September 2022 · Pegah Maham, Andy Wang

Data Science for Public Policy

How to Achieve Long-Lasting Impact

Stiftung
Neue
Verantwortung

Think Tank at the Intersection of Technology and Society

Acknowledgements

This paper is made possible largely because of discussions we have had over the past twelve months with experts who have implemented data labs or worked as data scientists in governments outside Germany. In particular, we would like to thank Jeffrey C. Chen and Alex C. Engler from the USA, Paul-Antoine Chevalier from France, Priit Võhandu from Estonia and Adam Bricknell from the United Kingdom for sharing their valuable experiences with us.

We also would like to thank the SNV team for their support throughout the research and publication process. In particular, we would like to acknowledge Luisa Seeling, Stefan Heumann, Shannon Reitmeir, and Andre Weisser.

The paper is part of the SNV topic area "Data Science" which is funded by Stiftung Mercator.

Executive Summary

Data science has thus far played only a minor role in public administration and governance, unlike in the private sector. However, data-driven approaches hold a lot of potential: data-driven tools can help ensure that administrative or political decisions are made on a better factual basis. Moreover, with the help of data tools, states can automate processes and utilise government resources more efficiently.

In Germany, the federal government has recognised the possibilities of data in governance. Its data strategy, published in 2021, sets a goal to significantly “increase innovative and responsible data provision and data use”. Among other provisions, the plan allocated around €240 million to establish data laboratories in all federal ministries and in the Chancellery itself by 2024.

However, for all of its potential, the establishment of data departments in administrative bodies also poses numerous challenges. Fundamentally, data scientists in this sector bear special responsibility: their work often affects large swaths of the population and can have an impact on critical areas of life over a long duration. In addition, there are hurdles to overcome in the practical implementation of data units. Data teams do not always succeed in finding suitable partners within their organization for the purpose of working together. Rigid administrative processes can also be burdensome when implementing data science techniques. And at the same time, high transparency standards accordingly impose constraints and limitations.

This impulse paper is meant as a guideline. In it, we address the question of how to successfully integrate data labs into public administration by compiling a series of concrete recommendations. Above all, we are guided by the maxim that data science in politics and administration should never be an end in itself. Rather, when developing data-driven products, data teams should focus on users and the concrete impact their products make in all phases of the process, from brainstorming to evaluation.

Data scientists in governance must fully understand their roles. They do not only have a technical task, as communication skills and interdisciplinary thinking are just as important for the success of their work. Data scientists in the public sector must be able to “translate” political and administrative problems into data science use cases that span a wide variety of users. They must be able to explain and evaluate their data products in

an accessible manner. Furthermore, they need to have an awareness of the risks of their methods: undetected biases in datasets can have major consequences—to the detriment of entire segments of populations.

Our paper follows the development of data products along the four phases of Ideation, Prioritization, Implementation, and Evaluation. Each chapter includes a list of the key "Do's and Don'ts" of building data labs. The final chapter focuses on recruiting and developing data teams. In the public sector, it can be difficult to find the right talent because data labs require a combination of skills and must compete with the private sector. We recommend making positions more attractive through cross-organizational dialogue and development. The search for talent should focus on candidates who have a strong orientation towards the public good and an interest in the greater social context of data work.

Table of Contents

Introduction

While data science has been well established within the commercial world, there has been considerably less attention paid to its usage and nuances in public policy. Yet, the potential to improve governmental processes and decisions is huge[1]: **Governments can utilise data-driven tools to streamline in-house processes and administrative services.** Such tools, for example, include language processing that extracts relevant information from text data and makes it usable. Visual data presentation, such as Covid-19-related dashboards, are another classic use case. This way, governments not only free up resources that can be used more efficiently on other tasks but also reinforce their citizens' trust in their ability to provide state-of-the-art digital services. **Governments should also take advantage of the benefits of data-driven methods (such as increased transparency, objectivity and speed) to improve their decision making, both for daily administrative decisions and long-term policies.** Moreover, with in-house data science units, governments can gain more technological sovereignty and independence from contractors and consultancies and obtain important data literacy skills across government bodies through spillover effects.

Having acknowledged the potential of data-driven methods, in January 2021, the German federal cabinet released its first-ever data strategy, which recognises that 'data is still not being used enough in today's Germany'.[2] Accordingly, the strategy laid out a plan to 'significantly increase innovative and responsible data provision and data use'.[3] A few months later, the German Recovery and Resilience plan allocated €239 million to build out data capabilities, including data labs in all federal ministries and the chancellery, enabling the strategy to be set in motion.[4]

Now, as the ministries work to build up their formal data capacities, Germany must navigate the opportunities – and unique challenges – at the intersection of data science and governance. **Such unique challenges include the more pronounced responsibility of governmental data science units, as the stakes are much higher in governance than in the private sector.** Government decisions affect all citizens, including the most vulnerable, and reach over long timeframes. Citizens cannot opt out or choose an alternative service provider for administrative processes. Therefore, governmental data units are confronted with higher expectations and standards regarding many aspects of their work.

1 see Engler, A. (2022). *Institutionalizing Data Analysis in German Federal Governance.*
2 p. 5, Federal German Government (2021). *Data Strategy of the Federal German Government.*
3 p. 6, Federal German Government (2021). *Data Strategy of the Federal German Government.*
4 p. 434, Federal Ministry of Finance (2020). *Komponente 2.1 Daten als Rohstoff der Zukunft In Deutscher Aufbau- und Resilienzplan (DARP).*

The increased responsibility and extra distinct challenges make the integration of data science units particularly difficult in governments compared to the private sector. These challenges include the following:

- The goals and considerations of a governmental institution are often more complex than those of a company. Diverse societal interests and large-scale interdependencies must be accounted for. This requires even more interdisciplinarity for the successful integration of data science methods.

- In the private sector, market feedback and competition pressure create the necessity of using innovative tools, automating processes and considering data for decisions. In companies, this creates incentives to co-operate with data science units in a way that it does not for governmental institutions, which makes it harder to find collaborators.

- Rigid administrative processes and restrictions allow less flexibility. Salary caps and software constraints, for example make it harder to compete with the private sector for talent.

- Due to technological and cultural gaps,[5] governmental organisations lack familiarity with data science methods and have less experience with modern workflows, such as feedback-intense iterations, which slows down the speed of data science units.

- To ensure equity, fairness and democratic accountability, standards for inclusion and transparency are higher compared to those in the private sector (such as accessible visualisations for people with vision impairment or limitations on inexplainable “black box” models), thus limiting the scope of potential data products and methods.

- Given the role of governments, the integration of data science is accompanied by public scrutiny and a critical opposition, making mistakes more costly, thus lowering incentives to take risks, share mistakes openly and learn from them.

5 Heumann, S. (2021). *Scheinlösung Digitalministerium.*

Given these unique challenges that governments face, it becomes crucial to approach the integration of data units strategically. **We recommend that data units prioritise a user-centric approach to build institutional legitimacy and ensure long-lasting results.** While this might sound obvious, the above-mentioned challenges can create significant obstacles. Therefore, in this paper, we have collected best practices to increase the chances of success. We divide the data science process into stages and provide a checklist with suggestions of what to do and what not to do for each stage. Given the distinct conditions and requirements, an additional fifth chapter is oriented towards the formation of data teams.

Our paper draws, firstly, from conversations over the last year with individuals[6] outside Germany who have created data labs in governments or worked as data scientists in ministries. And, secondly, from our own experience at Stiftung Neue Verantwortung. In the last two years, we have built a data team from the ground up and supported the think tank's independent analysis with data-driven products, such as an interactive visualisation of the German cybersecurity ecosystem[7], a trend analysis on semiconductor research[8] or a method to identify relevant stakeholders using social media networks ("Fishnet Method").[9]

Finally, while this paper is primarily geared towards data science in governance, our suggestions should apply equally to NGOs or think tanks, which face many of the same challenges.

6 Including Jeffrey C. Chen and Alex C. Engler from the USA, Paul-Antoine Chevalier from France, Priit Võhandu from Estonia and Adam Bricknell from the UK.
7 Herpig, S., Rupp, C., Semenova, A., and Maham, P. (2022). *Germany's Cybersecurity Architecture*.
8 Kleinhans, J., Hess, J., Maham, P., and Semenova, A. (2021). *Who Is Developing the Chips of the Future?*
9 Maham, P., and Semenova, A. (2022). *Fishnet Method and Code.*

Data Science and User-Centric Data Products

Data science spans disciplines from statistics to software development. In general, data science can be applied to automate processes and gain insights by using various methods. To name a few, these methods allow us to tap into new data sources (web scraping or APIs) and keep them up to date (data pipelines), analyse textual data (natural language processing), support forecasts (time series analysis), simplify complexity (cluster analysis), test hypotheses (statistical inference), make information more accessible (interactive visualisations) or support decisions (classification and prediction models).

Other countries have already started using data science methods and have illustrated the potential of data-driven methods for public policy and administration.[10] The following two data-based tools from the USA and the UK provide examples of successful data science applications in governmental institutions:

- Using pattern recognition to identify insider trading (USA)
  Data science can be very effective at helping bridge the 'law on the books' to the 'law in action', which is precisely the aim of a suite of products used by the American Securities and Exchange Commission (SEC) that identifies violators of insider trading. These tools must be explainable and fit within the current regulatory process to be a helpful partner to regulators. Notably, the Advanced Relational Trading Enforcement Metrics Investigation System (ARTEMIS) is not completely automated. Rather, SEC staff first identify 'trigger events' that would result in large changes in the value of a company's stock. Once potential insider traders are flagged, the ARTEMIS tool employs pattern recognition to compare previous trades to determine anomalies in trading behaviour.[11] By working with end-user expectations, ARTEMIS has become a valuable time-saving tool.

10 For instance:
- UK: Behavioural Insights Team. (2017). *Using data science in policy: a report by the Behavioural Insights Team. Report.*
- USA: Engstrom, D. F., Ho, D. E., Sharkey, C. M., and Cuéllar, M. F. (2020). *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies.* Administrative Conference of the United States.
- EU: Misuraca, G. und Van Noordt, C. (2020). *Ai watch - artificial intelligence in public services.* Publications Office of the European Union, Luxembourg

11 Engstrom, D. F., Ho, D. E., Sharkey, C. M., and Cuéllar, M. F. (2020). *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies.* Administrative Conference of the United States.

- Using data-driven segmentation to enable the specialisation of social workers (UK)
  In Newcastle, UK, data has been used to help minimise the casework of social workers for family- and child-related issues as part of the Family Insights programme. Upon intake into the system, families are grouped according to a variety of characteristics and then assigned to a case worker who specialises in similar cases.[12] In doing so, case workers avoid being 'spread too thin' and instead can deliver services accurately and efficiently across a specific sector.[13] While some have been initially doubtful about the effectiveness of segmentation as it relates to complex cases, the Family Insights programme hosts weekly case transfer meetings to refine segmentation criteria and ensure that segmentation works for social workers – and does not make social workers adapt to fit segmentation.

In this paper, we broadly use the term 'data products'. It can refer to visualisations such as COVID-19 infection dashboards, data acquisitions such as creating a central database for traffic data, tools supporting administrative tasks such as paperwork automation or analyses of energy consumption data to support policy decisions.

Regardless of the type of data product, each product should have an end user – someone for whom the data product should be useful. This may look like an administrator who must use a data workflow daily, a member of the public who examines open data sources or an analyst who uses dashboards to make decisions. Therefore, for data products to truly help the end user, they must directly address the user's needs. This process, by which the needs of the user guide the development of a data product, is called 'user-centred design'.[14]

12 p. 13, Symons, T. (2016). *Wise Council – Insights from the Cutting Edge of Data-driven Local Government.* Nesta and Local Government Association.
13 p.24, Beninger, Kelsey et al. (2017). *Newcastle City Council's Family Insights Programme*. Department for Education.
14 Chen, J. C., Rubin, E. A., und Cornwall, G. J. (2021). *Data Science for Public Policy*. Springer International Publishing.

The user-centric process is composed of four steps (see image below), each of which references the user and their needs:

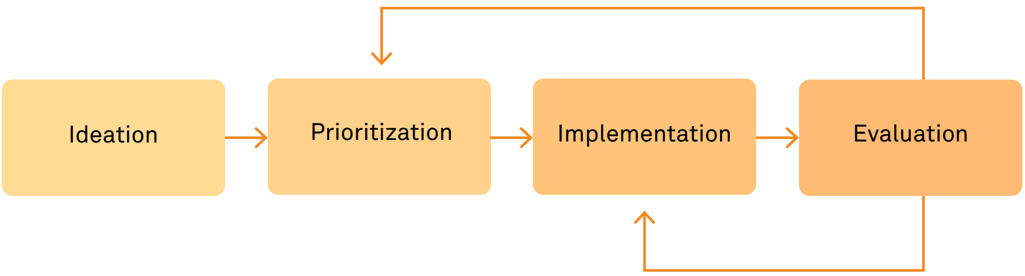
1. Ideation: The team identifies an end user and their needs.

2. Prioritisation: Together with the user, the team assesses the cost and potential impact of the project to decide whether to move forward.

3. Implementation: The team builds on minimal versions of the product, soliciting feedback from the user at each step.

4. Evaluation: The team tests the product and assesses its benefits, risks and limitations with the user.

Figure 1



The four steps of the user-centric process: Ideation, Prioritisation, Implementation and Evaluation. The evaluation impacts the two previous steps (based on Chen 2021).

In this paper, we examine each step individually, since each poses its own challenges and pitfalls.

1. Ideation

How can data science units identify project ideas, especially when they are building an institutional presence?

Identifying data science use cases is not straightforward, either in public policy or in the private sector: Compared to most other projects or departments, data science units are equipped with a method at hand and try to find a concrete problem to use it on – not the other way around, as usual.

Eventually, most projects should be identified by policy and administration experts who approach data science units with their requests. However, this is unlikely at the beginning. The habit of considering data-driven approaches has not been established yet, and with no experience or training, the awareness of data science applications is limited. **The data science team will be responsible for identifying use cases to a large extent. Therefore, the integration of the data science lab into the organisation becomes crucial, and it will not be sufficient to establish siloed units.**

Moreover, in their early stages, data units face the following challenges. They must prove their value to potential collaborators to incentivise working with them. Simultaneously, they need collaborators in the first place to identify use cases and to prove their value. **To create momentum, we recommend focusing on early adopters:** people excited to explore the potential of data science with the data team.

The following hypothetical example illustrates the importance of understanding users and their tasks to bridge the gap between data science on one end and public policy and administration workflows on the other. Imagine that administrative staff manually enter hundreds of zip codes individually into a search engine to obtain each of their geographical coordinates, which are needed for distance calculations. Given the team's lack of experience with data science, they do not realise this is a potential use case for a data science project. Conversely, the data science unit has experience using APIs as building blocks to automate generalisable tasks. In this case, zip codes could be automatically converted into coordinates using an API. Once a member of the data science team witnesses the described workflow, they can make a link to their tools. **By learning more about users, their tasks and their operating environments, data teams can find more opportunities for impact.**

The goal of the ideation step is to recognise pain points: ways in which data tools can optimise processes, augment understanding or support decision making. **By targeting potential impact as the metric by which your work is assessed, you will get more people on board eventually by proving the value that data science methods can provide.**

The following list provides suggestions of what to do and what not to do for the ideation phase:

Do

Make it your highest priority to understand the user and their context.
Work 'on the ground' at the level of workflows that already exist rather than trying to invent new ones. Conduct user interviews (formal or informal), and shadow their work, not only to learn what the user says may be helpful but also to understand the greater context of their work. Notably, users often do not know what they want from a data science perspective, especially when a data team is new and is just beginning its journey.

Expect the need to refine data science project ideas.
As your team starts, many other individuals in the organisation may have no idea how to interact with a data team or how data product development works. A good starting place is to conduct collaborative whiteboarding or brainstorming sessions with potential users or stakeholders about a specific problem. Build mutual understanding by working together.

Keep channels of communication open and flowing.
Potential stakeholders may not know how data science may help them. Create the conditions for open collaboration by proactively conducting outreach and maintaining an 'open-door' policy. This will not only make data science more accessible, but also will foster a culture where others can reach out to your team if they have ideas.

Don't

Never choose innovative techniques misaligned with user needs.
While it may be tempting to illustrate the possibilities of data science through cutting-edge methods or innovative tools, these are ultimately meaningless if they do not fulfil a need within your organisation. Data science methods should always remain a means to an end and not become a goal unto themselves. Always ask yourself: Does this project directly address users' needs?[15]

Set expectations within your team that initial projects will likely be technically simpler.
As your team starts, it is likely that the first wave of requests will require simpler methodologies, such as summary statistics or aggregated datasets. Help your team understand that this is just the beginning and that there will be room to demonstrate their technical abilities later. The successful delivery of basic projects in the beginning will create buy-in for more sophisticated projects later.

Others may have different understandings of the role of data science.
Similarly, the mandate of a data unit may be unclear to others in the organisation, and for some, it may be confused with the role of an IT team or even a software development team. Even if it is outside the domain that you expect your team to work in, remain open to all ideas. You may notice routines or processes that are valuable to automate or other insights that will serve your team well.

15 Office for Artificial Intelligence (UK) (2020). *Guidelines for AI Procurement.*

2. Prioritisation

Once you have multiple project ideas, how do you prioritise?

It is not uncommon for government technology projects to be developed at great cost, only to be discarded because an implementation case is never found. **Thus, prioritise your projects by weighing their impact and investment.** [16]

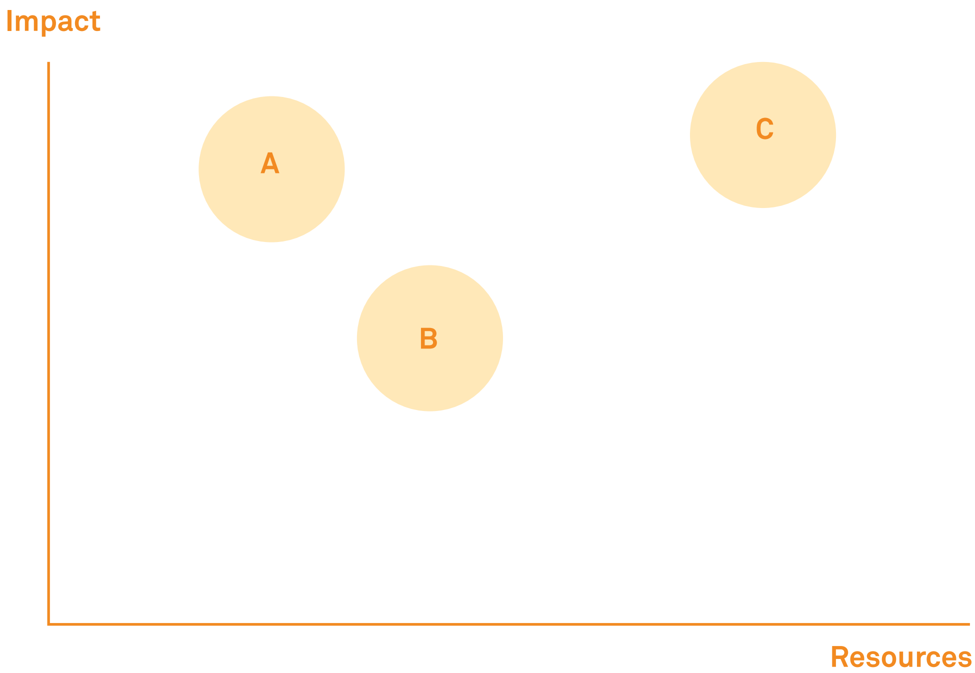
First, begin by discussing what the ideal outcome of the project is. **Evaluate what the impact of the project is in specific terms to the user:** How many hours can be saved, how much faster can decisions be made, or how will the insights of data analysis be used? In doing so, ensure that you avoid projects that fall within the 'so-what zone' – where a project would not achieve any meaningful change.

Next, take time to concertise the development parameters of the project, trying to be as specific and early as possible. What data will you use? Does it exist already, or must you create a dataset? What infrastructure do you need, and how readily accessible is it? What methods will you use, and is your team equipped for them? From these questions, you should not only have a better idea of the concrete outcome of the product, but also a sense of the resources – time, labour and money – required to commit to its development. Other factors to consider at this step could include identifying the likelihood of bottlenecks, the probability of success and legal or administrative barriers.

Finally, based on the expected impact and needed resources, assess whether this project is worth the investment right now. Focus on products that can have the most immediate impact, at least when you are still building your team. While moon-shot projects may be appealing, more ambitious projects require trust from agencies, which may be easier to earn through consistent and successful projects, even if they have a more limited scope.

16 p.79 provides a framework for conceptualising impact even more specifically. Misuraca, G., and Van Noordt, C. (2020). *AI Watch—Artificial Intelligence in Public Services*. Publications Office of the European Union, Luxembourg.

Figure 2



Sample decision grid for prioritisation. The horizontal axis indicates the amount of resources required to implement the project, while the vertical axis indicates the impact of the project. For example, Project A should be preferred over Project B or C because A has a greater impact than B while requiring fewer resources than B or C.

Do

Ensure clarity regarding user involvement in the development process. While assessing the feasibility of projects, be sure to factor in users' involvement, as their time resources are equally important to yours in priority. Clarify expectations concerning feedback, testing, domain knowledge and communication. Ensure that the teams you are working with have equal buy-ins, or you may face unexpected bottlenecks.

Assess project risks before beginning the development process. Before investing time into building out a project, consider the roadblocks – technical, political, ethical or legal – that may prevent you from reaching your ideal end product and its uptake. As early as possible, engage all relevant stakeholders, including the legal team and end users. Then, assess how you can avoid these pitfalls in your process. Ethical considerations arise naturally when working closely with and for the user.

Maintain a backlog of project ideas to which you can return.
Within your group, create a repository of unfeasible or deprioritised project ideas. Ensure that all relevant knowledge is documented to prevent doubled efforts in the future and to communicate adequately with stakeholders concerning why the project may not be feasible at that moment.

Don't

There are almost never 'ideal conditions' for beginning a project.
Especially in the initial stages of your data science unit, it is likely that data sources are messy or that you may even lack the right technical tools. Meanwhile, the data is never fully representative or unbiased. Before dismissing an idea, ask yourself what the next best alternative, the counterfactual, would be. (For example, COVID-19 case numbers depend on the number of tests taken, but this is better than no data at all.)

'Data exploration' should only be done to arrive at a specific, user-relevant question.
If presented with an interesting dataset, narrow the scope of your data exploration over time. Keep iterating the ideation and specification steps based on what a valuable data product for end users could be. Bear in mind that you are deviating from scientific best practices if you test your hypothesis on the same data that you used to generate it. As such, you risk that random correlations appear as statistically significant. [17]

Just because the data or tools are already prepared does not inherently mean that the product is useful.
Prevent the sunk cost fallacy: It is not worth starting a project just because a lot of work was already put into it. Make decisions considering the use case and its impact. Do not let a flashy dataset or innovative methods distract from the ultimate assessment of a project's worth: what it delivers to the end user. Manage expectations regarding the fact that many efforts will not come to fruition.

17 Generating hypotheses and testing them on the same data set can lead to false results. If enough hypotheses are tested (by exploring the data broadly), a random correlation will probably be found. See also discussions around "p-hacking" or "HARKing".

3. Implementation

How should data teams go about evaluating their products?

Data products[18] are best implemented in an agile fashion.[19] **In an agile development cycle, teams iterate between building minimum viable products (MVPs) and soliciting feedback from stakeholders. By doing so, data teams build trust in their units and avoid misallocating resources.** For example, when building a dashboard, start displaying the indicator as simply as possible. Once the data pipeline is established, you can refine the visualisation or add more indicators.

In each stage of an agile development process, known as a 'sprint', teams set a time limit to reach a certain set of key tasks. These time limits prevent teams from reaching too hard for perfection.[20] If you believe a project idea is viable, move forward by focusing on **creating a proof of concept** in your first sprint. This lean 'V1' should demonstrate that the end user's key goals can be met. The faster you can show a viable product, the faster you can obtain support for both the product and your team. Moreover, you minimise the tendency to overdevelop the project with unnecessary bloat for the end user.

After each sprint is over, the data team should present the working product to stakeholders. (Evaluation – in the short and long term – is the focus of chapter 4.) When starting your teams, you want to demonstrate not only that data teams can deliver on project specifications but also that other teams can effectively provide feedback. If a project is developed too far before soliciting feedback, the perception of sunk costs creates resistance to changing the product later – even when it is beneficial to the user.

18 Scientific studies should not be conducted in an agile fashion. As described in footnote 17, agile iterations and scientific best practices exist in tension with each other. Don't iterate the hypothesis testing and hypothesis generation process on the same data set.

19 Engelmann, J., and Puntschuh, M. (2020). *KI im Behördeneinsatz: Erfahrungen und Empfehlungen*. Kompetenzzentrum Öffentliche IT and iRights.Lab.

20 The primary goal is not to force a time limit, but to stay focused on minimal development and to constantly solicit feedback. It is difficult in software development (and especially data science) to accurately assess how long tasks will take.

Finally, based on the feedback, the team should evaluate their goals for the next cycle and build on the previous MVP. Teams break these goals into smaller steps and map them onto shorter deadlines. This cycle should continue until a satisfactory end product is reached. **If significant changes must be made along the way (such as a different data source or a rescoping of the user case), make sure to communicate with stakeholders and maintain a specific, clear use case.**

Do

Ensure structures for ownership within stakeholders.
Users play a critical role in developing a data product; their unavailability or unresponsiveness can block the progress on projects. To drive implementation and focus responsibility, you need clear ownership on the user side. Assign one person on the user side as the main point of contact for the data team, and structure collaboration up front (e.g. regular meetings).

Follow the KISS Principle: Keep It Simple and Straightforward.
Focus on the specifications of the project as guided by end users; do not do more! In fact, overengineering may make your product unusable for the intended audience. In deciding between methods, gravitate towards the more straightforward one: It is much likelier to be understandable by the end user. Only veer into complex methods if they are truly required.

Document not just code, but also context.
Work as if everyone from your data team could leave the next day. Many tools and data sets will be reused in future projects, ensuring that these are institutionalised and documented to prevent duplicated efforts in the future. Moreover, create an internal knowledge base for the context of your projects and work. Consider creating page-long "cheat sheets" for major datasets: Are there quirks to the data source? How frequently is it updated? How did the team obtain it, and are there limitations to its usage?

Don't

The largest aspects of user experience are often the smallest things.
Usability matters, but often, end users may care about different aspects of usability than the development team expected. Defer to end users on their experiences, and ensure you test often. Ease of use, visual presentation and accessible language can only be judged by the users. Take in all feedback with an open mind, even if it seems too nontechnical.

Do not hesitate to reach out to external expertise.
Especially as you start to build the capacities of the team, you are likely to have substantive or technical gaps in your in-house capacities. External data science units, civil society and specialised organisations may all be able to fill these gaps in various ways.[21] However, ensure that knowledge and expertise stay in your organisation so these partnerships also build future capacities.

Prevent data scientists from imagining potential use cases!
Be very precise about how the end user expects to use the product, and have this drive what your output looks like. This will provide a clear sense of direction for all efforts and smaller choices. Do not allow data scientists to veer into 'potential' territory: features that might be used or that might come from trying new methods.

21 Kupi, M., Jankin, S., and Hammerschmid, G. (2022). *Data Science and AI in Government.* Hertie School.

4. Evaluation

How should data teams go about evaluating their products?

Compared to other software products, mistakes in data products can be more difficult to detect.[22] In many cases, errors do not limit the functionality of the product, but rather can lead to distortions in its usage — whether subtle or pronounced. Take the example of a visa application prioritisation tool: If a biased data set is used to train the application, the bias carries over to its implementation, which may then discriminate against an entire population. An error like biased training data can cause profound damage, but it is much more difficult to detect as compared to a bug in the user interface of software. Such an example illustrates the importance of critically reflecting on both data collection methods and the context of product creation when implementing data-driven decision support tools. When insights from data are used to make policy decisions, decisionmakers must be able to interpret the significance of the results. This, in turn, requires a deep understanding of both the methodology and the domain context in which the evidence was generated. Creating this understanding is one of the most important tasks of data scientists in the public sector.

Simultaneously, governmental data science applications can be particularly impactful, and the responsibility placed on data science units is even more important. AI systems[23] in law enforcement that may interfere with people's fundamental rights or that are used for essential public services are defined as High Risk in the European AI Act – the second highest risk category after Unacceptable Risk – and will be subject to strict obligations. [24]

To ensure the necessary scrutiny, strive to allow a wide and diverse range of people to understand what you have done and to provide feedback. **Data products must be as explainable and transparent as possible. Place the responsibility for understanding on the data science team.**

22 See also this collection of failures or pitfalls in ML-based science, such as missing train-test splits: https://reproducible.cs.princeton.edu/#rep-failures
23 Data Science methods can include AI systems.
24 European Commission (2022, June 07). *Regulatory Framework Proposal on Artificial Intelligence.*

After each sprint – and once a final working product has been established – present your product to key stakeholders. This should include, at the very least, a sample of the end users and anyone else who may be managing or interacting with the product. Meet users where they are – structure your presentation according to each audience's expectations. How will they engage with the product? What level of technical detail must they know?

When presenting, start with a motivating problem or question for the project. Then explain your product as a solution (including methodology as necessary[25]) with as little jargon as possible. The end user – who may or may not be technically trained – needs to understand how the tool works. In cases of computer-assisted decision making, the administrator of the tool needs to trust the digital tool to be a reliable partner while understanding its limitations. In the case of data-driven insights, stakeholders must be aware of potential biases.

Finally, solicit as much feedback as possible. Begin with questions relevant to the user, perhaps following up on previous concerns. Then, allow end users to feel trusted in the process by opening the floor for general comments. Make sure you acknowledge all suggestions equally, and do not pass judgment hastily. Take feedback to iterate through the design process, setting new, narrow development goals. In the case of a final product, consider how feedback may be applied towards long-run implementation or revision goals.

By comparing the actual impact and invested resources against your initial estimates and predictions, you can consider these updates for upcoming projects and refine your team's prioritisation capabilities.

Do

Be transparent about the limitations and successes of your product. Provide as much information as possible to allow trust in your product.[26] What data was your product trained on? What model was used? How much more accurate was it compared to previous predictions? Acknowledging the successes and limitations of your product publicly answers questions and addresses scrutiny.

25 This includes relevant parameters (e.g. the false negative rate), their effect on results and what they are set to.
26 Recommendation 9, Engelmann, J., and Puntschuh, M. (2020). *KI im Behördeneinsatz: Erfahrungen und Empfehlungen.* Kompetenzzentrum Öffentliche IT and iRights.Lab.

Before deploying, recheck the risks of your product with diverse users.[27]
A bad product – even if fully developed – should not be set in action. Always take the time to understand what might go wrong with the product. Does it hurt any sector of the population? What ways could deployment go wrong? Ensure that everyone affected has a chance to comment on the product.[28]

Honestly compare your product to the next-best alternative.
Just as the ideation step deals with discovering pain points, your intervention should be specific about the improvement that your tool brings. How many work hours are saved? How much more accurate is a process compared to human judgement? These figures help manage expectations, recalibrate your estimations and increase buy-in.

Don't

Don't dismiss nontechnical feedback.
All aspects of the user experience – whether technical or not – are instrumental in designing a usable product. While expert audience members may comment on the project design or scientific specifications, some users may comment on cosmetic aspects or accessibility. Take the time to understand all the feedback provided to you.

You will have blind spots that can only be overcome with external evaluation.
Both your users and your team are likely to have blind spots, whether from established processes or perceptions of sunk costs. Regularly solicit external feedback – whether from other individuals in your organisation or external reviewers – to gain new perspectives. You may uncover a more efficient solution to a problem or discover the ethical limitations of your methods.

Long-term evaluation matters even more than short-term feedback.
he goal of the data science team is to create usable products. Check in on the usage of your products regularly over longer time periods, such as once a year. Understand how the product was used to improve decisions. If a product is not used anymore, investigate why (e.g. maintenance, ownership issues, mismatch with user needs). Consider these for future projects. If it is still being used, what feedback is there? How can the product be even more helpful?

27 Central Digital and Data Office (UK) (2021, March 31). *Make Things Accessible and Inclusive.*
28 Office for Artificial Intelligence (UK) (2020). *Guidelines for AI Procurement.*

5. Teams

How do you build strong data science teams in the public sector?

As already described, public sector data labs possess a special responsibility and need to combine various skill sets. It is important to recruit and develop the right talent sustainably. Compared to the private sector, data teams will not be developing cutting-edge technologies or offering equivalent compensation, at least in their beginning stages. Therefore, teams should look for impact-oriented candidates that have domain expertise and can learn quickly.

Data scientists in governance data science units will largely work as translators, connecting worlds and communicating effectively between them. While technical skills are essential for a data team, members of these units must also understand the broader context of their work and garner the trust of agencies. They must be able to translate algorithms and code into real-world policy contexts and administration, partly to less technologically literate audiences.

The projects and tools will change, but the goal of your work will not. **Focus on impact-driven candidates with sustainable motivation.** The best candidates are those who wish to create an impact through policy and who have an eye on using data as the means to do so. Candidates who only focus on data science may find their government structures stifling and their work limited in scope. Therefore, assemble a collection of data scientists with demonstrated interest in policy, data scientists with policy backgrounds or domain experts with data science skills.

Finally, 'data science' is a field that covers a wide range of skills. Depending on the projects that your team undertakes – which cannot readily be anticipated in advance – **skill needs may change rapidly.** Therefore, be clear about what your team needs when hiring: Is there a certain competence that the team urgently needs, or are you seeking a generalist who can adapt? In any case, value flexibility and the willingness to learn, whether new methods or substantive contexts.

Do

Ensure the right people know that your team is hiring.
Many people work at the intersection of policy and data, but they may not know that your data units exist. Put dedicated efforts towards reaching out to potential hires, emphasising the impact of the role. Create a central website or newsletter where governmental tech and data job openings are announced, and cross-post openings onto job boards focused on social good.[29]

Be clear about the advantages of working in a data lab.
The momentum for data science has long swung towards corporate jobs, including training and recruitment. Highlight why government work is a meaningful alternative. Emphasise the impactful nature of data work in governance with case studies of real-world impacts. Point to interesting data that you have collected and with which data scientists can work uniquely.

Connect data science teams across ministries.
Working together means better work. Establish channels to communicate between ministries, such as regular meetings, chat channels and shadowing. Build centralised resources and repositories to prevent doubled efforts. Provide training, and integrate lab members into a broader community of impact-oriented data scientists. The UK government, for example has established a cross-government data science chat channel and mailing list.[30]

Don't

Be realistic about expectations in hiring: Beware an 'innovation-first' mindset.
Hiring descriptions should accurately reflect the tasks that candidates are expected to perform. Do not set the expectation in job ads or interviews that data scientists will deploy cutting-edge methods or tools as a primary component of their work. Doing so may attract more people at first, but it is not sustainable if candidates leave due to unfulfilled expectations.

29 For example *Good Jobs* or *Gesines Jobtipps*
30 The UK has an extensive community programme for public-sector data scientists, including training sessions, meetups, conferences and a Slack chat channel. More information: https://www.gov.uk/service-manual/communities/data-science-community

Conventional criteria for government positions are much less relevant for data scientists.
While government hiring may focus on a set of well-established credentials (e.g. university degree, time in civil service etc.), data scientists are much likelier to have nontraditional backgrounds. Qualifications may come from online courses or specific course selections, which makes projects and experience more relevant than academic background. Beware of being too narrow in position requirements. For example, listing STEM (science, technology, engineering and mathematics) degrees as a requirement may deter qualified applicants who have significant data experience in policy fields.

Anticipate clashes with neighbouring teams or processes.
In organisations as large as ministries, many departments – IT, statistics or software development – may work on similar problems or processes as data science teams. Reach out early to clarify responsibilities so there is no confusion, especially in hiring. Identify possible collaborations (e.g. data scientists providing sources for statistical evaluation) and, when possible, work together to refer qualified candidates to each other.

Conclusion

The introduction of data teams into federal ministries represents a tremendous opportunity for the modernisation of German governance. However, these data teams are not self-acting. Rather, to ensure the lasting continuity of these teams, careful attention must be paid to how they are enacted and what projects they take on.

While it may be tempting to go for technological progress, implement the newest methods or show off the administration's digital skills, **the steadiest way to ensure that investments in digital advances pay off is to design for impact.** Begin by understanding the needs of ministries, their operating contexts and their workflows, and assess where data projects can make the greatest impact.

To ensure that these goals are seen through, focus on communication at every stage of the development process. In all four phases of the cycle, maintain close contact with your end users and other stakeholders, and above all, ensure you explain your expectations early in the process to avoid frustrations and misunderstandings.

Finally, it cannot be emphasised enough: Data scientists in the government possess a special responsibility regarding the data that they handle and the tools that they create. **Understand how the products you create fit within the legal, ethical and administrative landscapes.** Solicit feedback, and test your products consistently and constantly.

For years, data science was primarily applied in the public sector. Finally, it is gaining traction in public administration. We expect this field to develop strongly in the coming years, and we would like our paper to be understood as an offer for exchange and collaboration.

For any questions, the experts at SNV are very excited to open a dialogue on best practices and the administration of data units. Our European Policy Data Science Network also connects data-driven researchers at NGOs, think tanks and government organisations to support and research public policy.

If done properly, German federal agencies can recruit highly talented data scientists and enact long-lasting change. These data science teams can be the leaders of a better-equipped, digital German government. It is in everyone's interest to ensure they can do so.

About the Stiftung Neue Verantwortung

The Stiftung Neue Verantwortung (SNV) is an independent, non-profit think tank working at the intersection of technology and society. The core method of SNV is collaborative policy development, involving experts from government, tech companies, civil society and academia to test and develop analyses with the aim of generating ideas on how governments can positively shape the technological transformation. To guarantee the independence of its work, the organization has adopted a concept of mixed funding sources that include foundations, public funds and corporate donations.

About the Authors

Pegah Maham is the lead data scientist at the Stiftung Neue Verantwortung, where she has been building the data science team since 2020 and integrated data science into the think tank's work. In particular, she is concerned with the question of how trends in the development of artificial intelligence and its social consequences can be captured using data science methods. She is a co-founder of the Policy Data Science Network, which connects data scientists in public policy, think tanks, and nonprofits. Previously, Pegah worked as a data scientist at a startup in the renewable energy sector. She holds Master's Degrees in Statistics (Humboldt University of Berlin) and Data Science (London School of Economics).

Andy Wang is an intern with the Data Science unit of the Stiftung Neue Verantwortung. In the past, Andy has worked for the Research and Analytics Department of the New York State Office of the Attorney General and the Research and Data Division of the Legal Services Corporation. Andy studies Social Studies, Philosophy, and Computer Science at Harvard College. He is an Undergraduate Fellow at the Edmond and Lily Safra Center for Ethics as well as a Fellow of the Harvard Kennedy School Program for Science, Technology, and Society. His work largely focuses on the role of data in governance and the possibilities of digital knowledge.

Imprint

Stiftung Neue Verantwortung e.V.
Beisheim Center
Berliner Freiheit 2
10785 Berlin

T: +49 (0) 30 81 45 03 78 80
F: +49 (0) 30 81 45 03 78 97
https://www.stiftung-nv.de/en
info@stiftung-nv.de

Design:
Make Studio
www.make-studio.net

Layout:
Tom Semmelroth
Alina Siebert