Two-Dimensional Mapping of University Profiles in Research.

Joel Emanuel Fuchs[1] and Thomas Heinze[2]

[1] *jfuchs@uni-wuppertal.de*
University of Wuppertal, Institute of Sociology, Gauss-Str 20, D-42119 Wuppertal (Germany)

[2] *theinze@uni-wuppertal.de*
University of Wuppertal, Institute of Sociology & Interdisciplinary Center for Science and Technology Studies (IZWT), Gauss-Str 20, D-42119 Wuppertal (Germany)

Abstract
This paper presents a two-dimensional graphical mapping for institutional profiles of higher education entities (such as colleges and universities) in research, teaching, technology transfer, or internationalization. For the sake of simplicity, our illustrations focus on research profiles only. The new graphs embrace both the organizational field level and the organizational level. We illustrate how the new graphical representation improves existing ones, using examples from the German public university system.

Introduction
Mapping profiles in higher education, especially with regard to research and teaching, has proliferated in recent years. Such mapping is done in the context of an increased attention to large-scale visualization of science and technology (Borner, Bueckle, & Ginda, 2019; Fortunato et al., 2018). One focus in this literature have been comparisons at the system level (van Vught, 2009). An example is Huisman, Lepori, Seeber, Frölich, and Scordato (2015) who classify 24 national higher education systems in Europe with regard to their degree of horizontal differentiation in research, teaching, technology transfer, and internationalization. Another example is Harzing and Giroud (2014) who identify top-3/bottom-3 research areas in 34 national higher education systems using the *Revealed Comparative Advantage (RCA)* measure. At the country level, Teixeira, Rocha, Biscaia, and Cardoso (2012) found for Portugal that private higher education entities have research profiles complementary to those of public entities, using the *Relative Specialization Index (RSI)*.
A second focus in the literature is on organizational fields in higher education (Brint et al., 2011; Ruef & Nag, 2015). A well-known comprehensive mapping of bibliometric profiles has been done for the Nordic universities (Piro et al. 2017; 2014). Here, the *RSI* is used for the comparison of field-related publication and citation percentages with respective global field percentages, highlighting those colleges and universities with either below-average, average, or above-average contributions. In addition, Bonaccorsi, Colombo, Guerini, and Rossi-Lamastra (2013), by means of the *Activity Index (AI)*, show for Italy that universities specialized in applied fields and engineering have a positive impact on start-ups in their region, especially in the service industries. In contrast, universities with a profile in basic science fields are related to a greater number of start-ups in manufacturing. More recently, teaching and research profiles of public universities in Germany were mapped, using both the *RSI* and a modified version of it, the *RESP* (details below), with longitudinal data on scientific staff, funding, bibliometric indicators, and student enrolment (Heinze, Tunger, Fuchs, Jappe, & Eberhardt, 2019).
Mapping institutional profiles typically involves graphical representations, often *heatmaps*, that are meant to display degrees of specialization. Consider the example in Figure 1. Such heatmaps depict specializations with a color-spectrum ranging from yellow to green (Piro et al. 2017; 2014). Although such heatmaps show whether a university has either a below-average or above-average profile in a research field, they do not display the size of the respective field in the university under consideration. In other words: although a university might be specialized in a particular field, such as Aarhus University in "humanities" (Figure 1), the heatmap does not provide information with respect to the size of the humanities compared to the 7 other fields in

Aarhus. Of course, this information could be provided in a separate table or graph, but this would make comparisons of large numbers of higher education entities rather complex.

University	Geosciences	Health Sciences	Humanities	Materials Science	Mathematics & Statistics	Physics	Psychology	Social Sciences
Aalborg University	1.00	0.95	1.54	0.92	1.23	0.81		1.13
Aarhus University	1.19	1.26	1.60	1.18	0.88	1.25	1.16	1.16
Copenhagen Business School			2.32					1.44
Roskilde University						1.11		1.07
Technical University of Denmark	1.33	1.10		1.34	1.64	1.64		1.71
University of Copenhagen	1.27	1.27	1.42	1.32	1.27	1.61	1.13	1.11
University of Southern Denmark	1.81	1.06	1.37		0.89	1.55	0.82	1.22
Aalborg University Hospitals		1.01						
Aarhus University Hospitals		1.66						
Copenhagen University Hospitals		1.25						
University of Southern Denmark Hospitals		1.08						

Figure 1. One-dimensional heatmap, using a yellow-green color spectrum
Source: RSI scores for citation rates (2011-2014), Piro et al. (2017, p. 82).

This paper provides a new graphical representation that depicts both the specialization of a given higher education entity in a particular academic discipline (relative to all other relevant higher education entities) and its size within that entity. This new representation displays two types of information simultaneously. First, it shows whether a college or university either has a below-average, average, or above-average profile in a given research field. Second, it shows whether the research field is big, mid-size, or small compared to all other fields within that higher education entity. We provide examples from the German public university system, illustrating how the new graphical representation improves the existing one.

Data and Method

Our analysis builds on a dataset of 68 public universities in Germany, with information on scientific staff, basic funding, grant funding, publications, citations, and enrolment available for the years 1992-2018 (for details see Heinze et al., 2019). Based on these data, we calculated and then visualized institutional profiles, using both the RSI and the RESP. Results are available on the following website: https://fachprofile.uni-wuppertal.de/en.html. Both RSI and RESP are based on the aforementioned *Activity Index* (Narin, Carpenter, & Woolf, 1987). The *AI* captures the extent to which certain entities are specialized in certain activities (Formula 1). *AI* values lower 1.0 indicate a negative specialization (below-average scores), *AI* values greater 1.0 indicate a positive specialization (above-average).

Formula 1: General formula of the Activity Index (AI)

$$AI_{ij} := \frac{N_{ij}/\sum_i N_{ij}}{\sum_j N_{ij}/\sum_{ij} N_{ij}}$$

The AI's value range of [0.0,+∞] lacks an upper limit. Therefore, interpreting the RSI is easier than the AI due to its symmetrical value range of [−1.0,+1.0]: values lower 0.0 indicate negative specialization; values greater 0.0 indicate positive specialization. We use a modified version of the RSI that was introduced by (Grupp 1994; 1998). Its value range is [-100.0, +100.0] with an expected value of zero (Formula 2). We call this index RESP (for "Index of Relative Specialization"). It is based on the *hyperbolic tangent.* Consequently, its curve is steeper and reaches the upper limits of its value range more quickly than RSI.

Formula 2: Relative Specialization (RESP)

$$\text{RESP} := 100\,\frac{\text{AI}^2 - 1}{\text{AI}^2 + 1}$$

Note: The subindices i and j of the AI are omitted for the sake of simplicity.

Consider Figure 2 as an example that displays RESP scores for Web of Science publications of the University of Aachen from 1995-2018. At least three results can be inferred from the heat-map. *First*, Aachen has an above-average publication profile in psychology in all years. *Second*, Aachen's publication profile in economics has shifted from above-average to below-average. *Third*, Aachen's publication profile in chemistry has changed from below-average to average. As mentioned above, Figure 2 does not display the size of disciplines at the University of Aachen: the color bars of psychology, economics, chemistry are all of the same size. However, in reality, the three disciplines have very different publication outputs – and this is true also for other indicators, such as professorial staff or funding. As shown in Figure 3, Aachen's largest discipline in terms of publications is physics/astronomy ("Physik, Astronomie"), its bar has maximum size. All other disciplines are measured relative to it. For example, chemistry is smaller than physics/ astronomy but larger than psychology or economics. In other words: specialization in psychology requires a much smaller number of publications compared to physics/astronomy. As Figure 3 shows, physics/astronomy has most publications in Aachen, but their number is not as high as to constitute a particularly strong specialization in this discipline: its scores are yellow, meaning Aachen's number of physics publications are on average, compared to all other (here: technical) universities with publications in this discipline. The different bar sizes of disciplines in Figure 3 are based on Formula 3 for a given year. A university's largest discipline has bar size =1, because the enumerator equals the denominator, whereas all other disciplines have a bar size relative to the largest discipline. In Aachen, as mentioned, the largest discipline with regard to publications is physics/astronomy. The three other above-mentioned disciplines have the following bar sizes [min, max] between 1995 and 2018: chemistry [0,367; 0,830]; psychology [0,024; 0,140]; economics [0,010; 0,032]. All annual values are placed in the middle of the bar, and adjacent bars are connected using a natural cubic spline interpolation which has the effect of smoothening the bar graph.

Formula 3: Bar size calculation in Figures 3-5

$$\text{Bar size} := \frac{\text{Publications of University i in Discipline j}}{max_j\ \text{(Publications of University i in Discipline j)}}$$

Results

Figure 3 constitutes an improvement compared to Figure 2 because it allows to simultaneously evaluate both disciplinary specialization (measured at the organizational field level) and size of disciplines (measured at the organizational level). What are possible conclusions? *First*, Figure 3 depicts disciplines that have been for a long time below-average (geosciences) or have become so (economics, mathematics). *Second*, these three disciplines are marginal in terms of overall publication output; here four disciplines dominate: physics/astronomy, chemistry, mechanical and process engineering, and biology. *Third*, Aachen's increasing specialization in chemistry has been the result of a considerable growth in publications, relative to the discipline with most publications over time (physics/astronomy).

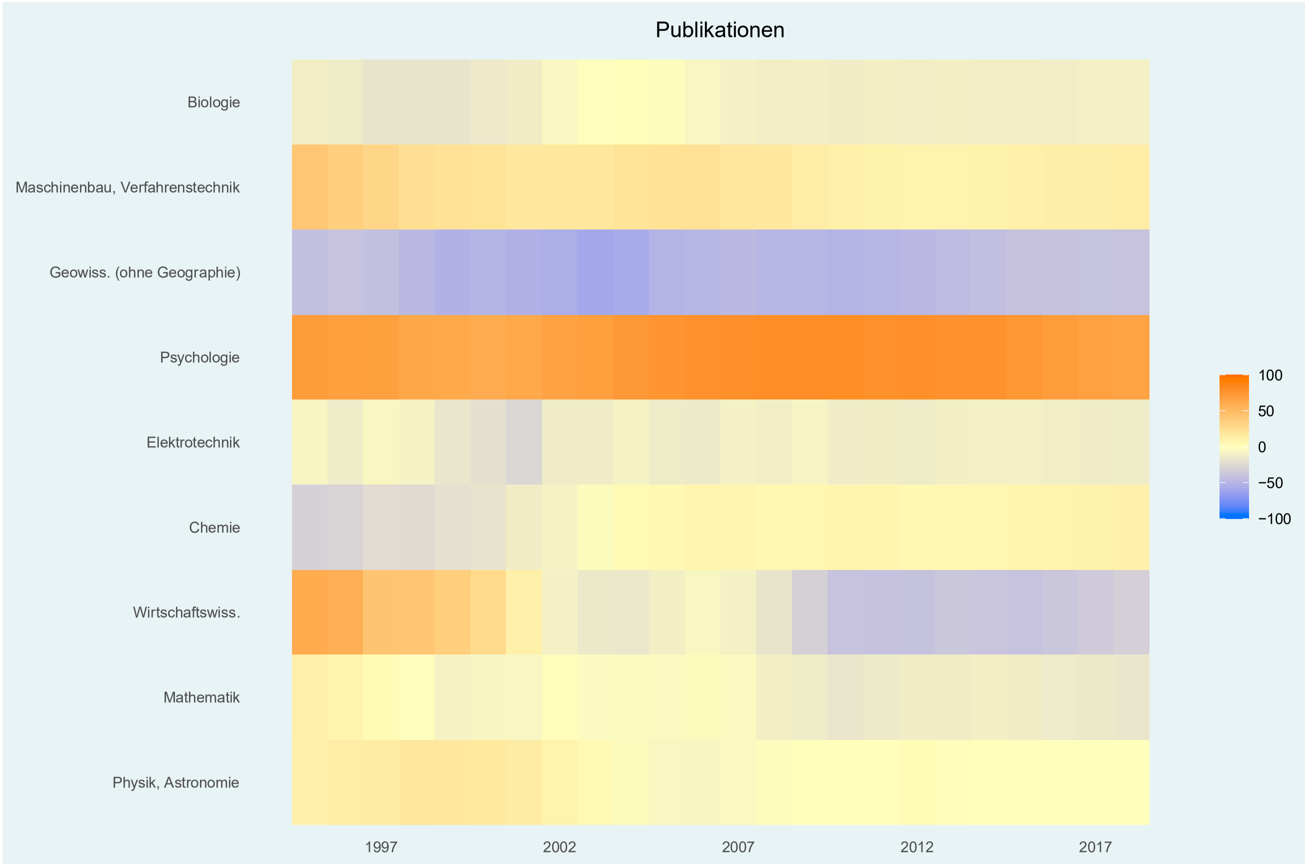

Figure 2. One-dimensional heatmap, Web of Science publications (Univ. of Aachen)
Source: RESP scores for WoS publications, available at: https://fachprofile.uni-wuppertal.de/en.html.

This interpretation for Aachen’s publication profile can be squared with other indicators, such as scientific staff. Consider Figure 4 that captures Aachen’s development of professorial staff. The overall picture looks quite different from that of publications in that almost all disciplines have grown relative to the discipline with most professors (mechanical/ process engineering). Two conclusions seem warranted regarding the comparison of Figures 3 and 4. *First*, in some disciplines, we observe double growth in input (professorial staff) and output (publication output) that leads to a stronger disciplinary profile (e.g., chemistry, physics/astronomy). *Second*, in other disciplines, input and output are decoupled, and there seems to be no coherent institutional profile (e.g., economics, geosciences, mathematics).

Another functionality of the new graphical representation is that it documents the connection between changes both in funding structure and research profiles. For example, consider the University of Wuppertal where grant funding has considerably changed since the late 1990s (Figure 5). In particular, the growth of externally funded research projects in mathematics moves the university’s research profile in this discipline from below-average (blue) to above-average (orange) within a few years. Note that mathematics is deeply orange although its grant funding is only a fraction of the largest discipline (electrical engineering). In contrast, chemistry takes almost the opposite route: within ten years only, the university switches its research profile in this discipline from above-average to below-average. Here, the decrease in grant funding occurs simultaneously: both on the organizational field level (other universities) and on the organizational level (grant funding, Wuppertal).

Discussion

We introduce a two-dimensional graphical mapping, with emphasis on research profiles of public universities. The new graphical representation can be applied to other dimensions, such as teaching, technology transfer, or internationalization as well. The key difference compared to existing *heatmaps* is that our new graphs capture comparisons on both the organizational field level (here: other universities) and the organizational level (university). In this way, we make a first step in better understanding the interplay between both levels. We are aware that our contribution is descriptive, and that further statistical analyses regarding the two levels and their function for both building and maintaining institutional profiles are necessary.

Regarding the graphical representation in Fig. 3–5, one might wonder about other technical possibilities, such as Sankey diagrams, stack area charts, or stream graphs (Wickham, 2016). Although it would be possible, for example, to stack the heat bars, and thus arrive at stack area charts, we decided not to do this. First, our prime motivation is to improve existing (and widely used) heat maps (as discussed above). Therefore, we used existing heat maps available for German public universities (https://fachprofile.uni-wuppertal.de). By adding one additional slice of information (here: the share of a field in comparison to all fields at one university), we wanted the new graphs to be as similar as possible to the existing ones, resulting in variable heights of the heat bars without stacking. Second, when using stack area charts or stream graphs, every research field would get their own color (so fields can be identified). But this is not our aim. In our graphs, colors represent RESP values. Consequently, different fields are displayed by the same color. In contrast, stalking the field would lead to a much more complex color spectrum in the graph, and thus counteract our main purpose: the production of intuitively understandable graphical representations that capture both the field and organizational levels.

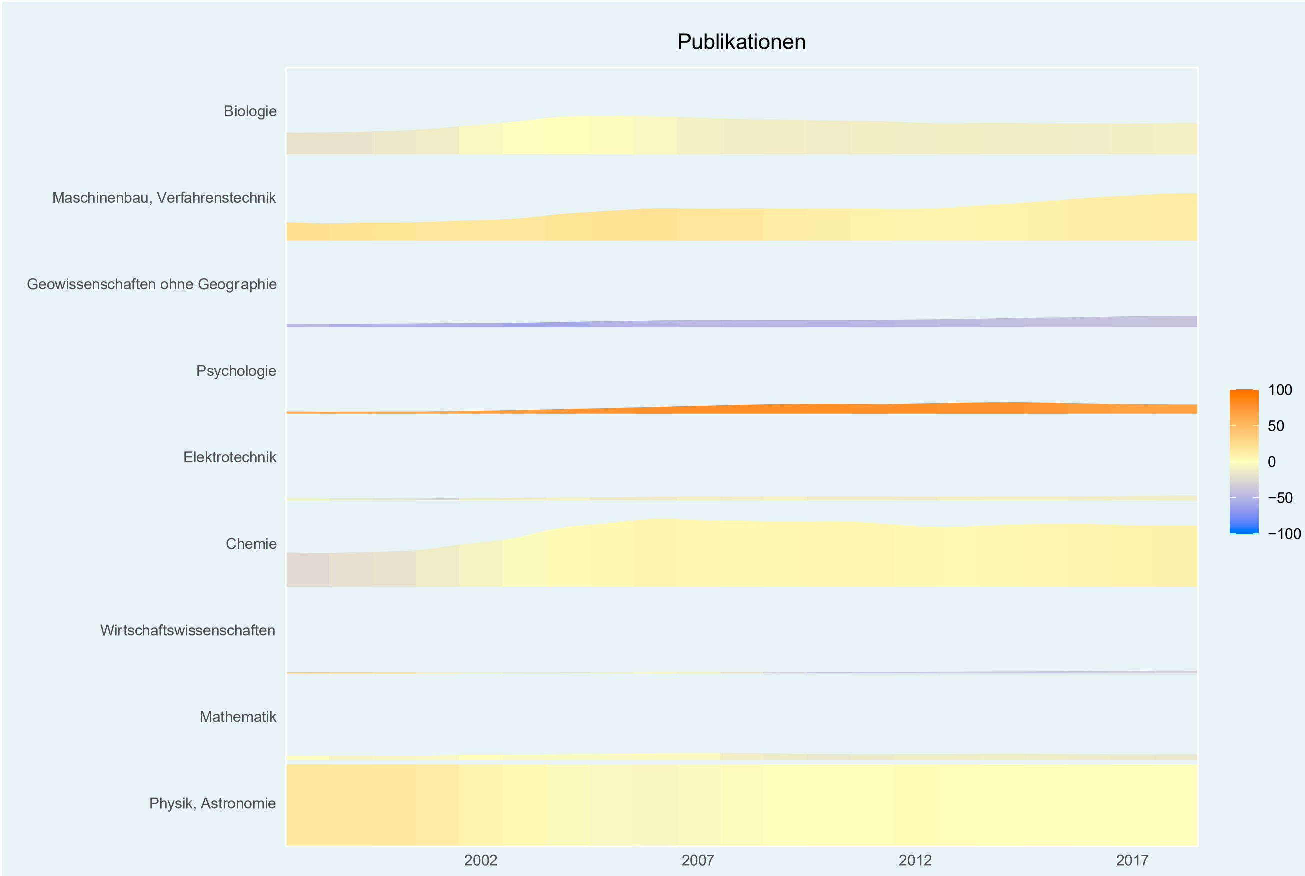

Figure 3. Two-dimensional heatmap, Web of Science publications (Univ. of Aachen)
Source: RESP scores for WoS publications plus respective size of disciplines.

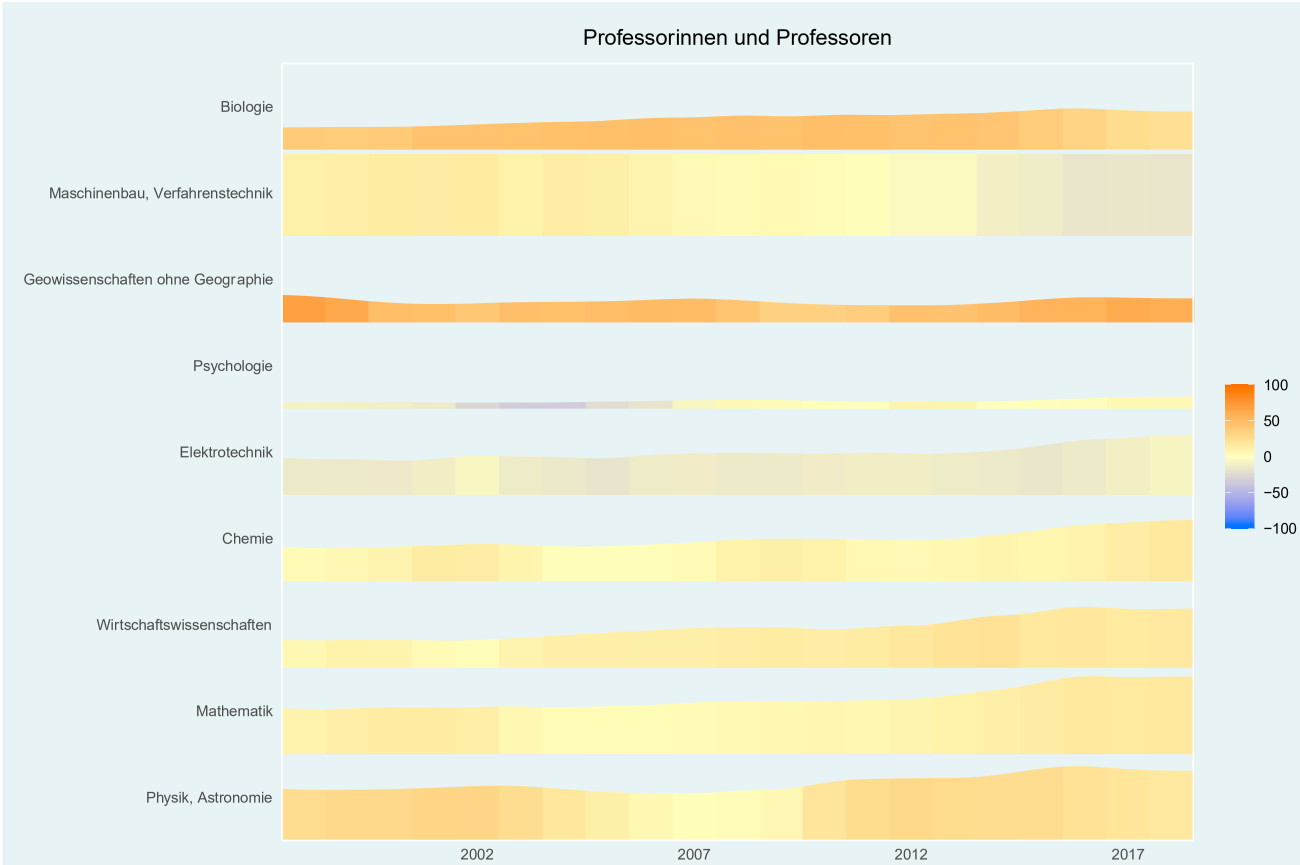

Figure 4. Two-dimensional heatmap, professorial staff (Univ. of Aachen)
Source: RESP scores for professorial staff plus respective size of disciplines.

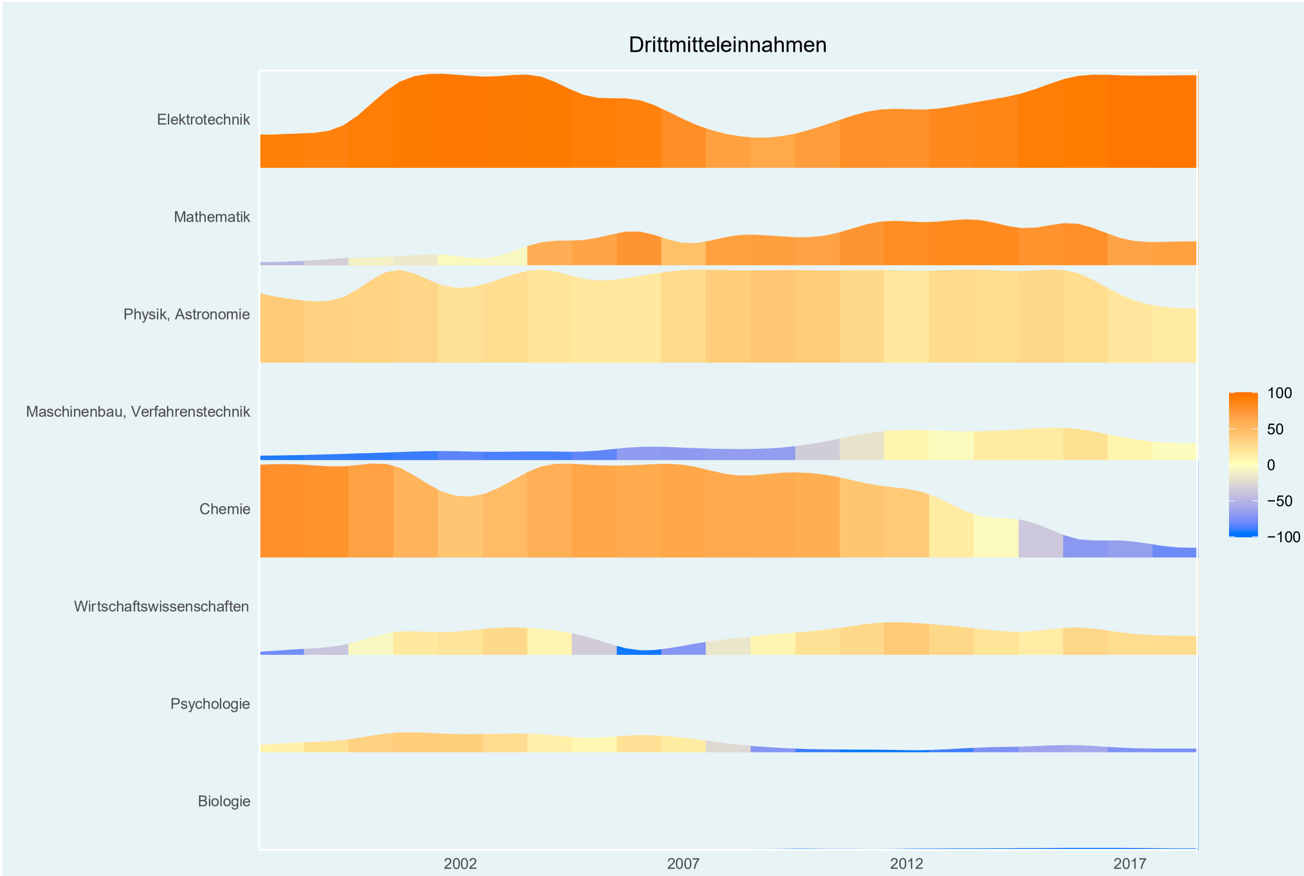

Figure 5. Two-dimensional heatmap, grant funding (Univ. of Wuppertal)
Source: RESP scores for grant funding plus respective size of disciplines.

Notes

All graphs can be found and used under the CC-BY-NC-ND-4.0 international license at https://fachprofile.uni-wuppertal.de/conferences/issi2021.html. Analysis was conducted in R (R Core Team, 2020) with data.table (Dowle & Srinivasan, 2019) and figures were produced using ggplot2 (Wickham, 2016). For smoothing the heat bars, the packages stats (R Core Team, 2020) and splines (ibid.) were used. The pseudo code for smoothing is (as geometric part of the ggplot command): stat_smooth(method = 'glm', method.args = list(family = gaussian), formula = y ~ splines::ns(x,df = years - 1), se = FALSE, geom = "ribbon", span = 1).

Acknowledgments

We thank Dirk Tunger and Paul Eberhardt for their support in data collection and curation.

References

Bonaccorsi, A., Colombo, M. G., Guerini, M., & Rossi-Lamastra, C. (2013). University specialization and new firm creation across industries. *Small Business Economics, 41*, 837-863.
Grupp, H. (1994). The measurement of technical performance of innovations by technometrics and its impact on established technology indicators. *Research Policy, 23*, 175-193.
Grupp, H. (1998). Measurement with patent and bibliometric indicators. In H. Grupp (Ed.), *Foundations of the economics of innovation. Theory, Measurement, Practice* (pp. 141-188). Cheltenham: Edward Elgar.
Harzing, A.-W., & Giroud, A. (2014). The competitive advantage of nations. An application to academia. *Journal of Informetrics, 8*, 29-42.
Heinze, T., Tunger, D., Fuchs, J. E., Jappe, A., & Eberhardt, P. (2019). *Research and teaching profiles of public universities in Germany. A mapping of selected fields*. Wuppertal: BUW.
Huisman, J., Lepori, B., Seeber, M., Frölich, N., & Scordato, L. (2015). Measuring institutional diversity across higher education systems. *Research Evaluation, 24*, 369-279.
Narin, F., Carpenter, M. P., & Woolf, P. (1987). Technological assessments based on patents and patent citations. In H. Grupp (Ed.), *Problems of measuring technological change* (pp. 107-119). Köln: TÜV Rheinland.
Piro, F. N., Aldberg, H., Aksnes, D. W., Staffan, K., Leino, Y., Nuutinen, A., . . . Sivertsen, G. (2017). *Comparing research at nordic higher education institutions using bibliometric indicators covering the years 1999-2014. Policy Paper 4/2017*. Oslo: NIFU.
Piro, F. N., Aldberg, H., Finnbjörnsson, Þ., Gunnarsdottir, O., Karlsson, S., Skytte Larsen, K., . . . Sivertsen, G. (2014). *Comparing Research at Nordic Universities using Bibliometric Indicators – Second report, covering the years 2000-2012. Policy Paper 2/2014*. Oslo: NordForsk.
Teixeira, P. N., Rocha, V., Biscaia, R., & Cardoso, M. F. (2012). Competition and diversity in higher education: an empirical approach to specialization patterns of Portuguese institutions. *Higher Education, 63*(3), 337-352.
van Vught, F. A. (Ed.) (2009). *Mapping the Higher Education Landscape. Towards a European Classification of Higher Education*. Dordrecht: Kluwer.
Bonaccorsi, A., Colombo, M. G., Guerini, M., & Rossi-Lamastra, C. (2013). University specialization and new firm creation across industries. *Small Business Economics, 41*, 837-863.
Borner, K., Bueckle, A., & Ginda, M. (2019). Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences of the United States of America, 116*(6), 1857-1864. doi:10.1073/pnas.1807180116
Brint, S., Proctor, K., Hanneman, R. A., Mulligan, K., Rotondi, M. B., & Murphy, S. P. (2011). Who are the early adopters of new academic fields? Comparing four perspectives on the institutionalization of degree granting programs in US four-year colleges and Universities, 1970-2005. *Higher Education, 61*(5), 563-585. doi:10.1007/s10734-010-9349-z
Dowle, M., & Srinivasan, A. (2019). data.table: Extension of data.frame. R package version 1.12.8. Retrieved from https://CRAN.R-project.org/package=data.table
Fortunato, S., Bergstrom, C. T., Borner, K., Evans, J. A., Helbing, D., Milojevic, S., . . . Barabasi, A. L. (2018). Science of science. *Science, 359*(6379). doi:10.1126/science.aao0185
Grupp, H. (1994). The measurement of technical performance of innovations by technometrics and its

impact on established technology indicators. *Research Policy, 23*, 175-193.
Grupp, H. (1998). Measurement with patent and bibliometric indicators. In H. Grupp (Ed.), *Foundations of the economics of innovation. Theory, Measurement, Practice* (pp. 141-188). Cheltenham: Edward Elgar.
Harzing, A.-W., & Giroud, A. (2014). The competitive advantage of nations. An application to academia. *Journal of Informetrics, 8*, 29-42.
Heinze, T., Tunger, D., Fuchs, J. E., Jappe, A., & Eberhardt, P. (2019). *Fachwissenschaftliche Forschungsprofile staatlicher Universitäten in Deutschland. Eine Kartierung ausgewählter Fächer*. Wuppertal: BUW.
Huisman, J., Lepori, B., Seeber, M., Frölich, N., & Scordato, L. (2015). Measuring institutional diversity across higher education systems. *Research Evaluation, 24*, 369-279.
Narin, F., Carpenter, M. P., & Woolf, P. (1987). Technological assessments based on patents and patent citations. In H. Grupp (Ed.), *Problems of measuring technological change* (pp. 107-119). Köln: TÜV Rheinland.
Piro, F. N., Aldberg, H., Aksnes, D. W., Staffan, K., Leino, Y., Nuutinen, A., . . . Sivertsen, G. (2017). *Comparing research at nordic higher education institutions using bibliometric indicators covering the years 1999-2014. Policy Paper 4/2017*. Oslo: NIFU.
Piro, F. N., Aldberg, H., Finnbjörnsson, Þ., Gunnarsdottir, O., Karlsson, S., Skytte Larsen, K., . . . Sivertsen, G. (2014). *Comparing Research at Nordic Universities using Bibliometric Indicators – Second report, covering the years 2000-2012. Policy Paper 2/2014*. Oslo: NordForsk.
Ruef, M., & Nag, M. (2015). *The classification of organizational forms: Theory and application to the field of higher education*. Stanford: Stanford University Press.
Team, R. C. (2020). *R: A language and environment for statistical computing.* . Retrieved from Vienna: https://www.R-project.org/.
Teixeira, P. N., Rocha, V., Biscaia, R., & Cardoso, M. F. (2012). Competition and diversity in higher education: an empirical approach to specialization patterns of Portuguese institutions. *Higher Education, 63*(3), 337-352.
van Vught, F. A. (Ed.) (2009). *Mapping the Higher Education Landscape. Towards a European Classification of Higher Education*. Dordrecht: Kluwer.
Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer.