Information Systems Journal

WILEY

SPECIAL ISSUE ARTICLE OPEN ACCESS

The Data Product Canvas: Designing Data Products for Sustained Value From Enterprise Data

M. Redwan Hasan[1] | Bastian Finkel[2] | Christine Legner[1]

[1]Faculty of Business and Economics (HEC), University of Lausanne, Lausanne, Switzerland | [2]SAP SE, Walldorf, Germany

Correspondence: Christine Legner (christine.legner@unil.ch)

Received: 1 July 2024 | **Revised:** 24 April 2025 | **Accepted:** 29 April 2025

Funding: This work was supported by Competence Center Corporate Data Quality (CC CDQ).

Keywords: AI use-case | data democratisation | data management | data product | data product canvas | data use case | data value | data-driven innovation | design science research | visual inquiry tool

ABSTRACT

Organisations are increasingly striving to become more data-driven by embedding data into decisions, interactions and processes and by leveraging advanced AI technologies to unlock innovative use-cases. However, many remain unprepared to meet the rising demands for data, analytics and AI. A data product mindset—combining, packaging and delivering data as a product—has emerged as a promising approach to meet the needs of an expanding user base. Despite their popularity, data products are often seen as a purely technical concept, with suitable methodologies and tools for designing them still underdeveloped. This paper introduces the data product canvas, a visual and versatile tool that helps cross-functional teams—comprising business, data, analytics and IT experts—collaboratively design new data products and assess existing ones. The canvas ensures that critical themes are addressed: *desirability* from the customer perspective, *feasibility* from the technical perspective and *viability* from the economic perspective. The practical application at SAP illustrates how the data product canvas supports its data democratisation initiative, showcases real-world examples and offers practical insights to guide future adopters: (a) tailoring designs to different data product types, (b) periodically refining data products to increase their value and (c) systematically assessing requests to build a cohesive data product portfolio.

1 | Introduction

> We have loads of data that we can really leverage for bigger goals [...] but the problem is consistency in its creation and management [...], which is not helped by the fact that we are such a huge company. The only way is to find some sort of structured way to do it.
>
> (Team Lead Democratised Governance, SAP)

As the volume of data continues to grow exponentially and analytics becomes increasingly pervasive, many organisations are striving to become more data-driven (Bean 2022). They are aiming to embed data into every decision, interaction and process across the enterprise (Assur and Rowshankish 2022), whereas leveraging advanced AI technologies to unlock novel use-cases. However, the rising importance of data confronts organisations with significant challenges, and many are not fully prepared to meet the growing demands: First, most companies operate with centralised data and analytics teams that typically address individual user requests one at a time. For instance, if a vendor dashboard is requested to conduct tail-spend analysis, the central team must locate, stage, clean and prepare the necessary data, upload it into an analytics platform and then develop the dashboard. Now imagine rinsing and repeating this for every single-data request coming from different functions! The sheer volume of requests overwhelms the central team, making it impossible to handle them reliably.

This practice also fails to scale because a dedicated pipeline must be built for every incoming demand, making this approach slow, expensive and unsustainable (Desai et al. 2022). Second, various business teams within the same enterprise build their own tailored solutions to address specific use-cases. As these solutions are typically not shared across teams, they cannot be widely reused to drive other use-cases in the firm (Davenport and Kudyba 2016). This lack of visibility and coordination leads to redundant efforts, with similar solutions being created repeatedly and often underestimating the effort required for ongoing maintenance. Third, creating value from data requires collaboration among different stakeholders from business, data and IT teams, responsible for preparing, delivering and using data within the organisation. However, diverging expectations and a lack of communication between these groups often leads to a mismatch between capabilities of data providers and expectations of data consumers (Schulze and Tiwary 2024).

In response to these key challenges, companies are increasingly adopting a product mindset towards data (Davis et al. 2020). Similar to how a firm manages physical products—from raw materials to final consumption—a comparable approach is applied to data (Wang 1998). Under this data product approach, data is industrialised, packaged and delivered in such a way that users can rapidly identify and access the right insights fitting their purpose (Assur and Rowshankish 2022). This comes with several advantages: On one hand, combining, packaging and delivering data as a data product allows them to effectively address information needs of targeted end user segments (Hasan and Legner 2023a). On the other hand, packaging data in a more standardised fashion allows providers to leverage existing capabilities to embed reusability across multiple use-cases. As such, organisations can deliver analytics to users 90% faster, lower their cost of ownership by 30% and productise data to align with various consumption patterns (Desai et al. 2022). Therefore, it is no surprise that frontrunners have already built and implemented data products within their organisations to generate and capture value, both in the short and long run (Davenport et al. 2024). As an example, Vista, the Dutch online printing firm, has globally unlocked around $100 million in incremental profit with their 120+ data products covering eight domains, including channel marketing, pricing and promotion and product and personalisation (O'Neill 2023). Regions Bank, a retail and commercial banking giant in the United States, has saved over $1 million in annual payments to vendors using just their 'Regions Voice of the Customer' data product (Davenport and Bean 2022).

Despite these promising prospects, however, companies still face many open questions when putting data products into practice. The concept remains nascent, and it is often unclear what a well-defined data product entails and how data products differ from existing datasets or assets (Davis et al. 2020). With the advent of the data mesh (Dehghani 2021), many companies are exploring and experimenting with data products. However, the concept is primarily regarded from a technological perspective, whereas suitable tools and methodologies for designing data products remain underdeveloped. Without integrating business and economic aspects into product design, organisations will waste resources and end up with ineffective data products, failing to generate value from their data (Desai et al. 2022).

This paper addresses these challenges by outlining *how organisations can effectively design and assess data products to generate sustained value from enterprise data*.

To this end, we propose the data product canvas, a visual inquiry tool that supports cross-functional teams—comprising business, data and analytics and IT experts—in collaboratively designing new data products and assessing existing ones. The practical application at SAP's internal data democratisation initiative, led by the Chief Analytics Officer, showcases concrete examples of data products and offers practical insights to guide future adopters. In doing so, we show how companies can design suitable data products to accelerate their data-driven projects and initiatives and foster sustained value creation from enterprise data.

2 | Towards the Creation of the Data Product Canvas

With the increasing importance of data and the advent of data mesh, many companies are actively exploring and experimenting with data products. Reflecting this trend, data products have also been a central focus in the Competence Center Corporate Data Quality, an industry-research collaboration with Fortune 500 companies across various sectors, including automotive, pharma, retail and consumer goods industries, over the past 3 years. Through our joint research activities and regular exchanges with data experts from 20 firms, we found that many of them struggled with the same fundamental questions: What constitutes a well-defined data product? And how can data products be designed effectively? A major challenge lies in the general tendency to view data products purely as a technical topic, often overlooking their alignment with key organisational goals and the specific business challenges they aim to address.

To help these companies in their data product journeys, we first analysed their emerging data products and derived a working definition and typical categories of data products (Hasan and Legner 2023a). Building on this foundation, we developed a visual inquiry tool, namely the data product canvas, to assist practitioners in designing data products. The canvas was created following a design science research (DSR) process (Peffers et al. 2007), in collaboration with 30 experts representing 15 firms. Although the initial version of the canvas was grounded in a comprehensive literature review on data products and incorporated established design principles for visual inquiry tools (Avdiji et al. 2020), it was iteratively refined based on practitioner feedback. This involved discussions in focus groups, demonstrations on real-world data products and expert evaluations, resulting in a stable version of the data product canvas after three iterations (Hasan and Legner 2023b). To showcase the use of the canvas, we built a data product library, documenting the design of 22 data products from 17 organisations, both in the public and private sector.

Since mid-2023, the data product canvas has been publicly available under a creative common licence (link). It has been

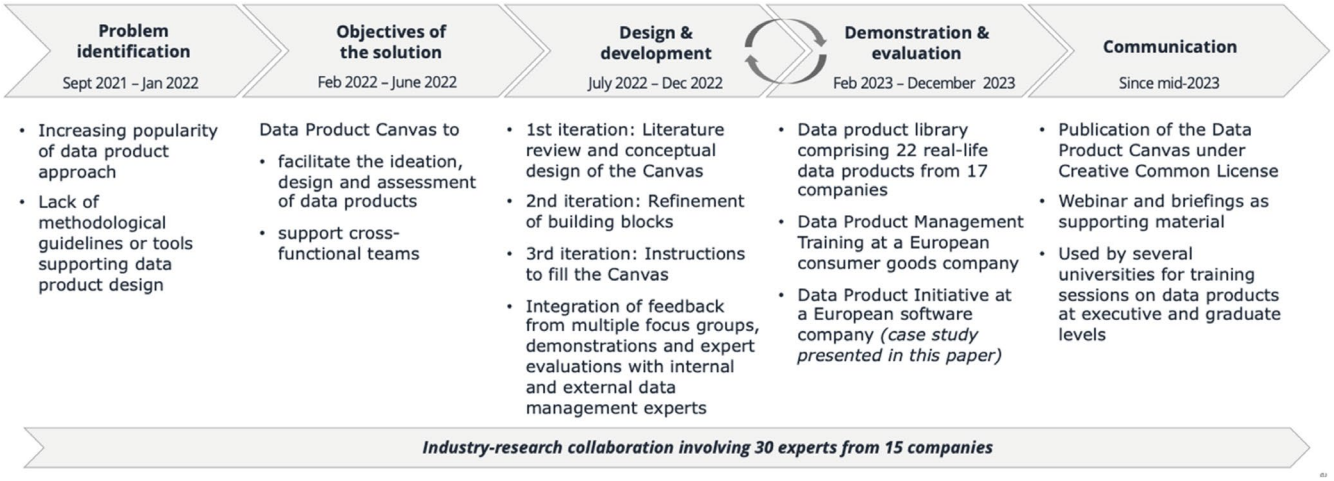


FIGURE 1 | Design science research process for the data product canvas based on (Peffers et al. 2007).

taught in several training programmes on data products at executive and graduate levels, with an increasing user base across various types of organisations, including small and medium-sized enterprises. Several companies have adopted the canvas to conceptualise, develop and evaluate their emerging data products. A large European food manufacturing company integrated the data product canvas into their 'Data Product Management' training programme where 60 practitioners were trained on the canvas and subsequently documented 10 emerging data products. Convinced of its effectiveness in guiding data product ideation and design, the data product canvas became part of their standard operating procedure (SOP). At a large software vendor, the data management team used the data product canvas to support the design of their initial data products, including a data quality dashboard ('Data Quality Mission Control Center') and a smart matching model ('Function and Department Engine'). The feedback was highly positive, emphasising its intuitive and interactive nature. This led to the broader application of the canvas as part of the data democratisation initiative, led by the Chief Analytics Officer, which we will present later. Figure 1 provides the timeline on the development of the data product canvas, and Appendix 1 provides more details on the research activities.

3 | Data Product Canvas

The data product canvas supports organisations in their journey to become more data-driven by facilitating the design and assessment of data products. It broadens the predominantly technical perspective on data by encouraging stakeholders to adopt a business-oriented view and focus on addressing end user needs. It does so by first bringing together cross-functional teams—composed of business users who consume data products, data and analytics teams who develop them and IT experts who deploy them—to collaborate on data product ideas and discuss relevant design aspects. As such, a valuable dialogue is established where experts from diverse roles can share their perspectives and reach mutual understanding of key questions underpinning the data product. Second, the canvas makes this exercise interactive so that participants not only reflect on critical questions surrounding data products but also discuss and have fun with their colleagues from other parts of the organisation while doing so. Third, the canvas aims to simplify the process of designing and assessing data products for any practitioner, regardless of their work profile and level of knowledge. Hence, the canvas is made intuitive by using colours, simple questions and possible examples to guide the thought process of any individual.

3.1 | Desirability, Feasibility and Viability: The Tri-Perspectives

The data product canvas builds on core ideas of design thinking methods and innovation tools, drawing particular inspiration from the business model canvas (Osterwalder and Pigneur 2010). Conceptualising a suitable business model allows firms to offer the most appropriate product that meets consumer needs, whereas considering the required resources and economic viability. Similarly, the data product canvas outlines the critical elements for designing data products that ensure a strong fit with the recurring information needs of the consumers, and ensures the technical capabilities needed to build and deliver the product. The nine building blocks are organised around three key perspectives (see Figure 2): *desirability* (in red) that gauges the consumers' requirements for the data product; *feasibility* (in yellow) that underscores various technical capabilities and tasks needed to address those requirements, and *viability* (in green) which addresses the economic suitability of investing in the data product akin to similar assessments made during product development.

These themes play a vital role in highlighting the main aspects of a data product. First, they help identify the specific people who are needed as experts to outline the details in each of these themes. For instance, one may need delegates from the business functions where the data product will be used to address the *desirability* theme, the data and analytics and IT experts to address the *feasibility* theme and managers who also have authority for sponsorship or funding to address the *viability* theme. Second, these themes intuitively harmonise the key concepts surrounding data products. They illustrate that data products are

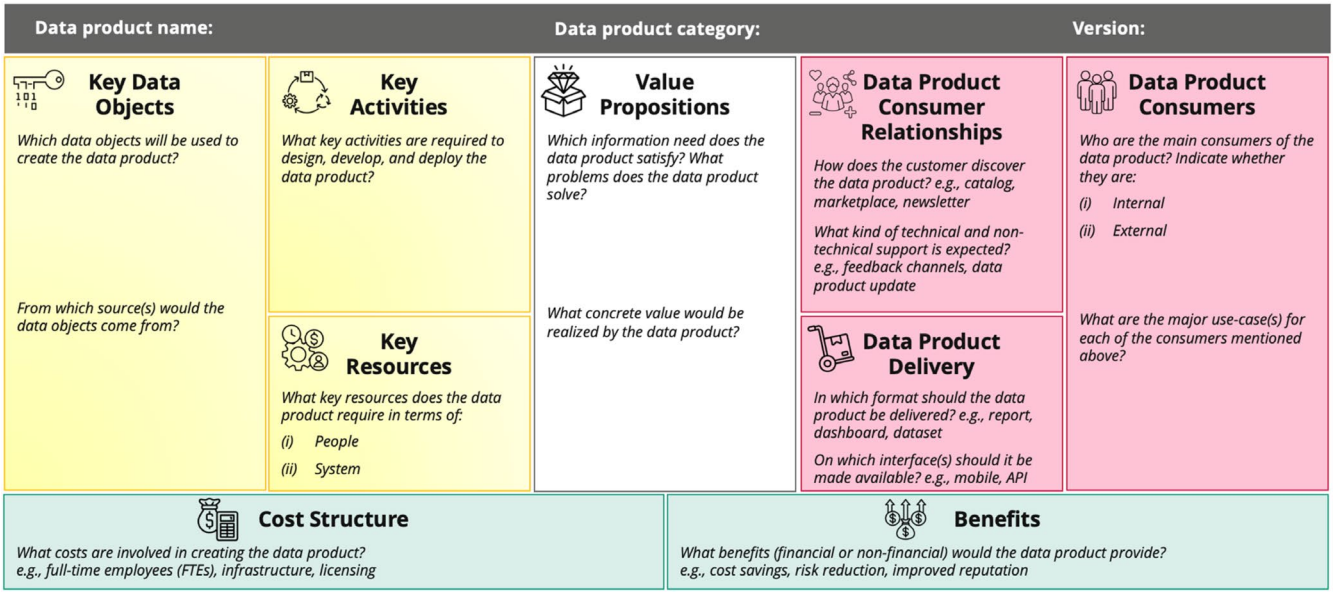


FIGURE 2 | The data product canvas (Hasan and Legner 2023b).

designed to address the needs and solve the issues of business users (*desirability* theme), leverage the most suitable resources (*feasibility* theme) to generate tangible value for the company (*viability* theme).

3.2 | Value Proposition—The Focal Point for Data Product Design

> As it comes with being part of a software company, we are very solution-oriented and love to think about technical problems, but not many of us try to understand the business value.
>
> (Team Lead Democratised Governance, SAP)

Designing data products with a technical mindset is a common pitfall, often resulting in solutions that are technically sound but fail to address real business problems or meet user needs. To emphasise the potential value generation of a data product, the canvas places a critical block at the centre: *value proposition*. This ensures the design process starts with a clear understanding of the value the data product aims to deliver by addressing consumers' information needs. The value proposition is purposefully positioned in the middle of the canvas to accentuate its influence on the surrounding blocks as it helps identify the key stakeholders who will use and generate value from the data product (*desirability* theme), specify exactly what data and resources are required (*feasibility* theme) and estimate the returns that will substantiate this value (*viability* theme). This central block prompts practitioners to reflect on the specific problems or needs and then outline the expected value that will result from addressing them. This assessment is crucial for linking data products to broader business objectives and evaluating whether they are aligned with consumer needs. For instance, a data quality dashboard for vendor data can help purchasing managers resolve issues such as vendor duplication or incomplete address information, which are critical for contracting and accurate invoice processing. Additionally, answering the questions in this block makes participants recall whether similar data products already exist within the firm, not only helping to avoid duplicative efforts but also reflecting on new possibilities to augment existing data products. As such, resource concerns can be embedded into their thought process from the very beginning.

3.3 | Desirability—'Who Are the Consumers and What Are Their Information Needs?'

> So, how many times have you seen that a fancy new dashboard initially generates a lot of interest, only for that fascination to quickly die down? That is because the solution does not properly address consumer pain points and the consumers realize this sooner than you expect.
>
> (Data Scientist, SAP)

One of the challenges accompanying the rapid growth of data and analytics in enterprises is the tendency to create dashboards, reports, machine learning models and other artefacts quickly for specific use-cases, often without a long-term vision for the broader information needs they are meant to address. The desirability perspective emphasises the importance of addressing consumers' core information needs and ensures that, like any other product, data products are designed to be desirable to the consumers. As different types of users have varying backgrounds, purposes, skills and preferences in how they consume data products, understanding who these data product consumers are and how they seek to use data products is paramount. The *data product consumers* block helps with this. First, it outlines the future users, whether they are internal or external and their primary information requirements, based on the use-cases they support. For instance, a pharmaceutical scientist may require minute-to-minute readings of the temperature of a newly created drug product, which can be addressed by a real-time dataset, whereas a marketing manager may need to understand why their campaigns fell

short in generating leads, postulating the need for a proper dashboard with marketing KPIs per region or customer segment. These details, therefore, hint towards the emerging consumption patterns and also prompt reflection on how the intended data product should be governed.

To support consumers in consistently accessing and using the data products, the *data product consumer relationship* block captures how consumers can discover available data products and what support is needed for effective use. An example could be using a data catalogue to skim through the list of available data products, or a data marketplace could enable users to purchase a data product. Additionally, a dedicated data product team can provide the necessary support to facilitate adoption. Building on the consumer profiles outlined above, the *data product delivery* block then captures the most fitting way to deliver the product. Specifically, it defines the format and channels through which the product should be offered so that it aligns with the consumers' aptitude level and preferences. This is part of the packaging of the data product—just like one would package and deliver a physical product in a specific way for a specific customer—in order to fulfil the need of the consumer in the way they desire. For instance, the pharmaceutical scientist may require the data through an API call for further analysis, whereas the marketing manager may need to access the dashboard on their company smartphone at all times.

3.4 | Feasibility—'Can We Deliver What the Consumers Want?'

> To enrich a data product, at SAP we try to capture and prepare data from various sources such as our S/4 HANA systems or internal data centers [...] allowing us to get a consolidated picture of the business problem in question.
>
> (Data Product Owner, SAP)

One of the key questions in creating successful data products is whether the appropriate data exists to meet their needs. Consequently, the feasibility perspective shifts from emphasising user requirements to addressing the data requirements essential for building the data product. The *key data objects* block enables the data, analytics and IT experts to specify all the data required to develop the product. This is an important exercise because it facilitates a gap analysis between existing data and the data required. It thereby expedites discussions about whether, for instance, currently available data are sufficient or additional data need to be collected or purchased from third-party sources. These decisions, however, carry implications for finances, timelines for building and deploying data products as well as data security and governance—all of which can now be reflected on and aligned during the ideation phase.

Having identified the data needs, participants reflect on how data must be transformed from its original state into a usable data product. For this, the *key activities* block captures the sequential steps needed to transition the data product from its design phase towards its development, deployment and consumption. Such a lifecycle view can show the current state of a data product and illustrate its lineage, enabling business users to see how the data product came into being—instilling confidence in the data product. To effectively organise these activities, naturally, companies must then consider the *key resources* required, especially in terms of systems and personnel. Such reflection facilitates the revaluation of the firm's resource allocation strategy, helping to determine whether existing resources are sufficient or if additional support—such as consultants, new hires, or emerging technologies—needs to be acquired.

3.5 | Viability—'Does It Make Sense to Commit?'

> It is always a significant challenge to secure funding for any sort of data-related project as we have to demonstrate that we can tangibly generate returns from data [...] which is often quite hard to estimate.
>
> (Data Product Owner, SAP)

The viability perspective underscores a critical aspect: the difficulty of building a strong business case and securing investments in data products. Here, the *cost structure* focuses on identifying development costs and potential areas where fixed or variable cost may arise during deployment and consumption. This will, first, help pinpoint key cost areas and provide an estimation based on historical data for each cost area. The *benefit structure* deliberates the concrete advantages that could potentially occur. These benefits may be either monetary, such as cost savings or revenue generation, or non-monetary, including factors like brand image, culture and mindset within the organisation. Understanding these different areas helps teams perform an assessment between the expected versus actual cost/benefit to estimate the performance of the data product. Moreover, such discussions remain critical to ensure transparency on who will bear the cost and who will fund future investments—questions that many organisations struggle with. Aligning on these aspects during the design phase ensures early commitment and lowers the possibility of cross-functional conflicts.

4 | Data Democratisation at SAP

> Our vision is to enable AI-powered, data-driven decision making by every SAP employee.
>
> (SAP's Chief Data and Analytics Officer)

With SAP's data democratisation initiative, a clear objective has been laid out: take the data from the hands of a few to the hands of many. The primary goal is to decentralise data across the company to generate novel insights through AI for driving informed decisions. The interest in leveraging data and AI is very high at SAP, and over time more than 1000 data assets—such as dashboards or AI models—have been created: 'As a software company, the analytical skill level existing within SAP is exemplary and people are always willing to do amazing things with data' (Team Lead Democratised Governance, SAP). However, certain challenges around data access and usage hinder the realisation of the full potential. For one, as a large organisation, SAP employees struggle to locate the appropriate data needed for their analysis. This difficulty often

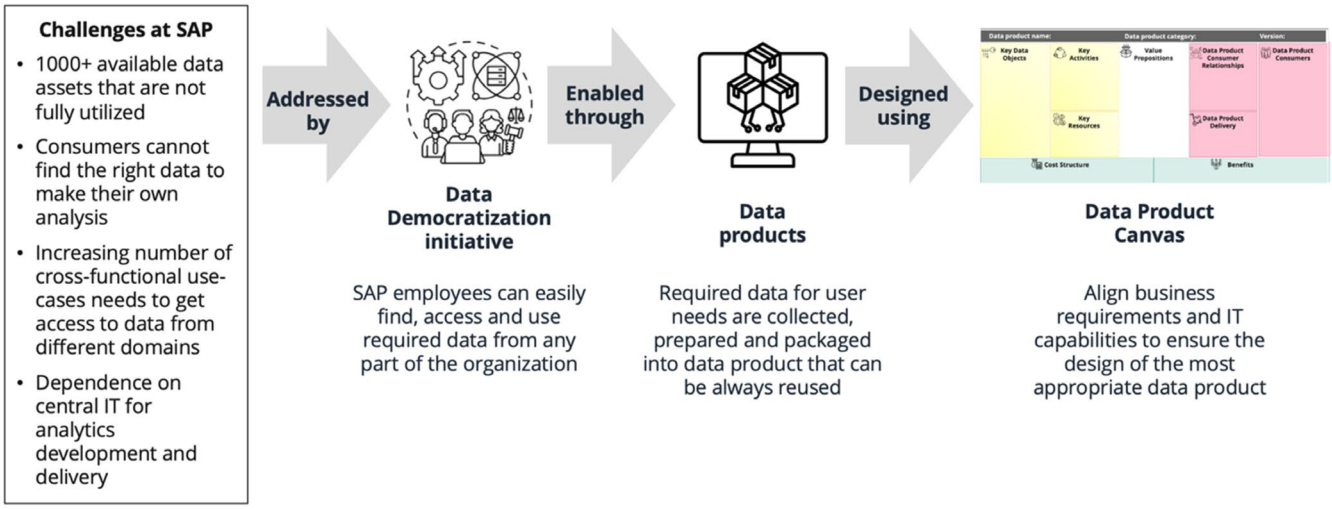


FIGURE 3 | The role of data product canvas in SAP's data democratisation initiative.

leads to the use of low-quality data from unreliable sources, compromising the quality of insights generated for critical use-cases. Moreover, thousands of data assets are built by employees for specific needs, which remain siloed and unusable for other purposes, eventually becoming costly to maintain for the company. Furthermore, employees still often rely on central IT to support them, quickly culminating in serious bottlenecks due to the sheer number of requests that pour in. Lastly, the focus on cross-functional processes requires data to be accessible from different domains—indicating a need to open up data across the entire enterprise.

To leverage this strong analytical skillset, SAP identified data products as a critical enabler for their data democratisation initiative. As the company's Chief Strategy & Operating Officer explained: 'Data products pave way for standardized and well-governed access to and sharing of data assets, ensuring a certain level of control on the data, while promoting its diffusion throughout the entire organization'. With the data democratisation initiative, SAP aims to enable employees to request and retrieve the needed data product quickly and accurately, allowing them to perform analytics for various use-cases. Moreover, data products will reduce the cognitive load on data consumers by eliminating the need for time-consuming tasks, such as pre-processing and cleaning the data. Employees can focus more on building meaningful analytics and generating valuable insights for the enterprise.

To drive this vision across the entire enterprise, SAP has defined 13 domains, each representing critical business capabilities to be enriched by AI and data insights. These domains are designed to group, enable and manage key data assets for the key business capabilities in a consistent manner, ensuring the successful delivery of SAP products and services on a global scale. SAP has committed to building data products in each of their 13 domains, not only to serve internal needs but also to support cross-domain processes. To guide the design of effective data products across all domains, SAP has identified four common aspects: *mindset change* to align varying perspectives (mainly among business stakeholders, data and analytics teams and IT experts), *catalogue* to find the available data products, *semantics* to ensure a coherent understanding of data products and *consumption* to expedite reusability across various use-cases.

The data product canvas was identified as a tool that integrates all four key aspects into the design of data products. By addressing the *desirability* and *feasibility* aspects of the data product, the business users, data and analytics teams and IT experts can align their perspectives early in the design. For instance, the data product consumer relationships block defines standardised modes of access for data products, that is, using a catalogue as a discovery tool for existing data products. The key data object block, on the other hand, captures all data sources and types required to build the data product, simultaneously outlining its definition and primary purpose as a way to improve semantics. Finally, the data product consumers block captures the different (or similar) use-cases where the data product could be (re)used, and the data product delivery block highlights how such reusability would be made possible from a provisioning perspective. Figure 3 summarises key SAP challenges and how the data product canvas plays a role in their data democratisation initiative.

4.1 | Setting Up to Work With the Data Product Canvas at SAP

Currently, SAP is in the exploration phase with data products. Over the past few months, numerous data product requests have been collected from the different domains. However, not all data product ideas can be realised, posing the key challenge of deciding which ideas to pursue. The head of each business unit from which the requests came evaluated the proposals based on their alignment with the organisation's broader strategic goals.

For the prioritised data products, the design process was facilitated by a team called Democratised Governance. This team brought together representatives from the responsible domain (business areas), as well as from the data and analytics and IT teams, to collaboratively complete the data product canvas. To accommodate participants located across different SAP offices worldwide, separate MS Teams meetings were organised

and supported by a Miro board. On average, completing each canvas took approximately 1 h. Afterward, the team conducted a 'reflection round' to discuss the challenges faced during the exercise and their key implications. Table 1 presents four exemplary data products that have been designed and assessed using the data product canvas and were eventually developed and deployed within SAP for practical use. They are:

- Account master data (DP1): A dataset containing information on SAP's customer accounts, which is critical for most of the processes existing within the company. This data product is highly reusable as it is offered in a foundational form, that is, ready-to-use dataset. By productising account master data, it is expected to feed into various descriptive and prescriptive analytics use-cases.
- Customer support case (DP2): A dataset containing information on various support cases registered by customers at the support centre. This data product would be required to support use-cases in multiple areas: such as customer services to quickly resolve pending issues, operations to improve efficiency of ticket resolution and sales to offer new or updated products or services.
- Data quality mission control centre (DP3): A dashboard to monitor and control the quality of customer address data, which can be sourced from the account master data product (DP1). Hence, this product is relevant for several domains, such as finance, for sending invoices to the correct location and sales, for ensuring products are delivered to the right address.
- Function and department engine (DP4): A machine learning model that matches job titles with their respective departments in order to help with creating targeted marketing campaigns. Despite being more use-case specific, at the very least, this data product must be shared with the marketing domain as well as the data management team.

4.2 | Designing Data Products With the Data Product Canvas—The Experiences at SAP

The starting point of the exercise was the *value proposition* block. Here, two key questions were discussed: What is the business problem the data product is intended to address, and what value is expected to be generated in return? Hence, the participants had to reflect on the main business issue(s) and possible outcomes from solving them. They found it more challenging than expected: 'Normally, we don't think about, for instance, what a dataset could bring in for the business. But I guess we need to do some brainwork from now on' (Domain representative for DP2, SAP). 'Well, this is hard! Let me think of some possible indicators to report value' (Domain representative for DP1, SAP).

This reflected the difficulty practitioners experience in answering the 'why' question. Being able to outline why their data product is needed will first accentuate its alignment with vital business objectives, highlighting a potential business case for its existence. One could think of it as an elevator pitch, where domains succinctly outline the key value of their data products to secure funding from senior management. Not only will this help key stakeholders see what concrete problems could be addressed by the data product but also make the case for its further adoption by new users across the enterprise stronger. The exercise eventually proved useful because it triggered creativeness among the participants: 'As a standalone data product, I think DP2 would have minimal contribution but now that I think about it, may be its interesting to actually aggregate this with the ticketing data to get a much richer picture of customer cases' (Domain representative for DP2, SAP). As such, the participants were able to detach themselves from their 'own little world' and reflect broadly on why their suggested data product makes sense from an enterprise-wide perspective. This aspect of value blending sparked engaging discussions and inspired participants to consider collaborating with colleagues to 'combine forces.' This approach not only promotes reuse of existing resources within SAP but also facilitates cross-domain conversations—an essential first step in bridging data silos.

Naturally, the next question from the *desirability* perspective is: Who is going to generate this value? A list of *data product consumers* was compiled, categorised mainly by their role and team, to identify those expected to regularly interact with the data product. Additionally, the consumers were classified as either internal or external to SAP. This distinction is critical, as it determines how access, security and privacy are configured, even for internal consumers. Subsequently comes the 'how' question. The obvious use-cases were captured first, but the participants were further encouraged to ponder additional use-cases that the same data product could address. The more usage scenarios a data product can handle, the higher its potential for reusability—a critical aspect for any data product (Blohm et al. 2024; Desai et al. 2022). For instance, during the exercise on DP4, the participants uncovered that the same matching algorithm could possibly be trained in order to map titles and departments in languages other than English. This extension could enable SAP to tap into potential leads in other countries, expanding opportunities for higher sales. For *data product customer relationship and delivery*, the Collibra data catalogue has become the de-facto platform at SAP for discovering data products. The primary reason for this choice is its stable connectors with the SAP technology stack. Regarding support services, the participants agreed that the level of support required depends on the specific data product types. As a dashboard, for instance, DP3 was guaranteed to have a 99% uptime, whereas as a dataset, DP1 was promised 100% accuracy and completeness. For more detailed questions on the data products, their respective owners remained available.

Next, the participants shifted their focus to the central element of building the data product—the data itself—and the *feasibility* perspective. For the *key data objects*, most of the data for the products came from internal SAP systems, such as SAP S4/HANA or SAP Master Data Governance. As a result, data acquisition remained a fairly simple process. However, depending on the use-cases supported by the data products, significant portions of data had to be sourced from external sources. For instance, with DP4, a large amount of unstructured data was collected to fine-tune the parameters, despite possible quality challenges: 'We collected data from online forms as well as from marketing events to train the ML

TABLE 1 | Exemplary data products at SAP.

Data product canvas building block	Account master data (DP1)	Customer support case (DP2)	Data quality mission control centre (DP3)	Function and department engine (DP4)
Value proposition	Establishment of single-source-of-truth to enable reusability of accounts data	Consolidation of customer case data to conduct advanced analytics	Transparency and control of address data quality	Smart segmentation of clients to facilitate targeted ads
Data product consumers (teams and use-cases)	Sales team: *Segmentation of the customers* Finance team: *Integrated financial reporting* Data operations team*: Improve and monitor the quality of the dataset*	Data scientists*: Forecasts on various operational metrics* Operational team*: Faster and efficient resolution of customer tickets*	Data (process) owners*: Conduct deduplication* Data service bureau and data advisors*: Adoption support to business users* Finance team*: Create, pay and manage invoices*	Data management team*: Manual mapping of clients' job title to departments* Marketing team*: Produce targeted ads and generate new leads*
Data product consumer relationship	Findable in data catalogue 24/7 availability of the dataset 100% completeness guaranteed Multiple updates per day	Findable in data catalogue Technical and user document on the attributes and definitions Near real-time updates of case data	Findable in data catalogue Trainings for users Link to report bugs 24/7 FAQs and user documentation on SharePoint	1-to-1 meetings to showcase data product Newsletter Model accuracy guaranteed at app. 80%
Data product delivery	As a dataset Login through browser using SAP intranet	As a dataset Login through browser using SAP intranet	As a dashboard or Excel file Login through browser using SAP intranet	As ML model Web browser and mobile device login using SAP intranet
Key data objects	Business partner data	Customer case data	Address data	Job designation and contact data
Key activities	Collect priorities from the business; collect, clean and transform data; build and deploy dataset; formulate a roadmap and deploy in catalogue	Replicate the data to the data foundation; transform and harmonise the data; develop and test stability of the dataset; deploy for usage	Brainstorm data quality rules with stakeholders; create and test MVP; build final product; collect feedback and enhance rules and parameters	Develop business case; evaluate product feasibility; check data availability and quality; develop and test MVP; deploy in cloud and go-live
Key resources	Intelligent data management team; data owner; enterprise architects; platform specialist	Data product owner; developer; enterprise architects	Analytics platform team; delivery team; IT product owner; data product manager	Data owners; business SMEs; marketing experts; data quality and governance expert

(Continues)

TABLE 1 | (Continued)

Data product canvas building block	Account master data (DP1)	Customer support case (DP2)	Data quality mission control centre (DP3)	Function and department engine (DP4)
(People and systems)	Collibra; SAP MDG	S/4 HANA; SAP datasphere; ServiceNow; salesforce	SAP Analytics Cloud; SAP InfoSteward	Azure for ML; Marketo; SAP BW and CRM
Cost structure	Full-time employees; Collibra licence; third-party data source; Maintenance cost	Collibra licence; development cost; time effort in maintenance and monitoring	7–8 full-time employees; time effort during sprint review	Azure hosting cost; full-time employees; time effort for manual mapping
Benefits	Higher adoption of the data product; improved decision making based on high-quality data	Aggregate with other data products; improved consumption rate of customer case data	Reduce data duplication report effort from 7 h to 2 min; reduce time for processing tickets; Reduce invoice bounce rate	Reduce 80 K/EUR running cost per year; automation of matching process

model properly, expecting it would do better matching' (Data Scientist, SAP).

Given the diversity of data, it must be cleaned, modelled, aggregated and transformed into the desired data product. Hence, there was strong consensus among the participants on adhering to the SAP data product lifecycle to guide these activities. Consisting of six steps, it provided a good framework to structure the *key activities* across specific phases. It also ensures that consumers are aware of the current stage of the data product and the expected timeline to go-live. When outlining tasks based on the lifecycle phases, we observed a natural discussion on the *key resources* required for each phase, categorised in either technical or personnel resources. Additionally, some participants took the opportunity to reflect on their own role in this context: 'Although I am taking the owner role for DP4 now, but it is secondary to my main job and I think we should probably acquire some additional resources to support us as the number of data products will grow' (Data Scientist, SAP). Based on such reflections, the participants also managed to identify and ponder on current resource gaps or possible scarcities that may impede the successful implementation of the data product. Overall, the most frequently mentioned technical resources included the Collibra data catalogue, which is critical for making data products more findable and accessible. Platforms such as Microsoft Azure were required for storage and analytics, particularly for DP4 that leveraged Azure's machine learning capabilities. Additionally, SAP Datasphere was also identified as a key platform. In terms of personnel, platform specialists (from IT department) were required to manage the Collibra data catalogue, data owners/stewards for providing access to the required data, as well as data architects to build consistent and high-quality data models to be used by everyone (from data and analytics team).

Eventually, it came down to demonstrating tangible returns from the data product, addressing the *viability* perspective. Simply put, the participants had to define the expected costs and benefits. In particular, the discussion of the *cost structure* led to lively debates on who is going to pay for it. Although the domains normally create or capture the data, the data and analytics team is responsible for transforming it into data products and the IT team is responsible for storing and delivering it to the entire organisation. Hence, a shared approach is probably more suitable. The usage of the canvas does not resolve this debate but rather opens up the possibility to further discuss and create a costing model that all parties can agree on. The costs associated with the data products arise from personnel costs for employees dedicated to various stages in the data product lifecycle, expenses for hosting and licencing needed to store and access the data product and monitoring costs to perform continuous health checks on the data product. With regard to *benefits*, the participants shared both monetary and non-monetary factors that would act as key selling points to obtain sponsorship for the data product. The most commonly identified benefits included savings in manual work, increased data product adoption rate—a critical indicator for SAP's data democratisation initiative—and improvements in data quality. To put some numbers to it, DP3, for instance, significantly reduced the processing time for creating a report on address data deduplication from 7 h to 2 min by connecting the dashboard to the data source and refreshing the data periodically. Similarly, DP4 led to operational cost savings of 80 K EUR per year.

5 | Recommendations to Practitioners for Applying the Data Product Canvas

The strength of the data product canvas lies in its visual, easy-to-understand format, making it a practical tool that helps cross-functional teams—comprising business, data, analytics and IT experts—collaboratively design data products. Product-thinking around data is broadly applicable across industries and a wide range of organisational contexts, including small and medium-sized enterprises. As a versatile tool, the data product canvas guides the development of successful data products by emphasising consumer-centricity (*desirability*), the alignment of consumer requirements with provider capabilities (*feasibility*) and the formulation of a compelling business case to secure funding (*viability*). Based on our experience, firms of any size or type can therefore leverage the canvas as a de-facto tool to instantiate and embed product-thinking, whether as part of enterprise-wide data initiatives or in targeted data, analytics and AI-related projects.

Through applying the data product canvas in SAP's data democratisation initiative, we identified three important recommendations related to its use that offer valuable guidance for future adopters. Table 2 summarises the recommendations, including their rationale and practical guidelines for effective use of the data product canvas. Each recommendation is further illustrated with examples from the SAP case.

5.1 | Recommendation 1: Tailoring the Design to Different Data Product Types

Data products can take various forms—such as datasets, dashboards, or machine learning models—and must be shaped, packaged and delivered in the most fitting way to fulfil end-users' information needs effectively. To address this diversity, we recommend tailoring the design to the specific type of data product and propose two patterns for applying the data product canvas in this process, as illustrated in Figure 4.

For the most basic form of data products (e.g., SAP's DP1 and DP2), practitioners should start the design from the left side of the data product canvas, namely, the *feasibility* theme. These data products are source-aligned, meaning they are closely tied to their originating data sources and are intended to make data reusable for different types of data consumers and across multiple use-cases. Starting with *feasibility*, practitioners should assess how existing data is provisioned and whether it can be integrated into comprehensive, ready-to-use datasets for wider organisational use. Next, they should deliberate on the *desirability* theme and identify opportunities to increase reuse across different teams and their use-cases. By uncovering these prospects, they can ensure discussions on how to appropriately package the data product and integrate it into numerous downstream systems that have varying technical capacities. For example, DP2 aggregates ticketing data that is originally stored in ServiceNow and scattered in other legacy systems, where it had been captured inconsistently. A data product offering an easy-to-understand overview of the customer cases can support multiple customer-facing teams and help enhance issue resolution, enable operational efficiency and drive customer service improvements.

For analytical and advanced analytical data products (such as SAP's DP3 and DP4), practitioners are advised to begin the design on the right side of the data product canvas, that is, the *desirability* theme. These data products are consumer-aligned, designed to fulfil specific information needs of end-users and must be built to create direct business value. Here, the practitioners should prompt critical questions such as: Who are the key data product consumers? What use-cases will the data product support? How do consumers prefer the data product to be delivered and supported? Having a clear idea about the key use-cases, including any overlaps and knowing how the consumers want the data product delivered from the onset is essential for designing a value-creating data product. After clarifying these questions, the focus should shift to how the data product can be efficiently developed and delivered with high quality—corresponding to the *feasibility* theme. If the practitioners realise that there are too many similar data products or the key consumers cannot be identified, they may decide against investing in the data product—even if it is feasible and accurate and reliable data exists for it. For instance, with DP3, the initial concern was whether an additional customer dashboard would be worthwhile, given the presence of similar dashboards. Upon further exploration, it became clear that the time and effort for creating a dedicated data quality dashboard focusing on customer address data would be well spent due to compelling business reasons, such as correct invoicing.

5.2 | Recommendation 2: Reassessing and Redesigning Data Products

Data products naturally evolve over time, as underlying data sources, processing methods and analytical models are subject to continuous change. These changes are further accelerated by technological advancements, emerging demands for novel insights and evolving functional and non-functional requirements. Hence, data products must be designed with adaptability and scalability in mind. To support practitioners in navigating rapidly evolving requirements and technological capabilities, we recommend using the data product canvas not only during the initial design phase but also as a tool for periodically reassessing and, if necessary, redesigning data products once they have been deployed and are in use.

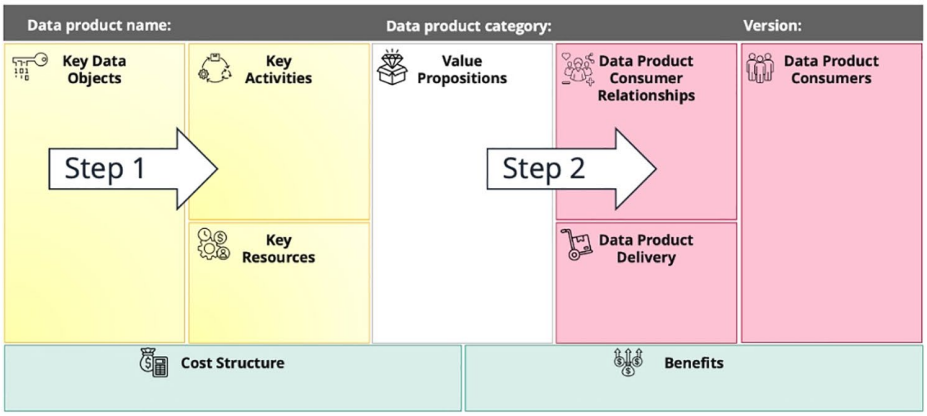
To do so, practitioners should revisit the data product design using the canvas and revise the relevant blocks to reflect any changes—resulting in an updated version of the data product, along with corresponding updates to its documentation with the canvas. Reassessing data products periodically initiates a learning process and encourages a reflection on how data products can be further improved and scaled to increase their value. This is specifically important as many data products begin as a proof of concept with a restricted scope and, if proven successful, are gradually expanded and rolled out to a larger user base. Moreover, it ensures that evolving business needs, technical capabilities and quality standards are reflected in the design, enabling data products to remain relevant, valuable and scalable over time. Figure 5 illustrates the evolution of DP4, originally developed as a local solution (Version 1.0) for marketing teams focused on lead generation and later expanded into a global solution (Version 2.0). The initial version used an ML model for

TABLE 2 | Recommendations for applying the data product canvas.

Recommendations (what)	Rationale (why)	Guidelines (how)	Examples from SAP
Tailoring the design to different data product types	• Data products have different drivers and complexities which impact the design process • Consumers operate in different domains requesting different types of data products, such as datasets, metrics, dashboards, or machine learning models	• For basic data products, such as a dataset, start from '*feasibility*' to discuss how the data products support multiple use-cases. At the next step, shift focus towards '*desirability*' followed by '*viability*' • To address specific needs with analytical insights, start from '*desirability*' to better grasp the user perspectives and requirements. Then, the practitioner should deliberate on the '*feasibility*' and then '*viability*'	• DP1 and DP2 are basic data products built on data coming from several internal sources. Hence, the first check was done on the availability and quality of the data (*feasibility*) • For DP3 and DP4, that provide more aggregated and specific insights, the first task was to understand the user base and how a data product could support their needs (*desirability*)
Reassessing and redesigning data products	• Data products and their outputs must evolve to accommodate continuous changes in the underlying data sources, processing methods or analytical models • Business priorities as well as technical capabilities change over time which impacts the way data products are managed and used	• Reassess the data products periodically based on the canvas and then update the documentation with the latest changes • Create updated versions of the data product canvas as new requirements are incorporated and then implement these changes into the design of data products	• DP3 focuses on 26 attributes for address data currently but the next version aims to incorporate the new details from the account data • For DP4, an updated version of the canvas is produced with every agile iteration, helping create a transparency of the modifications made
Building a cohesive data product portfolio	• An increasing number of new use-cases require the creation of numerous data products • High demand for limited resources poses challenges in supporting all data products within the organisation • Ensuring alignment of data products with key business goals is essential for sustainable value generation	• Use the canvas to document and register data product requests (new data products, enhancements of existing ones and productisation of existing data assets) • Derive decision-making criteria directly from the canvas by drawing on its three central themes: number of consumers and use-cases, synergy between use-cases (*desirability*), number of data sources and availability of resources (*feasibility*), cost vs. benefit estimates (*viability*)	• SAP business users co-create the canvas with data, analytics and IT experts to document initial requests • Various criteria are derived based on the canvas blocks to manage portfolio selection: (a) Is the data product request assigned an owner? (*feasibility*), (b) What are expected costs and benefits? (*viability*), (c) Will it support enough use-cases to be valuable? (*desirability*)

For basic data products
(highly source-aligned):

1. Start by assessing the sources and data availability, combine them to address recurring information needs
2. Increase data reuse of the data product by extending the consumer base with adapted packaging and delivery



For analytics and advanced analytics data products
(consumption aligned):

1. Start by understanding the consumers' information needs and use cases
2. Find efficient ways to build the product by creating synergies and ensuring high quality

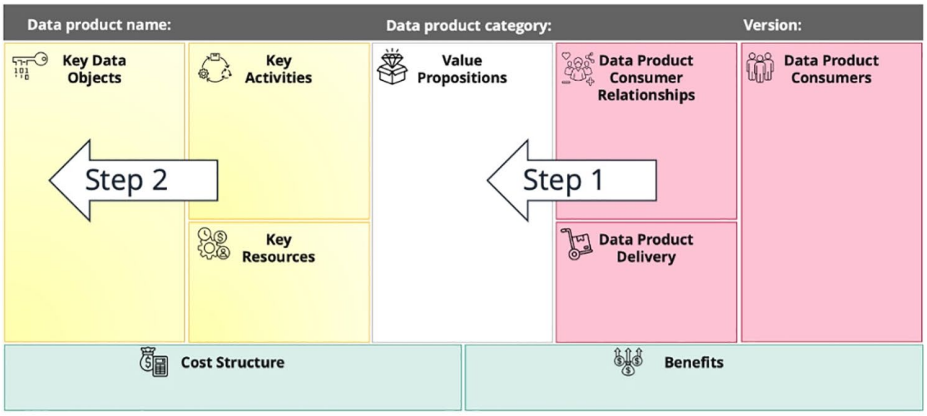


FIGURE 4 | Using the data product canvas for different data product types.

mapping job titles and departments in English, with data collected from customer contact forms and events. In Version 2.0, the data product will be transformed into a standardised global solution to be deployed in other SAP local offices worldwide. This global expansion requires mapping job titles and departments in non-English languages and involves purchasing and integrating third-party contact data to enhance coverage and accuracy. As the number of contact data increases, the data management team could emerge as a new user group, responsible for ensuring data quality and enabling reusability across other teams.

5.3 | Recommendation 3: Building a Cohesive Data Product Portfolio

As the number of data products created and utilised across the enterprise grows, it is critical to manage them in a structured and controlled manner, ensuring they remain relevant, effective and capable of consistently delivering the expected value. To build a cohesive data product portfolio, we recommend systematically using the canvas to manage the data product idea funnel and to collect, assess and prioritise data product requests. This approach comprises three steps, as illustrated in Figure 6.

As a first step, registering and documenting the incoming requests with the canvas ensures that all critical perspectives—*desirability*, *feasibility* and *viability*—are comprehensively considered. Typical requests may involve creating new data products, enhancing existing ones, or productising existing data assets, ultimately generating a pipeline of ideas for potential investments, either immediately or at a later point in time. As requests are submitted to the data product idea funnel independently from different teams, some of these ideas may overlap, closely resemble one another, or even be duplicates.

In a second step, the standardised documentation enables a structured comparison of incoming requests and facilitates the identification of opportunities for consolidation. Using the data product canvas documentation, requests are assessed against a set of criteria that reflect the three core dimensions and similar or overlapping data product requests can be combined into a unified data product (as shown as the intersecting dotted lines in Figure 6). This consolidation minimises redundancies between data products and enhances resource utilisation during their creation, whereas enhancing the reusability of the resulting data products. Under the *desirability* theme, we recommend evaluating the expected number of users for the data product, the number and range of use-cases it addresses and the synergies between those use-cases. These criteria help gauge the business value the data product will contribute to the portfolio. Under the *feasibility* theme, the availability and number of required data sources must be considered, as they determine the type and number of input ports needed to ingest data into the data product and highlight major technical dependencies. Additionally, the availability of key resources, both systems and personnel, is a key criterion to determine whether the data product can be effectively managed throughout its lifecycle. Finally, under the *viability* theme, practitioners should estimate the costs

and benefits and evaluate whether the requested data product will result in a net gain or loss.

As the final step, practitioners should consolidate the assessment results into a prioritised list of final requests, ready to advance to the subsequent development and deployment stages. Through this structured approach, the most appropriate data products are selected for the portfolio—ensuring they provide the highest value for business users (as data product consumers) while enabling efficient development, delivery and maintenance by data, analytics and IT teams (as data product providers). Unselected requests are archived for future consideration, creating a well-documented pipeline of forward-looking data product ideas for the organisation.

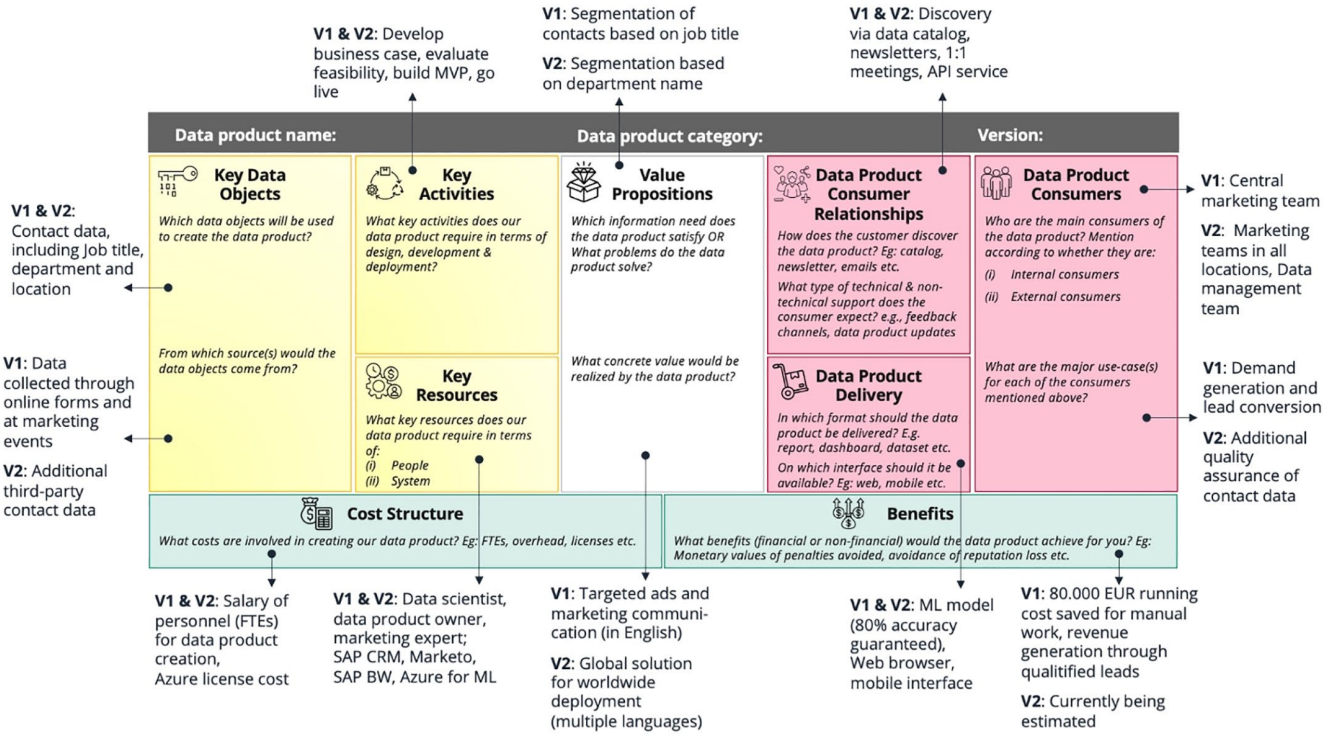


FIGURE 5 | Incremental enhancements of a data product—function and department engine (DP4).

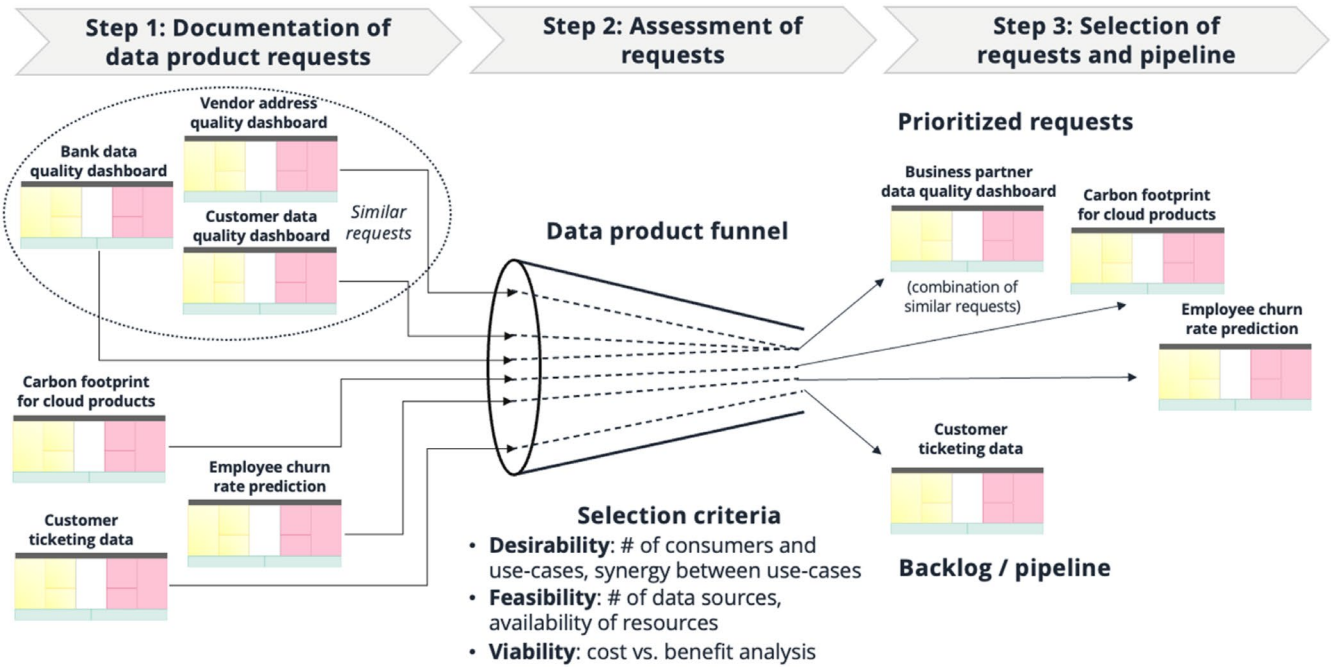


FIGURE 6 | Managing the data product idea funnel using the data product canvas.

6 | Conclusion and Future Direction

As companies strive to become more data-driven and democratise data, they must establish a standardised and scalable approach to provisioning data, whereas addressing the evolving information needs from an increasing number of users. Data products play a crucial role in this by combining, packaging and delivering data and insights into a consumable form so that they can be easily accessed and (re)used across the organisation. The data product canvas addresses the lack of methodological guidelines and tools to effectively develop data products, and as a visual and versatile tool, it enables cross-functional teams to collaboratively design and assess data products while ensuring consideration of three critical themes: *desirability* to evaluate the business need for the data product, *feasibility* to assess the data requirements and technical capabilities needed for its creation and delivery and *viability* to estimate the potential returns from the data product. Thereby, the canvas brings data consumers and data providers together, harmonising their perspectives on data products and ensuring alignment of requirements and capabilities. It facilitates vital early-stage conversations—clarifying key questions around its business purpose, potential implementation approaches and required resources—and helps achieve consensus about the data product vision between important stakeholders, particularly in business, data and analytics and IT.

From the experiences at SAP and other companies, applying the canvas not only supports experts in collaboratively designing well-defined data products but also facilitates enterprise-wide data initiatives by highlighting underlying capabilities. As the canvas helps identify various organisational elements of a data product, such as roles and responsibilities (key resource block), lifecycle-related tasks (key activities block) as well as literacy and training needs (data product consumer relationships block), practitioners can organise these aspects to steer effectively the creation and consumption of data products. For managers, the insights provide valuable inputs for addressing weaknesses and help scale data, analytics and AI across the organisation. To succeed in this endeavour, managers must also recognise that a data product mindset requires the development of data product management at two levels: the single-data-product level and the portfolio level. The canvas serves as a starting point by facilitating the assessment of individual data product requests, whereas supporting portfolio decisions and guiding the organisation's shift from a single-data-product towards a multi-data-product perspective. However, its effective application must be complemented by additional processes, guidelines and new responsibilities: data product managers to oversee the lifecycle of individual data products, and a domain or portfolio manager to manage the data product portfolio within a specific business domain. This underscores the transformational nature of adopting a data product approach.

Acknowledgements

We would like to thank the experts from SAP for their excellent collaboration without which this case study would not have been possible. We would also like to acknowledge the Fortune500 companies who co-innovated with us for the past 3 years on data products and contributed to the development of the data product canvas as part of our collaboration in the Competence Center for Corporate Data Quality (CC CDQ).

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request. The Competence Center Corporate Data Quality (CC CDQ) has made the Data Product Canvas available under a creative commons license at the following link.

References

Assur, N., and K. Rowshankish. 2022. "The Data-Driven Enterprise of 2025. McKinsey." https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-data-driven-enterprise-of-2025.

Avdiji, H., D. Elikan, S. Missonier, and Y. Pigneur. 2020. "A Design Theory for Visual Inquiry Tools." *Journal of the Association for Information Systems* 21, no. 3: 695–734. https://doi.org/10.17705/1jais.00617.

Bean, R. 2022. "Why Becoming a Data-Driven Organization Is So Hard." *Harvard Business Review*. https://hbr.org/2022/02/why-becoming-a-data-driven-organization-is-so-hard.

Blohm, I., F. Wortmann, C. Legner, and F. Köbler. 2024. "Data Products, Data Mesh, and Data Fabric." *Business and Information Systems Engineering* 66, no. 5: 1–10. https://doi.org/10.1007/s12599-024-00876-5.

Davenport, T. H., and R. Bean. 2022. "Developing Successful Data Products at Regions Bank." *MIT Sloan Management Review*. https://sloanreview.mit.edu/article/developing-successful-data-products-at-regions-bank/.

Davenport, T. H., and S. Kudyba. 2016. "Designing and Developing Analytics-Based Data Products." *MIT Sloan Management Review* 58, no. 1: 83–89.

Davenport, T. H., R. Bean, and R. Wang. 2024. *CDO Agenda 2024: Navigating Data and Generative AI Frontiers*, 1–24. Amazon Web Services.

Davis, J., D. Nussbaum, and K. Troyanos. 2020. "Approach Your Data With a Product Mindset." *Harvard Business Review*: 1–6. https://hbr.org/2020/05/approach-your-data-with-a-product-mindset.

Dehghani, Z. 2021. *Data Mesh—Delivering Data-Driven Value at Scale*. Vol. 1. First ed. O'Reilly Media.

Desai, V., T. Fountaine, and K. Rowshankish. 2022. "A Better Way to put Data to Work." *Harvard Business Review* 100, no. 4: 100–107.

Hasan, M. R., and C. Legner. 2023a. "Understanding Data Products: Motivation, Definition, and Categories." In *Proceedings of the Thirty-First European Conference on Information Systems (ECIS 2023)*, 1–17. AIS eLibrary. https://aisel.aisnet.org/ecis2023_rp/229.

Hasan, M. R., and C. Legner. 2023b. "Data Product Canvas: A Visual Inquiry Tool Supporting Data Product Design." In *Design Science Research for a New Society: Society 5.0. DESRIST 2023. Lecture Notes in Computer Science*, Vol. 13873, edited by A. Gerber and R. Baskerville. Springer, Cham. https://doi.org/10.1007/978-3-031-32808-4_12.

O'Neill, B. T. 2023. "110—CDO Spotlight: The Value and Journey of Implementing a Data Product Mindset With Sebastian Klapdor of Vista." https://designingforanalytics.com/resources/episodes/110-cdo-spotlight-the-value-and-journey-of-implementing-a-data-product-mindset-with-sebastian-klapdor-of-vista/.

Osterwalder, A., and Y. Pigneur. 2010. *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*. Wiley.

Peffers, K., T. Tuunanen, M. A. Rothenberger, and S. Chatterjee. 2007. "A Design Science Research Methodology for Information Systems Research." *Journal of Management Information Systems* 24, no. 3: 45–77. https://doi.org/10.2753/MIS0742-1222240302.

Schulze, K., and K. Tiwary. 2024. *Shifting Mindsets: Why You Should Treat Data as a Product*. Thoughtworks. https://www.thoughtworks.com/insights/e-books/modern-data-engineering-playbook/data-as-a-product.

Wang, R. Y. 1998. "A Product Perspective on Total Data Quality Management." *Communications of the ACM* 41, no. 2: 58–65. https://doi.org/10.1145/269012.269022.

Appendix 1

Development of the Data Product Canvas

The design of the data product canvas started in late 2021 as part of the Competence Center Corporate Data Quality, an industry research consortium working in the field of data management. For this project, we collaborated with 30 practitioners from 15 global companies from different industries who are members of this consortium. These practitioners are highly experienced in the topic of data and analytics, typically with 5–10+ years of experiences in the field and represent firms that are currently experimenting with or implementing data products. Furthermore, they were also willing to share concrete details about their organisational context, challenges and expectations with regards to data products. Leveraging our initial focus group conducted on February 2022, we collected various challenges currently persisting within the organisations related to data products. They are as follows:

- Data products have been primarily viewed as a technical topic, overlooking its alignment with key organisational goals and the specific business challenges they are expected to address.
- Data consumers across the enterprises build their own tailored solutions to address each of their use-cases, underestimating the effort required to maintain them and undermining the potential of reusability which data products can bring.
- Data product providers and consumers often have different perceptions and understanding of what a data product entails, mainly in terms of its functionalities and implications, leading to conflicting views on how it should be designed.
- Persisting challenges in estimating ROIs from data products to demonstrate their tangible value, impeding the opportunities to secure funding from organisational sponsors.

With these challenges in mind, the practitioners agreed that the potential solution should aim to bridge the primary gaps around the concept of data products and support its proper design. Hence, the solution was decided to be a canvas that would support the design of data products by bringing together business users (who will consume the data product) and data, analytics and IT experts (who will provide the data product).

To develop the data product canvas, we conducted three design iterations in May 2022 (conceptual design), September 2022 (refinement of building blocks) and November 2022 (instructions for using the canvas). We built the canvas by incorporating the design principles for visual inquiry tools into its design. To ensure alignment with practitioner needs, we adopted the design science research (DSR) method (Peffers et al. 2007) and conducted multiple iterations including focus groups, demonstrations and expert evaluations by following the guidelines provided by the authors. The focus groups were conducted both onsite and online (using MS Teams and were recorded by getting permission from the participants), running roughly for 1.5–2 h. It focused on the various design aspects of the canvas and how they can be further improved to address existing challenges. During the design, the canvas was also used for documenting a variety of data products, initially from three different companies and at later stages with 22 data products from 17 other companies that are both within and outside our research setting. Furthermore, we conducted expert evaluations with not only the study participants but also with external audiences (consultants from a data management firm and participants at a data product-related industry summit), using surveys where they rated various design elements of the canvas and also provided their reflections, allowing us to gain further credibility. Based on this feedback, a stable version of the canvas was established (Hasan and Legner 2023b) and has been publicly available under a creative common licence at the following link. Since 2023, it has been taught in several training programmes on data products at executive and graduate levels, with an increasing user base across various types of organisations, including small and medium-sized enterprises. It has been adopted by several large companies who use it in their data product initiatives and have made it an integral part of their standard operating procedures for data product design.