.Gallen Tourism Dashboard R Shiny app	Felix Luginbuhl	

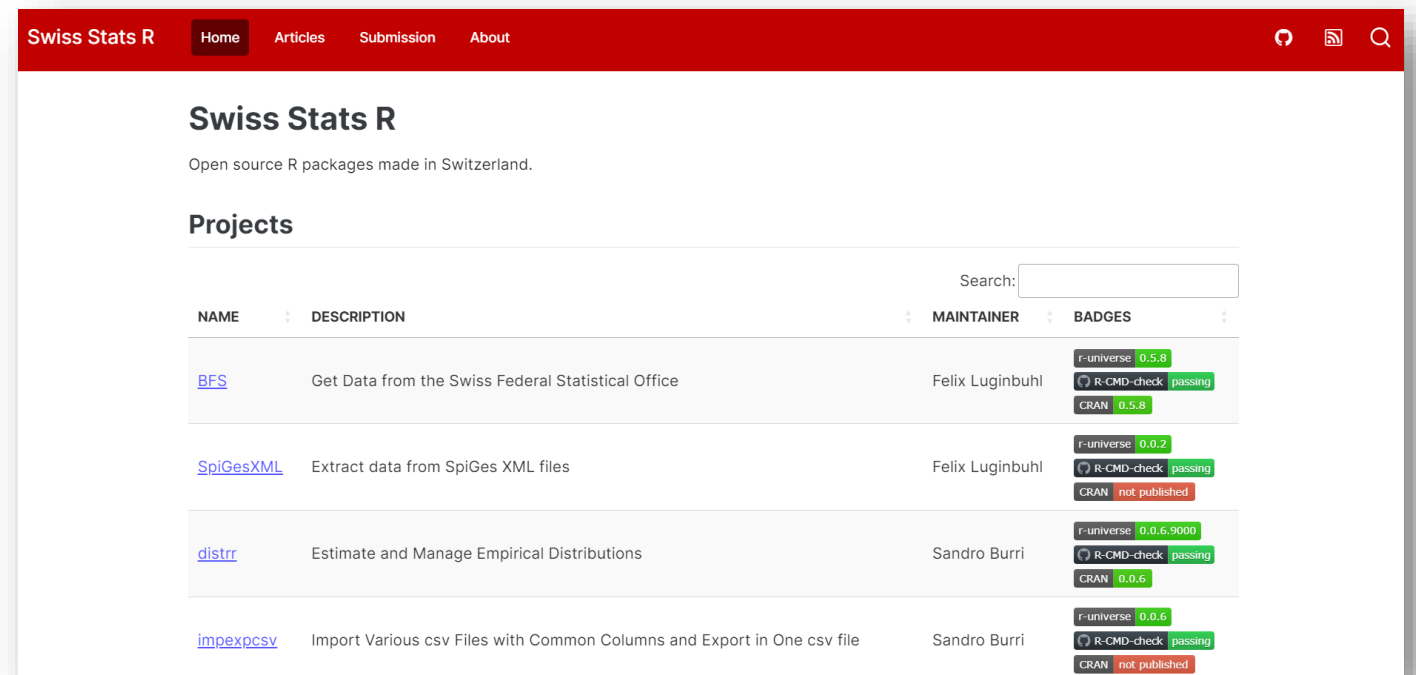
Data from the [Swiss Stats R universe](#).

Developed with ❤️ in [St.Gallen](#) using [Quarto](#) and [Zephyr](#) theme.



Collaboration: SwissStatsR.org

- All available open source R packages are accessible in the «home» page.
- The link in the **Name** column gives access to the GitHub sources of the R packages.
- The **Badges** provide information about the R Universe version of the R package, based on the R-CMD-Check and its availability on CRAN.
- The *Search* tab at the top right of the interactive table allows to quickly find any R package.



The screenshot shows the SwissStatsR.org website. At the top is a red navigation bar with links: Home, Articles, Submission, and About. Below the navigation bar is the title "Swiss Stats R" and the subtitle "Open source R packages made in Switzerland." The main content area is titled "Projects" and features a search bar. Below the search bar is a table with the following columns: NAME, DESCRIPTION, MAINTAINER, and BADGES. The table lists four projects: BFS, SpiGesXML, distr, and impexpcsv. Each project row includes a link to its GitHub source, a description, the maintainer's name, and a set of badges showing the R-universe version, R-CMD-check status, and CRAN availability.

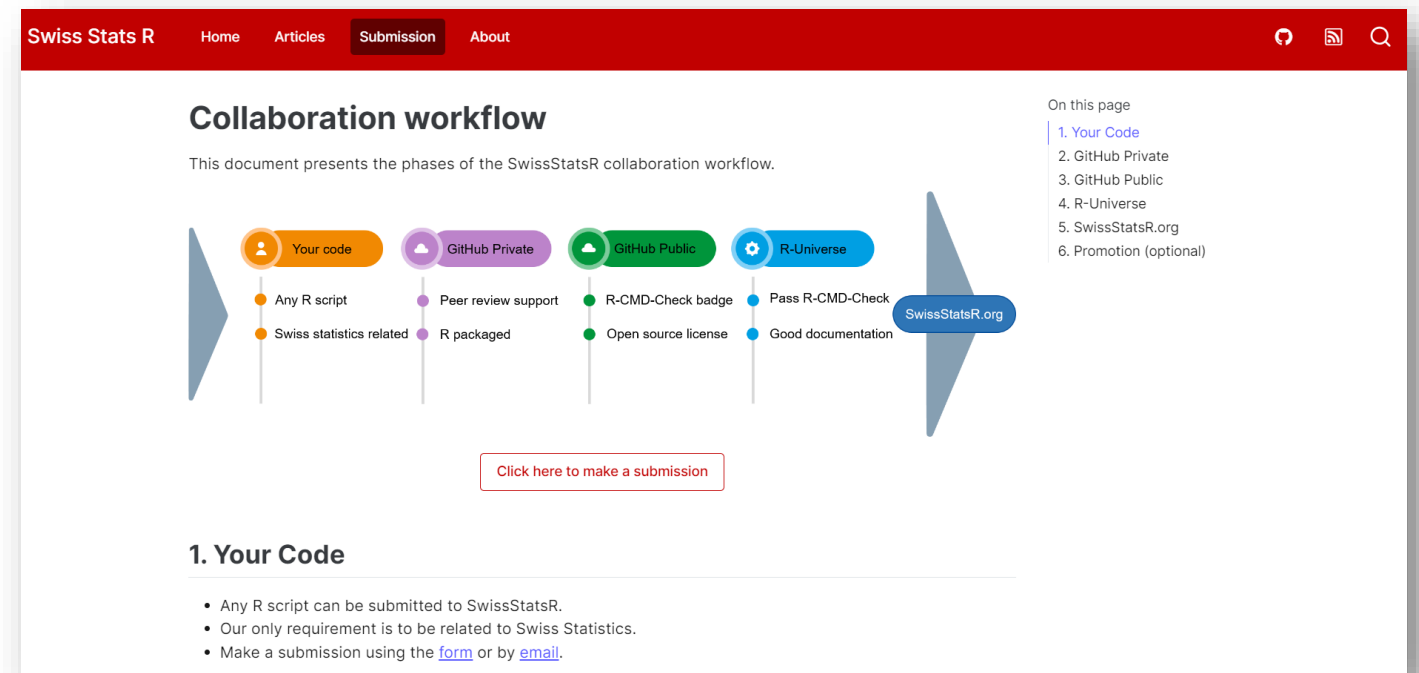
NAME	DESCRIPTION	MAINTAINER	BADGES
BFS	Get Data from the Swiss Federal Statistical Office	Felix Luginbuhl	R-universe 0.5.8 R-CMD-check passing CRAN 0.5.8
SpiGesXML	Extract data from SpiGes XML files	Felix Luginbuhl	R-universe 0.0.2 R-CMD-check passing CRAN not published
distr	Estimate and Manage Empirical Distributions	Sandro Burri	R-universe 0.0.6.9000 R-CMD-check passing CRAN 0.0.6
impexpcsv	Import Various csv Files with Common Columns and Export in One csv file	Sandro Burri	R-universe 0.0.6 R-CMD-check passing CRAN not published

<https://swissstatsr.org>



Collaboration: SwissStatsR.org

- The «Submission» page presents the phases of the formalized collaboration workflow.
- **Phase 1: Your Code.** Any R script related to Swiss statistics can be submitted.
- **Phase 2: GitHub Private.** Once accepted, peer review starts on GitHub.
- **Phase 3: GitHub Public.** Only open source R packages can be released publicly.
- **Phase 4: R-Universe.** R packages should pass R-CMD-Check and be well documented.
- **Phase 5: SwissStats.org.** Possible promotion on Substack blog.

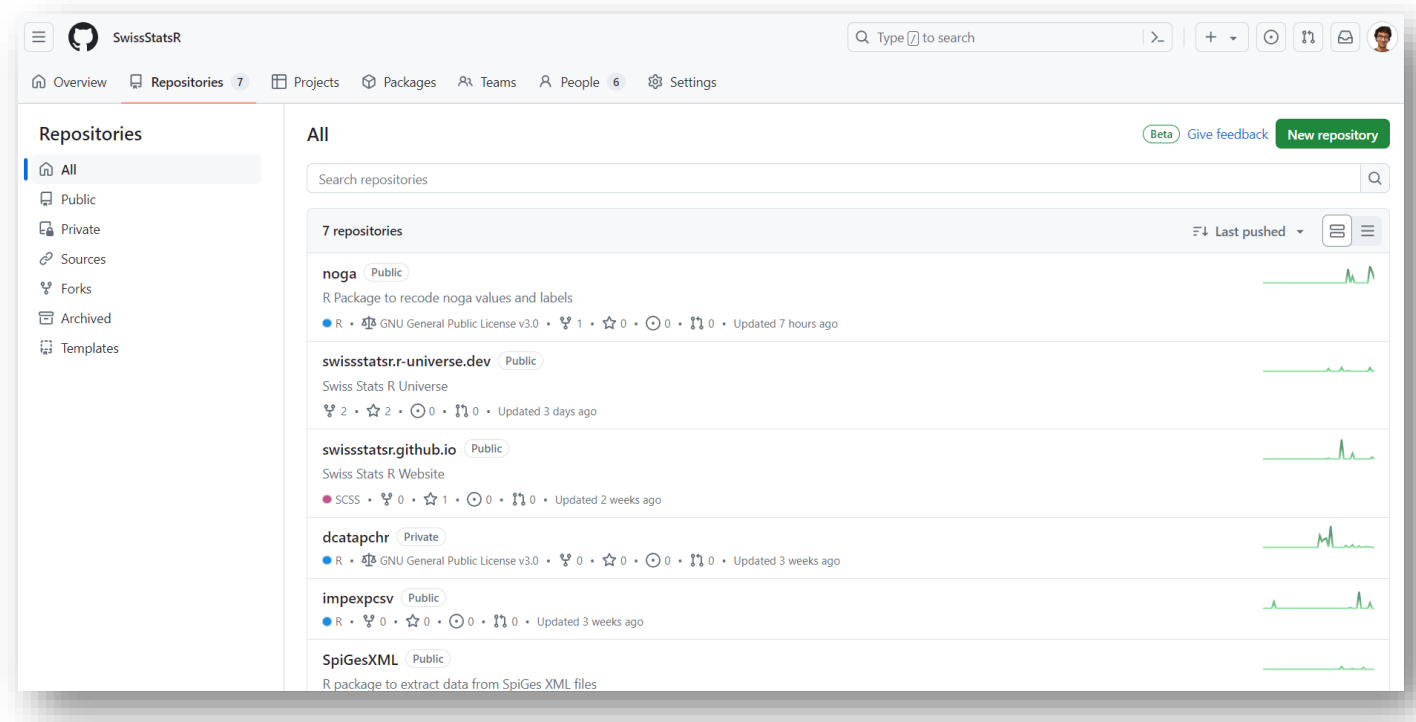


<https://swissstatsr.org/collaboration>



SwissStatsR GitHub organization

- R packages can be hosted on SwissStatsR GitHub organization, but **it is optional**.
- R packages can be hosted on any GitHub repository.
- It is **mandatory** to add R-CMD-Check GitHub Actions to the repository:
 - `usethis::use_github_action()`
- R packages should have an open source license such as MIT, GPL or LGPL:
 - `usethis::use_gpl3_license()`

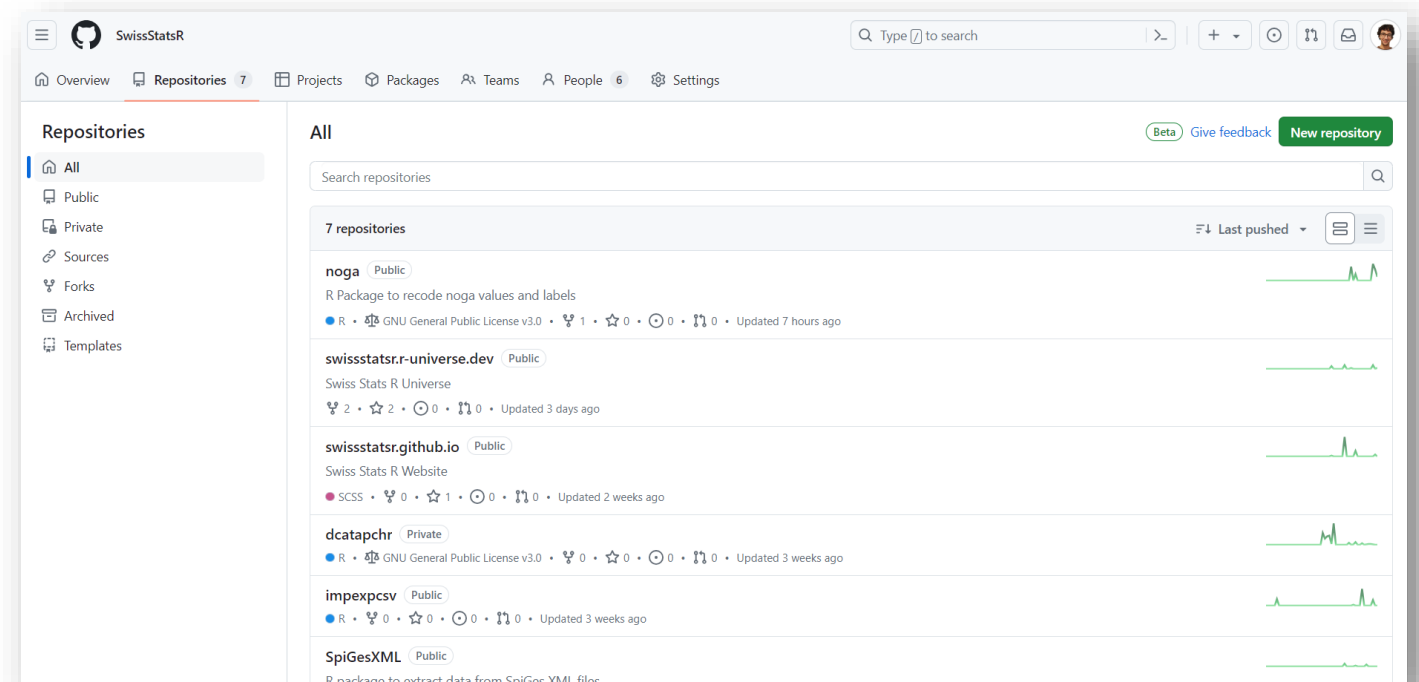


<https://github.com/SwissStatsR>



The SwissStatsR Universe

- The R-Universe service, provided by rOpenSci, builds the R packages in all available platforms (Windows, Mac, Linux).
- The R-Universe tests if R packages can be correctly build in all machines.
- It centralizes metadata information about the R packages, which are pulled by SwissStatsR Quarto website.
- SwissStatsR R packages can be installed easily with `install.packages()`:
 - `install.packages("BFS", repos = "https://swissstatsr.r-universe.dev")`



<https://swissstatsr.r-universe.dev/>



Two years ago...

Towards a National Code Base

...building a new universe!

The R-universe system

The R-universe platform is a new umbrella project under which we experiment with various new ideas for improving publication and discovery of research software in R.

The system consists of several components and subprojects, that come together on the front-end dashboard on <https://r-universe.dev>.

Currently serving 3778 packages and 3713 articles by 1066 maintainers and 5729 contributors.

Not sure what to search for? Why not try: *maps, bayesian, ecology, climate, genome, gam, spatial, database, pdf, shiny, rstudio, machine learning, prediction, birds, fish, sports, ...* (more popular topics)

giscoR: Download Map Data from GISCO API - Eurostat

Tools to download data from the GISCO (Geographic Information System of the Commission) Eurostat database <<https://ec.europa.eu/eurostat/web/gisco>>. Global and European map data available. This package is in no way officially related to or endorsed by Eurostat.

Last updated 7 days ago

35 stars 3.56 score 20 dependencies 4 dependents

sotkanet: Sotkanet Open Data Access and Analysis

Access statistical information on welfare and health in Finland from the Sotkanet open data portal <<https://sotkanet.fi/sotkanet/fi/index>>.

Last updated 2 months ago

5 stars 1.38 score 8 dependencies

rOpenGov

Open government data analytics with R

[r-universe/ropengov](https://r-universe.dev/ropengov)

<http://ropengov.org>

Links to ropengov

Leo Lahti

Diego Hernangómez

<https://r-universe.dev/search/> <https://ropensci.org/r-universe/>

adminR Spring Meetup 2022 in Aarau



Collaboration: CORSTAT «R» working group

Interested in R code, reviewer work / further information – please reach out:



Jan Wunder
Dr. Sc. ETH
Stellenleiter

Fachstelle für Statistik

Davidstrasse 35
9001 St.Gallen

📞 +41 58 229 87 35

✉ jan.wunder@sg.ch



Félix Luginbühl
M.A.

Wissenschaftlicher Mitarbeiter

Fachstelle für Statistik

Davidstrasse 35
9001 St.Gallen

📞 +41 58 229 65 01

✉ felix.luginbuehl@sg.ch

Stat. Amt Kanton Zürich

Andrea.Schnell@statistik.ji.zh.ch
Philipp.Bosch@statistik.ji.zh.ch



Thank you for your attention! Questions?

Interested in R code, reviewer work / further information – please reach out:



Jan Wunder
Dr. Sc. ETH
Stellenleiter

Fachstelle für Statistik

Davidstrasse 35
9001 St.Gallen

📞 +41 58 229 87 35

✉ jan.wunder@sg.ch



Félix Luginbühl
M.A.

Wissenschaftlicher Mitarbeiter

Fachstelle für Statistik

Davidstrasse 35
9001 St.Gallen

📞 +41 58 229 65 01

✉ felix.luginbuehl@sg.ch

