Lucerne University of
Applied Sciences and Arts

HOCHSCHULE
LUZERN

Wirtschaft

CARF Luzern 2020
Controlling.Accounting.Risiko.Finanzen.

Konferenzband

Konferenz Homepage: www.hslu.ch/carf

Kronos: A Visual Analytics Controlling Tool

Type of Contribution: Research Paper

Dr. Roland Schregle
Lucerne School of Engineering and Architecture, Institute of Civil Engineering, Horw, Switzerland
E-Mail: roland.schregle@hslu.ch

Prof. Dr. Stephen Wittkopf
Lucerne School of Engineering and Architecture, Horw, Switzerland
E-Mail: stephen.wittkopf@hslu.ch

Abstract

This paper presents *Kronos*, a novel data aggregation and visualisation tool currently under development at the Lucerne University of Applied Sciences and Arts. Its development was motivated by the need to obtain an overview of selected key performance indicators (KPI) to support decision making, control of measures, and internal communication.

Previously, the necessary data was only available in scattered form from disparate sources and formats. *Kronos* aggregates the data imported from the various sources and visualises it as temporally resolved KPI in a unified manner, across all hierarchical organisational units (OU), that is intuitive to directorial decision makers.

Primary KPI include visualisations of the manpower an OU dedicates to performance areas such as teaching, research, and consultancy / services. Secondary KPI include overheads, compliance with SBFI 50-20-20 rule, time balances, demographics, birthdays and terms of contracts, as well as tabular listings with additional head-counts and full-time-equivalents, sorted by all staff categories.

1 Introduction

Universities of all types are faced with growth, complex environments, and increasing control by external institutions (Bundesrat, 2020). Key performance indicators and reporting systems are management controlling instruments providing an important basis by creating transparency (Azma, 2010) (Tasić, 2017). They are used to analyse the current situation, as a basis for decision making, to monitor the effectiveness of measures and for internal and external communication.

This research paper presents the prototype *Kronos*, an example of a management system developed within and for the organisational unit "Fachbereich Bau" of the School of Engineering and Architecture of the Lucerne University of Applied Sciences and Arts. It serves as an experimental framework to evaluate various KPI within the department. It also contributes towards the "Digitale Agenda 2030", a strategic project of the University, by using digital tools for visual analytics to aggregate information from data scattered across specialist databases.

2 *Kronos* Overview

The primary design criteria for *Kronos* were:

– Import and aggregation of heterogeneous data from various sources via separate Excel imports

– Platform-independent interactive visualisation in a web browser via HTML output

– Visualise all data within a fixed timespan of the current (reporting) year.

– Intuitive presentation of the data aggregated at various levels in the organisational hierarchy.

– Provide a controlling summary tailoured to an academic environment, on the basis of personnel factors attributed to individual employees or groupings thereof, such as: working time, workload, contract duration, and demographics.

– A portable and extensible software framework using open-source libraries.

In addressing these criteria, the *Kronos* framework is implemented entirely in Python using the Pandas library (McKinney, 2010) (Pandas development team, 2020) for data aggregation and wrangling, and the Bokeh library (Bokeh development team, 2020) for interactive visualisation in HTML/Javascript. Choosing open-source software fulfils the design goal of sustainable development that is not bound to a specific and potentially obsolescent platform.

Pandas is a powerful Python library that has become the *de facto* standard in data science applications. Its interface with numerous other libraries has opened further applications in adjunct fields such as visual analytics and machine learning, providing the foundation for software tools geared towards these tasks. Given its very large user community, Pandas was a natural choice in the context of a Python framework.

The Bokeh interactive visualisation library interfaces with Pandas and directly supports generation of plots from Pandas data types. Bokeh applications can run as a server backend to a Python web framework, providing dynamic visualisation updates. Alternatively, they may also run as standalone application, processing the data and generating interactive output via static HTML pages that may be distributed to decision makers within the organisation.

Figure 1 (see appendix "Figures and Captions") presents a schematic overview of the entire *Kronos* framework, with arrows denoting the data flow from its import (red arrows), through aggregation (blue arrows), and finally its visualisation. It consists of the following main components:

– Excel import from the various sources (SAP Protime, HR-Tool, and Controlling). Some imports are subject to conditioning, categorisation and automatic cleaning within these modules. Some imports are merged from multiple exports from the same source attributed to different organisational units or timespans. All import and associated preprocessing routines reside in a dedicated Python module.

– The main data aggregation modules; this includes selective filtering and aggregation of imported data from different sources to generate the various visualisation types. The visualisations are generated in top-down fashion over the entire organisational hierarchy, with data correspondingly filtered and grouped for every OU and its employees at a given level within the hierarchy.

– The data visualisation modules; each module generates a specific visualisation type from data pre-filtered by the main module. The output is generated as HTML files and deposited in a tree of nested subdirectories for each OU, reflecting the organisational hierarchy. Plots for individual employees are generated in a separate subdirectory *0-Personnel* within their associated OU, with the numeric prefix ensuring it will be listed first in a file navigator. Similarly, aggregated plots for every staff category are generated in a sibling subdirectory *0-PersonnelGroup*. This hierarchy provides a mechanism for centralised deployment by granting directors of OUs access to only those directory subtrees under their management, but not its parents.

2.1 Data Import and Aggregation

Currently, two databases (SAP ProTime, HR) and two controlling Excel files (Kostenträgerplan, Organisationsstruktur) provide the main input data. ProTime contains the hourly reporting of each employee against all cost units and centres, the former being SAP accounts with incomes and expenses and the latter being accounts with expenses only such as overheads. HR contains the mapping of each employee to hierarchical OUs and adds demographic data. The Excel files map SAP account numbers to the four main performance mandates of Universities of Applied Sciences (research, Bachelor/Master teaching, continuing education, consultancy and services) and overhead costs incurred by administration, laboratory, facilities, infrastructure and selected other cost units/centres.

These input data are exported from SAP to spreadsheets using tailoured export templates to capture only the relevant data fields. The export frequency is typically monthly, with data from all previous months within the current reporting period being consolidated by mid-month for import and merging in *Kronos*. The other excel files need regular updates for new cost units/centres. The effort for providing this import data is less than half a day per month.

All data is imported into Pandas DataFrames from Excel files generated directly from the aforementioned sources. The Pandas DataFrame implements an associative table much like an SQL database, and provides the basic building blocks for all operations. Similarly to SQL, DataFrames may be joined via common keys, for example to pair an employee's time management and HR data, such as demographics and contract details, using the full name or staff ID as common key.

Of central importance to *Kronos'* data processing is its titular time management in order to analyse how staff invest their performance time. All times reported by staff are booked into dedicated cost centres associated with the university's mandates, but also overheads. Based on these cost centres, *Kronos* categorises the performance reported by staff into one of the KPI-relevant performance (and some non-performance) areas summarised in Table 1.

The categorised performance time can be aggregated for staff associated with specific OUs within the organisation's hierarchy, as identified from their HR data. This categorised data acts as a KPI as it reveals how each member of staff or OU is aligned with respect to the university's mandates. An OU that emphasises teaching at the expense of research or vice versa, for example, may need to broaden its activities.

Mandates	**Overheads**	**Other**
Bachelor/Master teaching Continuing education Research and development Consultancy and services	Administration and management Laboratory Facilities and infrastructure	Internal projects Vacation Absences

Table 1: Performance area categories for staff working time reported by Kronos as KPI.

The point of reference of all of *Kronos'* output is the reporting date; this is the most recent date extracted from the staff time management data after import. Several visualisations relate to this date on their temporal axes. For example, the staff birthday and contract plots are expressed in months from the reporting date to indicate the urgency for action based on the most recent data. Staff ages are updated relative to the reporting date, since the imported HR data may be up to a year old. All data is temporally clipped to lie within the reporting year, with the objective being to present projected results over the course of the current year.

Where there is the need to obfuscate uniquely identifiable personal data, e.g. for public dissemination or demonstration, *Kronos* provides an optional anonymous mode that generates fabricated names using the Faker Python library (Faraglia, 2014). In this mode, all occurrences of a name in every DataFrame are consistently replaced by a generated name at the time of import, via a lookup table. As Faker supports localisation, some of these generated names are almost quintessentially German.

2.2 Visualisation Types

Kronos generates various types of visualisations as individual Bokeh plots accumulated in one large layout to provide controlling data at a glance. These visualisations can be grouped into the following categories and addressed queries:

- Staff time management and performance areas:
  - How much time do staff spend in each of the relevant performance mandates per month and over the entire year?
  - How long are they absent or on vacation?
- Staff contracts, time balance and workloads:
  - Are there any members of staff with gross over/undertime or unused vacation?
  - How long do current staff contracts run and how soon are they due for renewal?
  - When are staff due for retirement?
  - Which staff qualify for SBFI's *50-20-20 rule* for bonus federal funding?
- Staff demographics:
  - What is the staff age/gender distribution?
  - What percentage of women make up the staff body?
  - How diverse are the staff? Which nationalities are represented, and by what percentage?

Each visualisation is discussed in the following subsections using sample plots shown in the appendix. Note that all sample output has been anonymised for exposition.

2.2.1 Staff Time Management and Performance mandates

These visualisations present an overview of how staff spend their cumulative reported performance time, notably in which performance areas These visualisations are comprised of a stacked bar plot of monthly and annual hours keyed by performance areas (see Figure 2), and a compass or “radar” plot quantifying the relative hours invested in each of the 8 main performance areas (see Figure 3). A hover tool enables the controller to interactively and quantitatively inspect the cumulative hours in each bar segment associated with a particular performance area.

The radar plot is of particular interest and warrants a more detailed explanation. It provides insight into an OU’s “orientation” in terms of its performance fields. The radar plot quantifies the hours invested in performance are on an arbitrary scale normalised at 100% at the circumference, with concentric lines extending in 10% intervals from the origin. The performance areas are positioned as compass points along the circumference of the radar “screen”, such the four main mandates (teaching, continuing education, research, and consulting/services) lie in the upper semicircle, while overheads (administration, laboratory, facilities/infrastructure, and miscellany) lie in the lower semicircle. Since the activities in the former group generate revenue, while those in the latter generate expenses, a blue radar patch biased toward the bottom would indicate excessive overheads, and therefore a problem. Ideally, an OU should exhibit equally strong performance in the mandated areas represented in the upper semicircle.

2.2.2 Terms of Contracts, Time Balance and Workloads

These visualisations provide insight into the terms under which staff are employed, and how staff balance their working times against their vacation. In addition, a plot specific to Swiss applied universities provides insight into each staff member’s qualification for SBFI’s 50-20-20 rule to receive bonus federal funding. All requisite data for these plots is extracted from the imported HR records and time balance.

Staff contracts are depicted as a scatter-plot (see Figure 4), where each point represents a member of staff, ordered vertically according to their category. The remaining time until contract expiry from the reporting date can be read off horizontally from a logarithmic axis in months; coloured regions along this axis ranging from red to green indicate the urgency for timely action (e.g. contract renewal) from the time of reporting. A hovertool provides details of individual contracts, including date of entry, contract type (permanent or temporary) and expiry date. In case of permanent contracts, the expiry date coincides with the staff member’s regular retirement date on his/her 65th birthday.

Staff time balance is similarly depicted as a scatter-plot (see Figure 5), where the horizontal axis indicates each staff member’s remaining vacation (in hours), while the vertical axis indicates his/her working time balance centred around the dashed zero axis. The coloured regions indicate the urgency to compensate for over/undertime and use up remaining vacation. This is particularly relevant towards the end of the year, as only a limited balance may be carried over into the next. Ideally, all staff should lie within the green region at year’s end. Note that some staff, notably associate lecturers, are not directly employed, and therefore do not report their vacation; these are visible along the vertical axis on the extreme left of the plot.

Each staff member’s qualification for the SBFI 50-20-20 rule is again depicted as a scatter-plot (see Figure 6). According to this rule, SBFI grants additional federal funding for all staff who dedicate at least 20% of their workload in the mandated performance areas of teaching and research, and are employed full-time or part-time (as is quite common in Switzerland) with an overall workload of at least 50%. A specific mapping of cost centres to the two performance areas mandated by SBFI is imported for this plot. Kronos uses the effective workload accumulated from each staff member’s reported performance time, rather than the nominal workload as specified in his/her contract; this value is then scaled for the reported time period if under 12 months to obtain a projected annual workload. The coloured regions in the visualisation indicate the qualification limits, where green indicates qualification in both teaching and research, yellow indicates qualification in only one area, and red indicates qualification in neither area. Note that employees with workloads under 50% are omitted from the plot, as they are categorically disqualified.

2.2.3 Staff Demographics

These visualisations provide insight into the demographics of staff with respect to age, gender, and nationality. In particular, these visualisations are indicators of diversity and gender equality. The necessary data is extracted exclusively from each staff member's imported HR record.

The staff age/gender distribution, or "pyramid" plot (see Figure 7) plots the number of staff (horizontal axis) within various age intervals (vertical axis), with male and female staff counts represented by coloured bars extending from the centreline towards the left and right, respectively. This plot is particularly useful to assess the percentage of women in the workforce in order to meet potential target quotas.

The staff birthday scatter-plot (see Figure 8) informs of upcoming birthdays from the reporting date, as indicated in months along the horizontal axis. As in the previous visualisation, staff are grouped vertically in age intervals. Coloured regions along the temporal axis indicate the immediacy of upcoming birthdays using a "traffic light" scheme. The hovertool provides details of each staff member's birthday and age at the time of reporting (which may be updated from potentially stale HR data). Milestone birthdays (corresponding to ages that are multiples of ten) are highlighted in enlarged blue points, if congratulations are due on behalf of the directorate.

Staff nationalities are visualised as a ring plot (see Figure 9) depicting the percentage of staff representing each nation. The nationalities are represented as colour-coded angular wedges and referenced in the legend by their 2-letter country code for brevity. The hovertool lists the names of all staff belonging to a nationality, and the percentage of the workforce they represent.

2.3 Current Status

Kronos is currently a fully functioning prototype demonstrator used in the OU "Fachbereich Bau" only. Current outputs include KPI for the whole year 2019, and for the first 4 and 6 months of 2021, which are securely shared with the heads of institutes of this OU via SWITCHdrive. All personal information has been anonymised for reasons of data protection. A short video introduction is available here: https://drive.switch.ch/index.php/s/tM8MqMcCOJ5qmVp.

3 Outlook

The University's new strategy for the period 2020-2023 sets forth both qualitative and quantitative KPI in addition to the ongoing digitalisation (Hochschule Luzern, 2020). *Kronos'* visual analytics controlling shall contribute towards a transparent and easy access to the KPI's up to date status for decision making, monitoring of measures and for internal and external communication.

4 Acknowledgements

The authors would like to thank staff from controlling and HR for their contributions in data/information delivery and aggregation, namely Stefan Bachmann, Raja Koller, Elvira Meyrat, Andreas Amrein, and heads of institutes and other organisational units, namely Albin Kenel, Fabian Ille, Mirko Kleingries, Dave Degonda and others for feedback during the development.

References

Azma, F (2010): Qualitative Indicators for the Evaluation of Universities Performance. Procedia – Social and Behavioural Sciences, 2(2): 5408–5411. DOI 10.1016/j.sbspro.2010.03.882.

Bokeh development team (2020): Bokeh: Python library for interactive visualization. https://bokeh.org. Accessed 23.07.2020.

Bundesrat (1. April 2020): Verordnung zum Hochschulförderungs- und –koordinationsgesetz (HFKG). https://www.admin.ch/opc/de/classified-compilation/20161646/index.html. Accessed 27.07.2020.

Faraglia, D (2014): Faker fake data generator Python package. https://faker.readthedocs.io/en/stable/. Accessed 23.07.2020.

Hochschule Luzern (2020): Mission, Vision, Werterahmen, Strategie. https://www.hslu.ch/de-ch/hochschule-luzern/ueber-uns/portraet/mission-vision-werterahmen-strategie/. Accessed 27.07.2020.

Hochschule Luzern (2020): Strategie 2020-2023 – Dachstrategie im Detail sowie Teilstrategien der Departemente und Services (Intern): https://hsluzern.sharepoint.com/sites/rs_rk_informationen-zur-strategie/SitePages/Dachstrategie-im-Detail.aspx. Accessed 27.07.2020.

McKinney, W (2010): Data Structures for Statistical Computing in Python. In: van der Walt, S; Millman, J (Ed.), *Proceedings of the 9th Python in Science Conference*: 56–61. DOI 10.25080/Majora-92bf1922-00a.

Pandas development team (2020): The Pandas Python Data Analysis Library. https://doi.org/10.5281/zenodo.3509134. Accessed 23.07.2020.

Tasić, N.; Delić, M; Maksimović, R; Lalić, B; Ćukušić, M (2017): Selecting Key Performance Indicators in Universities – Academic Perspective. In: Anderla, A (Ed.), *17th International Scientific Conference on Industrial Systems (IS'17)*, Novi Sad, Serbia. http://www.iim.ftn.uns.ac.rs/is17. Accessed 28.07.2020.

Appendix: *Kronos* Figures and Captions

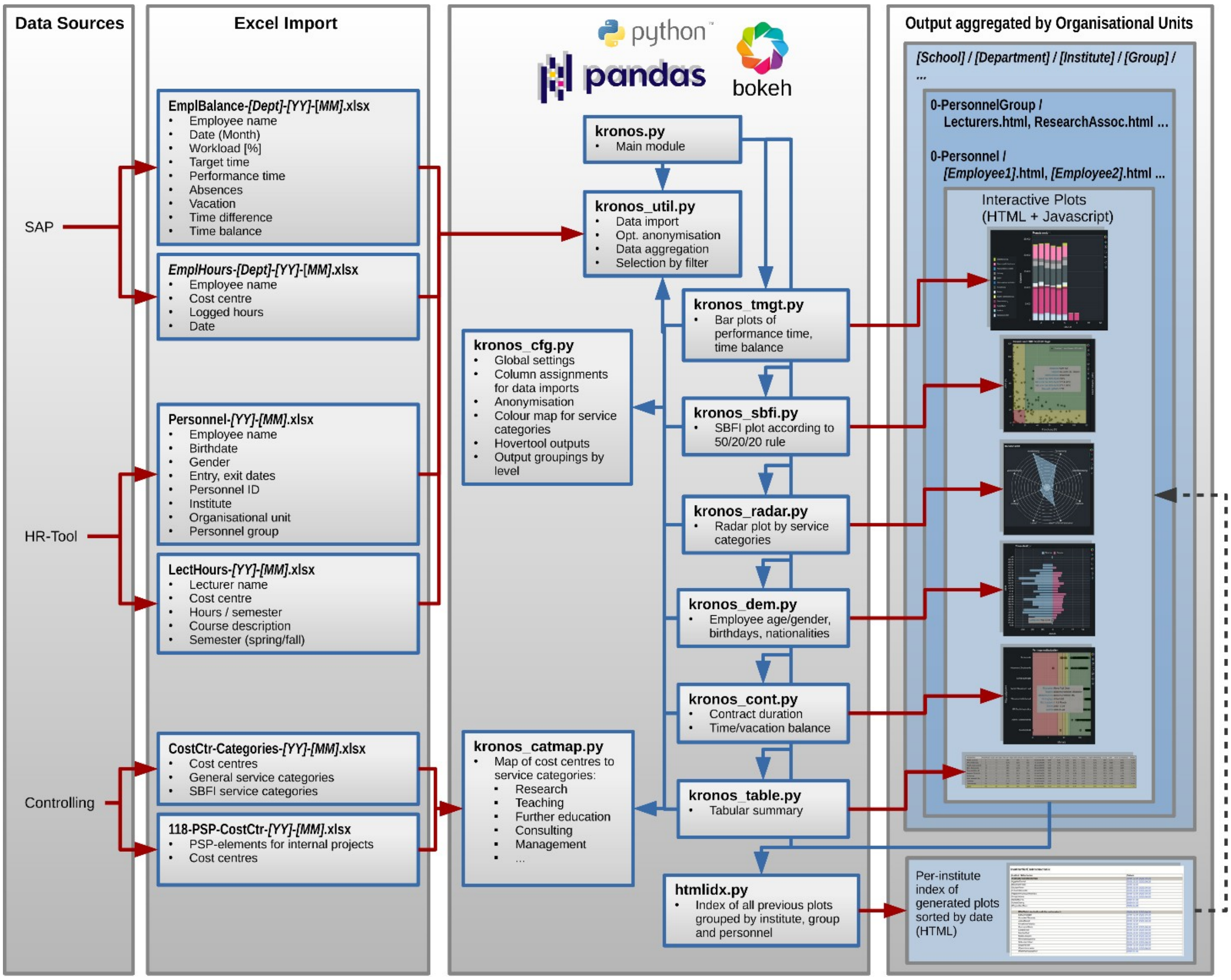

Figure 1: *Kronos* flowchart depicting constituent software modules and data import / aggregation / visualisation paths.

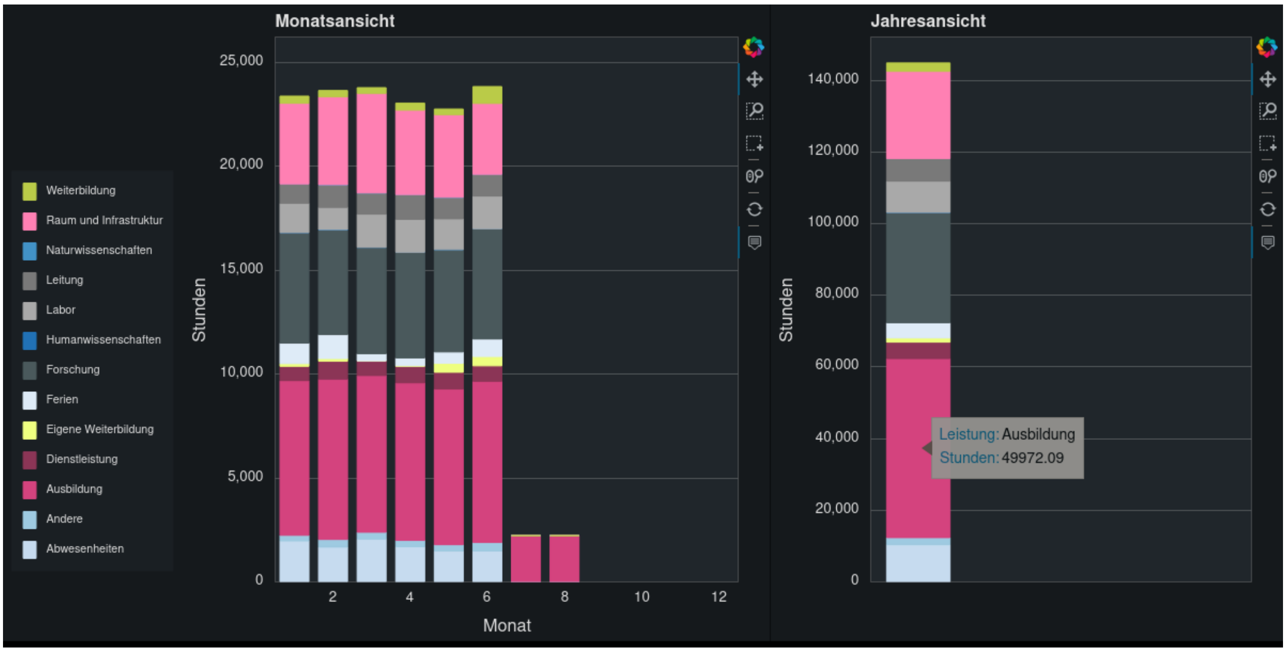

Figure 2: Sample time management output. Hours logged by staff are presented on a monthly (left) and annual (right) basis for the duration of the (in this case partial) reporting year. Reported hours are attributed to one of several performance areas as keyed by colour, and plotted as stacked bars.

Figure 3: Radar plot categorising staff cumulative performance time into 8 main areas, consisting of the four mandates in the upper semicircle, and overheads in the lower. In this example, most of the performance time was invested in (in decreasing order) teaching, research, and infrastructure.

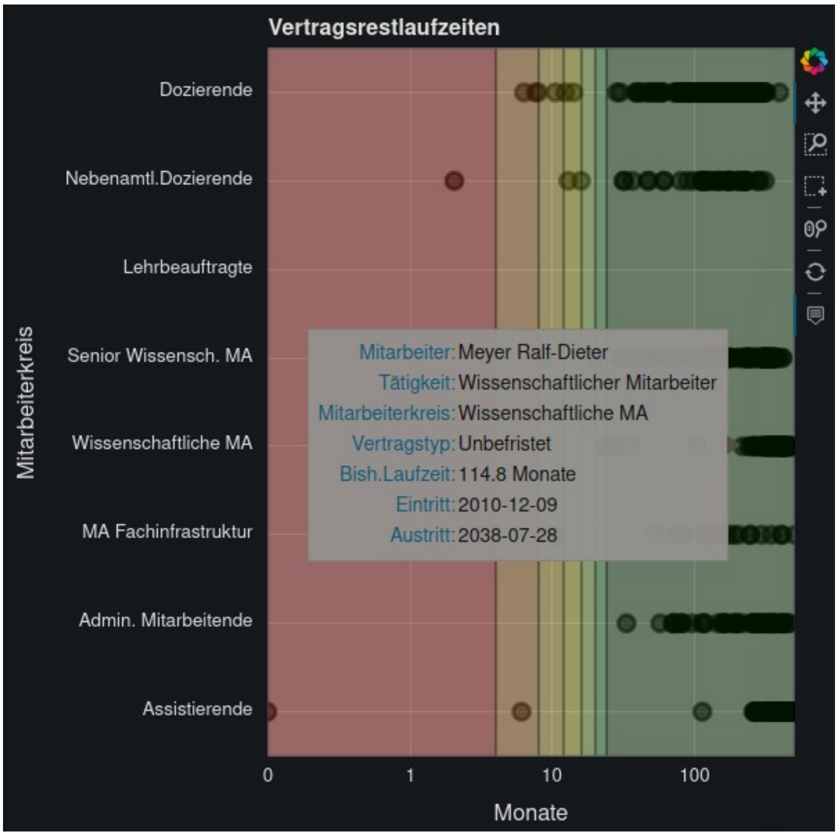

Figure 4: Staff contract plot. The coloured regions indicate the remaining duration of contracts (expressed in months on a logarithmic scale). Each member of staff is grouped vertically according to their category. The hover tool highlights a (rather lucky) senior research associate with a permanent contract since 2010, which is due to expire in 2038 when he retires.

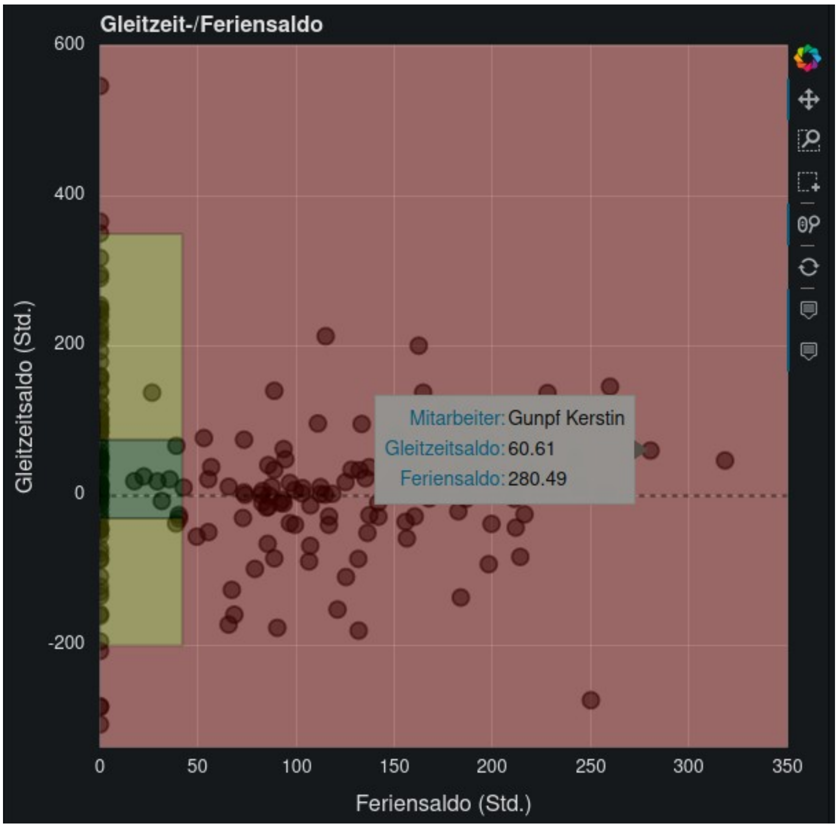

Figure 5: Staff time balance plot. This depicts each employee's vacation balance on the horizontal axis vs. his/her over/undertime on the vertical axis. Both balances are expressed in hours. The coloured regions indicate the urgency to equalise over/undertime resp. use up remaining vacation towards year's end. In the highlighted example, an overworked Ms. Gumpf has accumulated 60 hours of overtime and hardly taken any vacation.

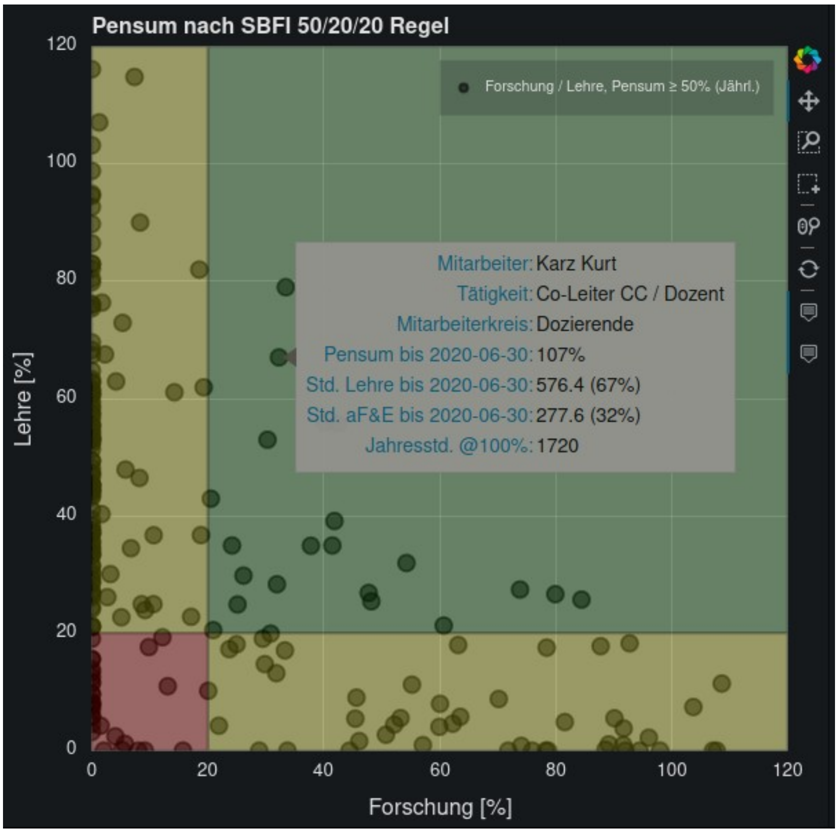

Figure 6: Staff qualification for SBFI's 50-20-20 rule. This plots the percentage of performance time reported staff relevant for the two primary mandates, research (horizontal axis) and teaching (vertical axis). A member of staff qualifies for the 50-20-20 rule if both percentages exceed 20%, and his/her effective workload is at least 50%; this corresponds to the green region.

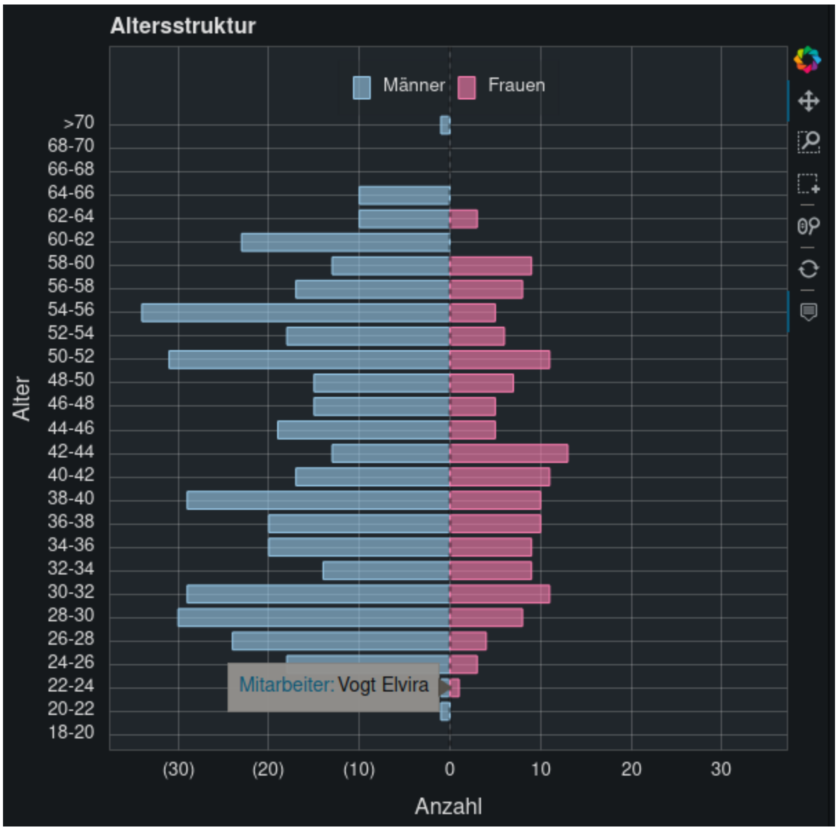

Figure 7: Staff age/gender distribution as "pyramid" plot. This organisational unit is male-dominated (note the predominant blue bars), with a relatively uniform age distribution into the mid 60s along the vertical axis. The youngest female employee (age 22–24) is highlighted. A single male employee well past retirement age (>70) is also shown.

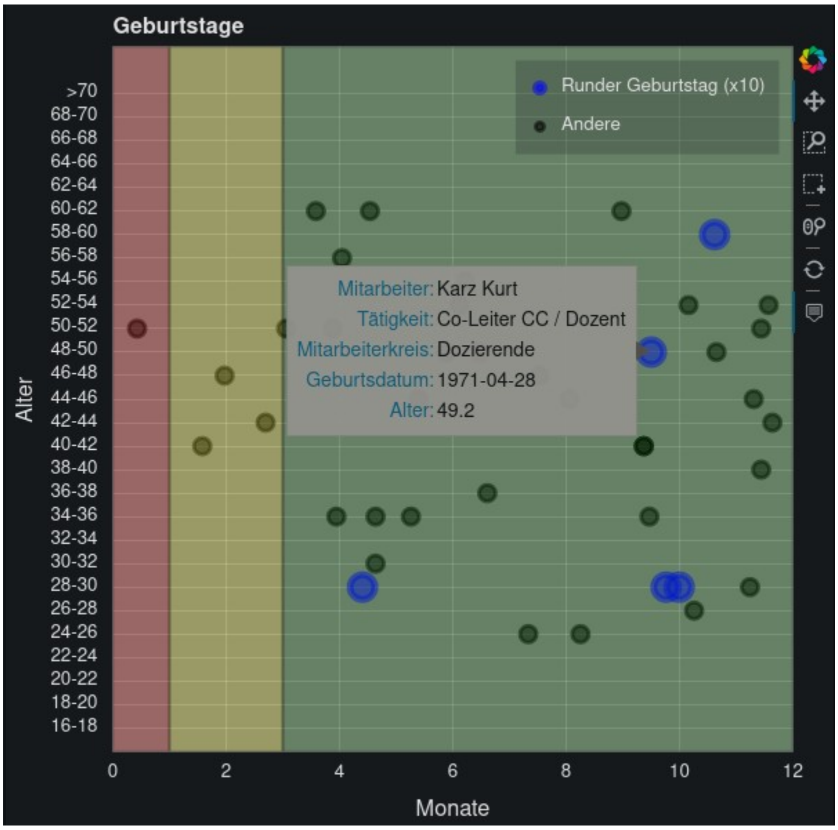

Figure 8: Plot of upcoming staff birthdays within the next 12 months (horizontal axis) from the date of reporting, encoded into red/yellow/green regions according to timeliness. Similarly to Figure 7, the vertical axis groups employees by age ranges. Milestone birthdays (multiples of 10) are emphasised as larger blue plots. In the highlighted example, lecturer Prof. Karz turns 50 in 9.6 months, so congrats are in order.

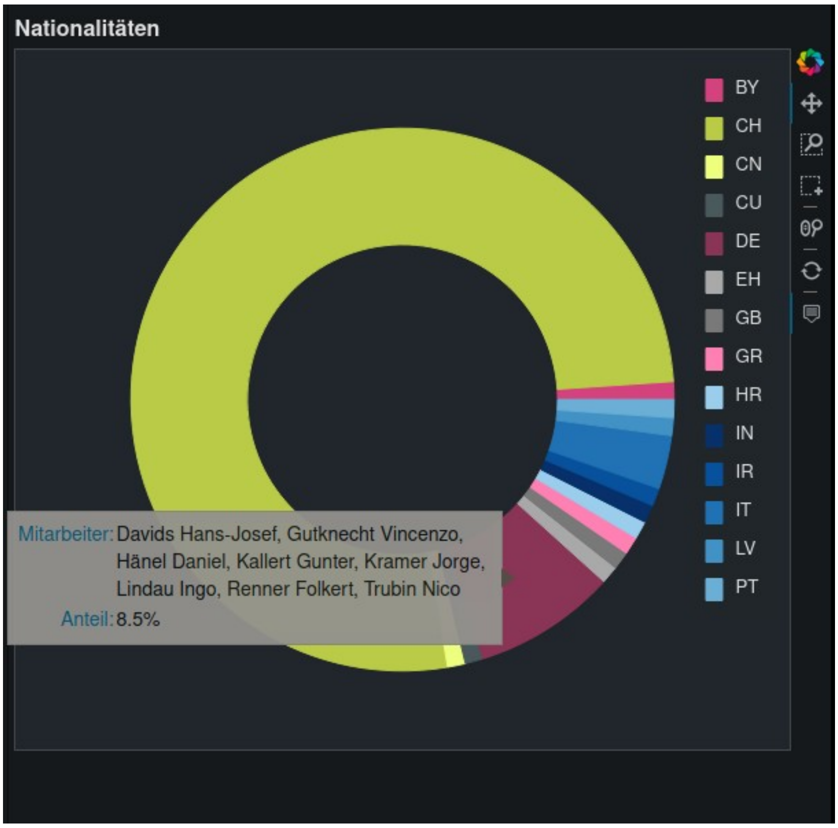

Figure 9: Staff nationalities as ring/donut plot. Each represented nationality is keyed by colour and country code. The hovertool lists the corresponding staff names and the percentage of the workforce represented by them. In this example, the three major nationalities are 76.6% Swiss, 8.5% German (highlighted in hovertool), and 3.2% Italian.