

Data Mesh - Jenseits des Hypes

Die Probleme von heute stammen von den »Lösungen« von gestern.
Peter M. Senge, The Fifth Discipline

Sascha Lorenz
lorenz@psg.de

ZIELE DIESER SESSION

WAS HABEN WIR VOR?

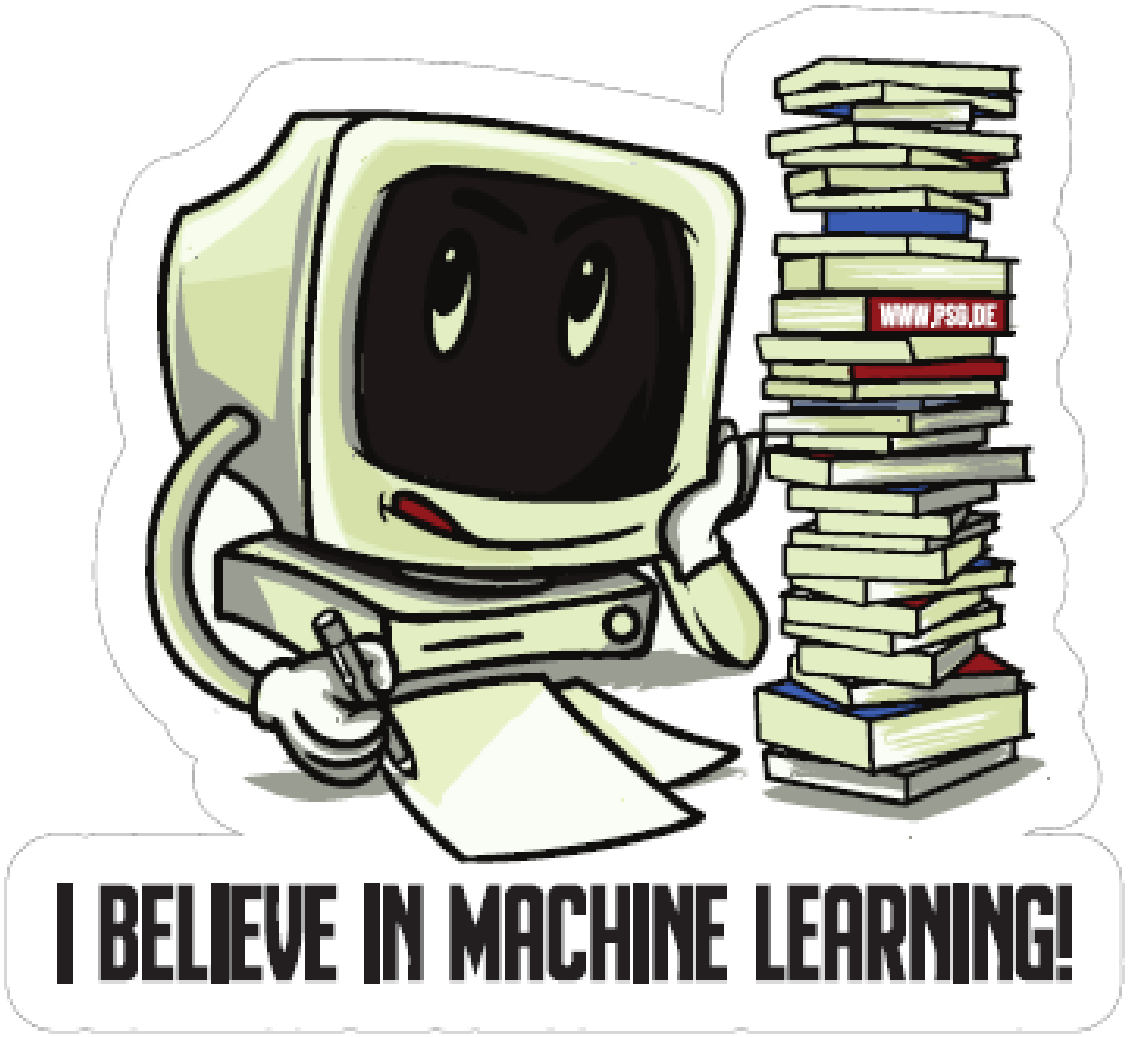
- Prämissen des Vortrags:
  - Keine Buchzusammenfassung. Sorry!
  - Schlüsselkonzepte
  - Was können wir übernehmen?
  - Herausforderungen

- Ziele für das Publikum
  - Tieferes Verständnis der praktischen Aspekte und Herausforderungen der Data Mesh Architektur
  - „No-Hype“ Entscheidungsfindung, ob es denn ein Data Mesh sein sollte…
  - Leitplanken für die Implementierung in Ihrem spezifischen Unternehmenskontext, sofern sinnvoll.

Über mich (Sascha)



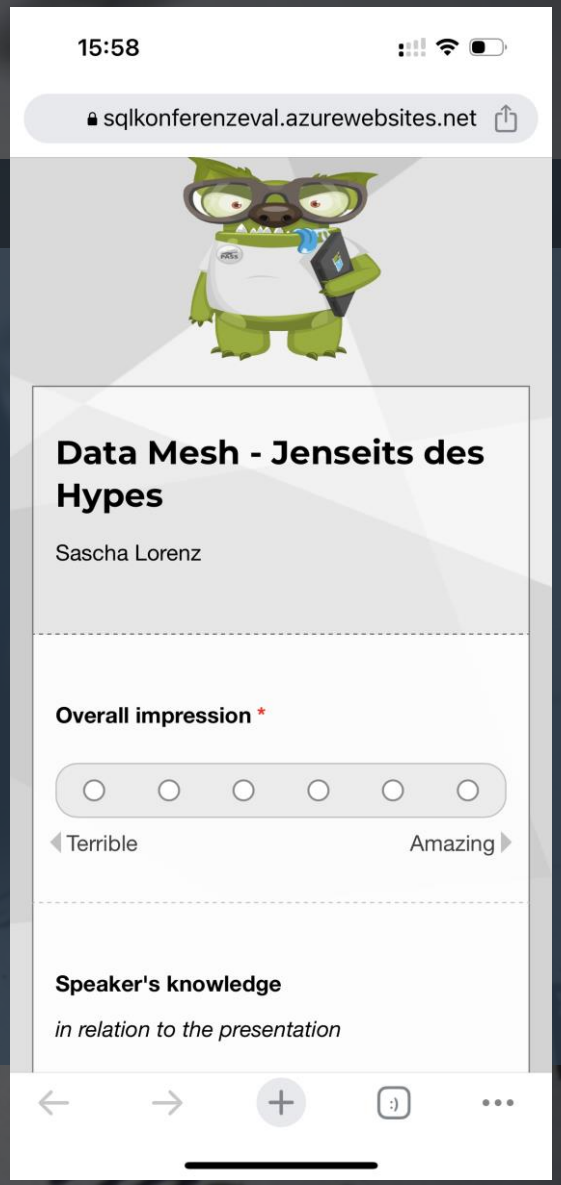
- Sascha Lorenz ein SQL Server Lead Architect, Trainer und manchmal auch Berater in DACH-Raum. Er hat eine Leidenschaft für komplexe Datenbankumgebungen und die Vereinfachung ihrer Bedienung.
- Er begann seine Karriere mit dem Schreiben eines Tabellenkalkulationsklons für Commodore C64 und später eines objektorientierten Pre-Compilers für Nantucket Clipper (dbase III+ Compiler). Seit 1994 ist er Miteigentümer der PSG Projekt Service GmbH, auch bekannt als "The SQL Server Company".
- Sascha ist seit 2005 aktiv in der deutschen SQL Server Community PASS Deutschland e.V., wo er verschiedene Funktionen innehat und seither weit über 100 Community-Gruppentreffen in Hamburg organisiert hat. Er wurde dreimal mit dem MVP Award ausgezeichnet.
- Derzeit ist er Programm-Manager für den PSG Tool Stack bei der PSG Projekt Service GmbH. Zu seinen Aufgaben gehören das Trainingsframework PDD für agile Datenbankentwicklungsprozesse und die Entwicklung des PSG Tool Stacks, mit dem seine Kollegen und Kunden weltweit Datenbankserver analysieren und optimieren.
- Seitdem (2019) sein Unternehmen Go als primäre Entwicklungssprache einsetzt, engagiert sich Sascha auch in der Hamburger GDG (Google Developer Group) und ist seit Oktober 2023 Mitorganisator des GCP & Go Meetup.



- If you are interested in this one, and there is no more available here?
- Just ask me!

FEEDBACK




15:58
sqlkonferenzeval.azurewebsites.net
Data Mesh - Jenseits des Hypes
Sascha Lorenz
Overall impression *
Terrible
Amazing
Speaker's knowledge
in relation to the presentation

DATA MESH


O'REILLY
Deutsche Ausgabe
Data Mesh
Eine dezentrale Datenarchitektur entwerfen
Zhamak Dehghani
Vorwort von Martin Fowler
Übersetzung von Jochen Christ und Simon Harrer

WAS IST DATA MESH?

- Ein Data Mesh ist ein soziotechnischer Ansatz für die Erstellung einer dezentralen Datenarchitektur.
- Ein Ziel dabei ist die Förderung einer Innovationskultur. ☺

- Welche Technologien?
- Welche(s) Cloud Produkt(e)?
- Ist das (nur) ein Architekturansatz?

- Und überhaupt…?

WAS IST DATA MESH?

Auf den Punkt gebracht:

Meines Erachtens geht es darum, klassische „Engpässe“ in der analytischen Datenverarbeitung und -aufbereitung durch den Ansatz einer Dezentralisierung zu reduzieren!

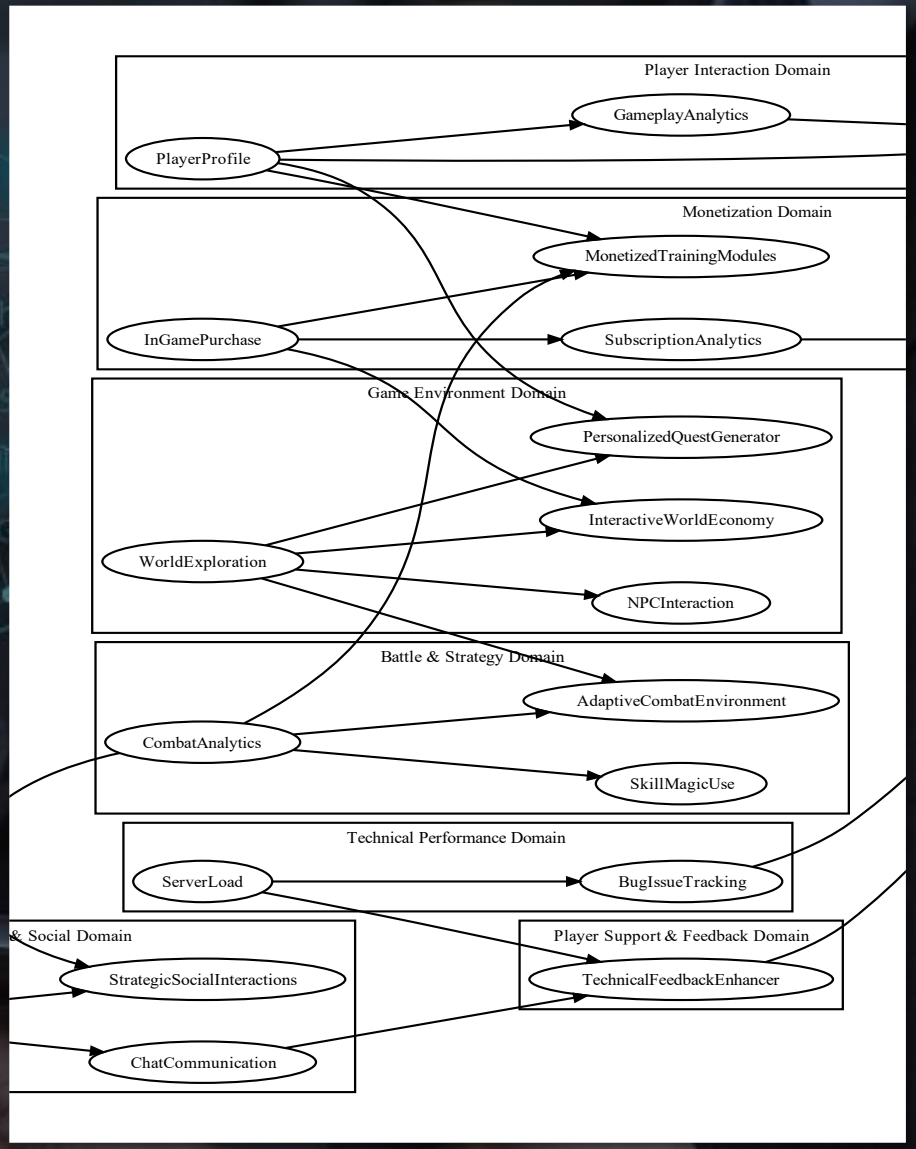
Der durch eine Dezentralisierung naturgemäß entstehenden Komplexität wird durch den Einsatz moderner Infrastruktur, Monitoring-/Observability-Tools und standardisierter Entwicklungsprozesse (CI/CD etc.) begegnet.

WAS IST DATA MESH?

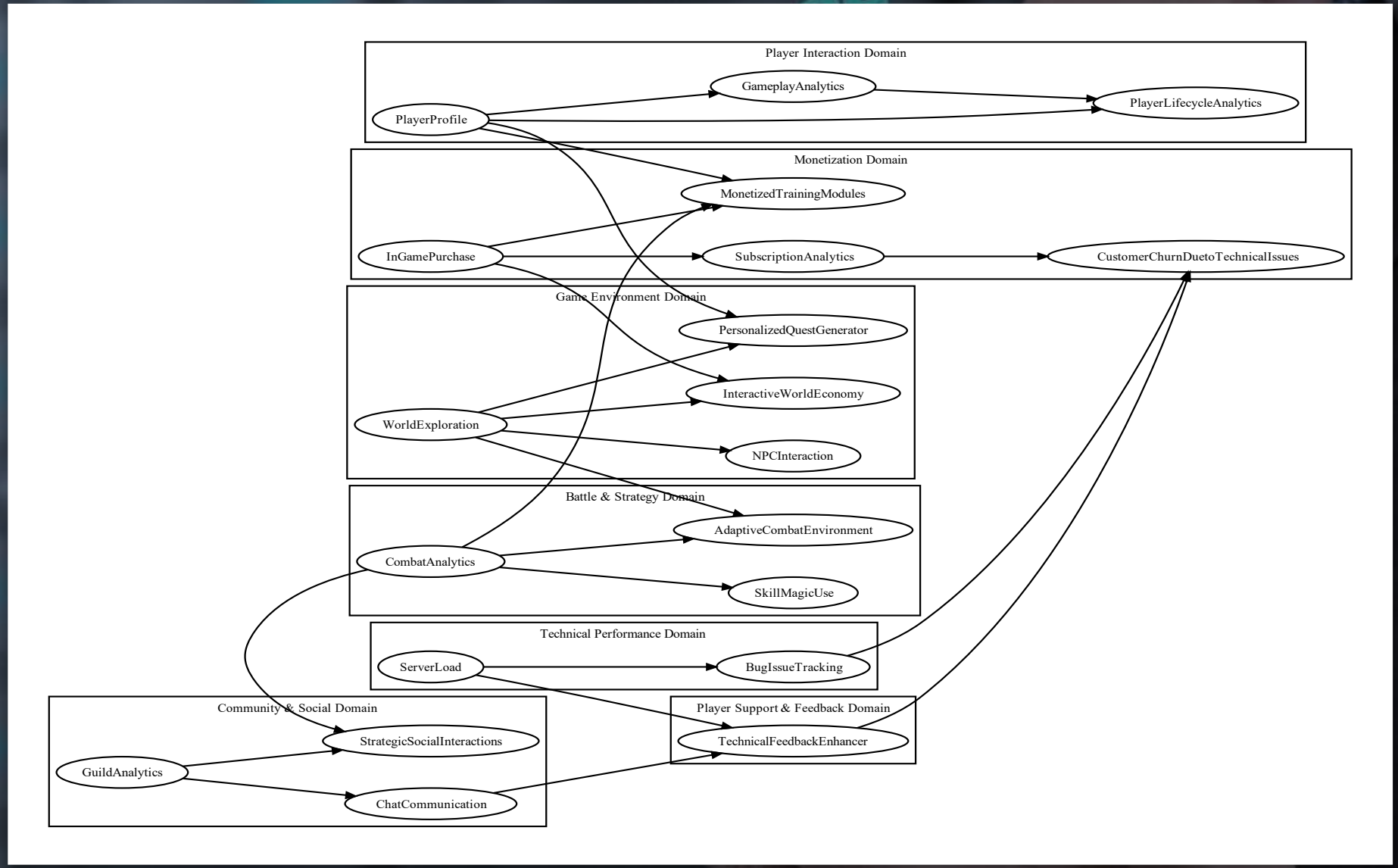
BEISPIEL “DATA DUNGEONS”

"Data Dungeons" ist ein Marktführer im Bereich Online-Rollenspiele. Als das Unternehmen expandierte, wurden seine zentralisierten analytischen Datensysteme zu Engpässen.

Um diese Probleme zu lösen, wurde eine Data Mesh-Architektur umgesetzt.



WAS IST DATA MESH?



WAS IST DATA MESH?

"4 SÄULEN"

- Domain Ownership
- Data as a Product
- Self-Serve Data Platform
- Federated Governance

WARUM IST DATA MESH WICHTIG?

- **Skalierbarkeit:**
  Traditionelle Data-Warehouses oder Data Lakes stoßen in der modernen, datengetriebenen Welt schnell an ihre Grenzen. Data Mesh bietet einen skalierbaren Ansatz.
- **Agilität:**
  In einer schnelllebigen Geschäftswelt ist die Fähigkeit zur schnellen Anpassung essentiell. Data Mesh fördert Agilität durch dezentrale Data Ownership und schnellere Time-to-Insight.
- **Data Governance:**
  Die Data Mesh-Architektur ermöglicht eine effektivere Datenverwaltung und -qualität durch den Einsatz von Federated Data Governance.
- **Demokratisierung von Daten:**
  Statt Daten in Silos zu sperren, ermöglicht Data Mesh den einfachen Zugang und die gemeinsame Nutzung von Daten über verschiedene Geschäftsbereiche (aka Domänen) hinweg.

TRADITIONELLE DATENARCHITEKTUREN UND IHRE LIMITIERUNGEN

- Zentralisierte Datenlager (Data Warehouses, Lakes etc.)
  - Limitierung: Hohe Kosten, schwierig zu skalieren, Single Point of Failure
- Daten-Silos
  - Limitierung: Schwierige Datenintegration, mangelnde Konsistenz
- Monolithische Architekturen
  - Limitierung: Nicht flexibel, schwierig zu erweitern und zu warten
- ETL/ELT-Prozesse (Extract, Transform, Load)
  - Limitierung: Zeitverzögerung, Komplexität steigt mit der Datenmenge
- Top-Down Governance
  - Limitierung: Trägheit bei Anpassungen, Flaschenhälse in der Entscheidungsfindung

HAUPTKOMPONENTEN/SÄULEN EINES DATA MESH

- Domain-Driven Design (Domain Ownership)
  - Jedes Geschäftsfeld oder jede Abteilung fungiert als eigener "Datenproduzent".
  - Beispiel: Vertrieb, Marketing, Produktion.
- Data Product Management (Data as a Product)
  - Daten werden als Produkte betrachtet, für die ein spezielles Management erforderlich ist.
  - Beispiel: Datenkatalog, Daten-Qualitätsmaßnahmen.
- Federated Data Governance
  - Datenmanagement und -zugriff werden dezentralisiert, aber durch gemeinsame Richtlinien geregelt.
  - Beispiel: Meta-Daten-Management, Datenschutz-Richtlinien.
- Self-Serve Data Platform
  - Daten Produkte können von allen Domänen gefunden und unkompliziert genutzt warden.
  - Beispiel: “Portal Lösungen”

HAUPTKOMPONENTEN/SÄULEN EINES DATA MESH

1. Organisationsstruktur
   - Traditionell: Zentralisierte Data-Teams sind für Datenverwaltung, -speicherung und -zugriff verantwortlich.
   - Data Mesh: Dezentrale, domainorientierte Teams kümmern sich um ihre eigenen Daten als "Data Products".
2. Datenverwaltung
   - Traditionell: Daten werden oft in einem großen Data Lake oder Data Warehouse gespeichert.
   - Data Mesh: Daten sind in verschiedenen Domains verteilt und werden als eigenständige Produkte behandelt.
3. Governance
   - Traditionell: Zentrale Governance für alle Daten.
   - Data Mesh: Federated Governance, bei der die Verantwortung für die Datenqualität und -sicherheit auf mehrere Teams verteilt ist.
4. Skalierbarkeit
   - Traditionell: Skalierung oft schwierig und teuer, da alles zentral verwaltet wird.
   - Data Mesh: Einfachere Skalierung durch Verteilung der Verantwortlichkeiten und Ressourcen.
5. Flexibilität & Anpassungsfähigkeit
   - Traditionell: Änderungen sind oft langsam und erfordern die Koordination mehrerer Teams.
   - Data Mesh: Schnellere Anpassung an neue Anforderungen dank dezentraler, agiler Strukturen.

DOMAIN-DRIVEN DESIGN IM DATA MESH

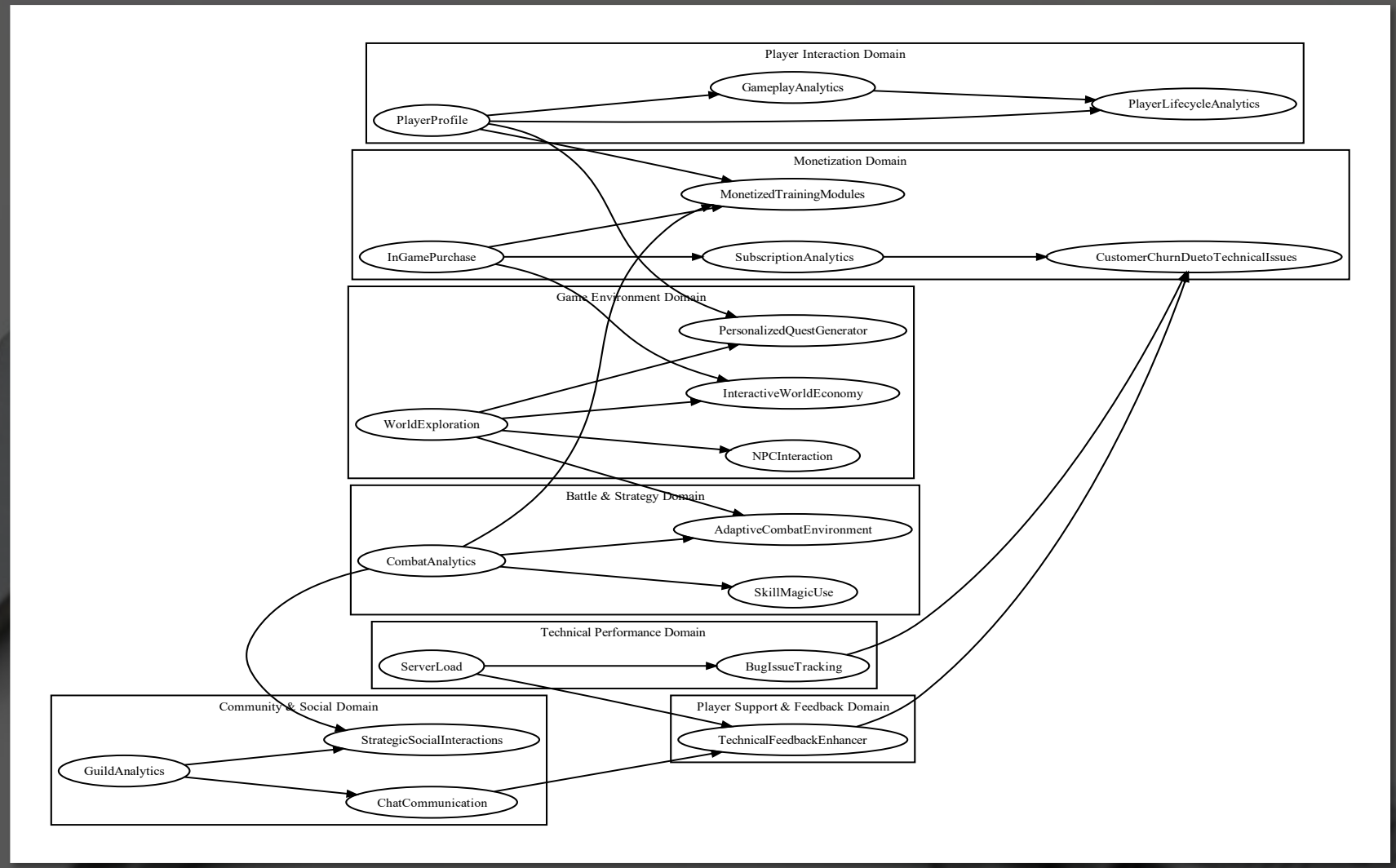
WAS IST DOMAIN-DRIVEN DESIGN UND WARUM IST ES WICHTIG FÜR DATA MESH?

- Domain-Driven Design (DDD) fokussiert sich auf das spezielle Geschäfts- oder Tätigkeitsgebiet (die "Domäne"), um eine Softwarelösung zu entwickeln, die eng mit der Geschäftslogik und den Fachexperten abgestimmt ist. Durch die Verwendung einer „Ubiquitous Language“ wird eine einheitliche Fachsprache in Code, Design und Kommunikation geschaffen. Mit taktischen und strategischen Entwurfsmustern wird die Komplexität der Domäne in beherrschbare, klar definierte Softwarebausteine zerlegt.

- Im Data Mesh-Ansatz wird DDD verwendet, um eine gemeinsame Sprache und ein einheitliches Verständnis für Daten-Domänen zu schaffen, die die Datenerzeugung und -nutzung in einem verteilten System organisieren. Durch die Anwendung von DDD können komplex verwobene Datenstrukturen und -beziehungen in klar definierte, domänenspezifische Bausteine zerlegt werden, wodurch die Datenhandhabung vereinfacht und die Zusammenarbeit zwischen verschiedenen Teams verbessert wird.

(siehe Martin Fowler)

DOMAIN-DRIVEN DESIGN IM DATA MESH



DOMAIN-DRIVEN DESIGN IM DATA MESH

VORTEILE UND HERAUSFORDERUNGEN VON DOMAIN-DRIVEN DESIGN IN DATA MESH

- **Wie DDD hilft, Domains und autonome Data Products zu sehen:**
  Domain-Driven Design ermöglicht es, die Geschäftsdomäne in klar abgegrenzte Sub-Domänen zu zerlegen, die als autonome Data Products entwickelt und verwaltet werden können. Dies fördert die Eigenständigkeit der Teams und ermöglicht eine schnellere Anpassung an sich ändernde Geschäftsanforderungen.

- **Warum DDD bei der Federated Governance hilfreich ist:**
  DDD unterstützt eine federierte Governance, indem es für jede Domäne spezifische Regeln und Richtlinien bietet, die jedoch im Kontext eines übergreifenden Systems kohärent sind. Durch die klare Definition von Domänen und deren Verantwortlichkeiten kann eine effektivere und dezentralisierte Steuerung der Daten und Services erreicht werden.

- **Mögliche Herausforderungen:**
  Die Anwendung von DDD kann zu Beginn komplex erscheinen, da ein tiefes Verständnis der Geschäftsdomäne und der dazu gehörigen Prozesse erforderlich ist. Darüber hinaus erfordert die erfolgreiche Implementierung von DDD Fachkenntnisse sowohl im Bereich der Geschäftslogik als auch in technischen Aspekten der Softwareentwicklung, was eine enge Zusammenarbeit zwischen Fachexperten und Entwicklern voraussetzt.

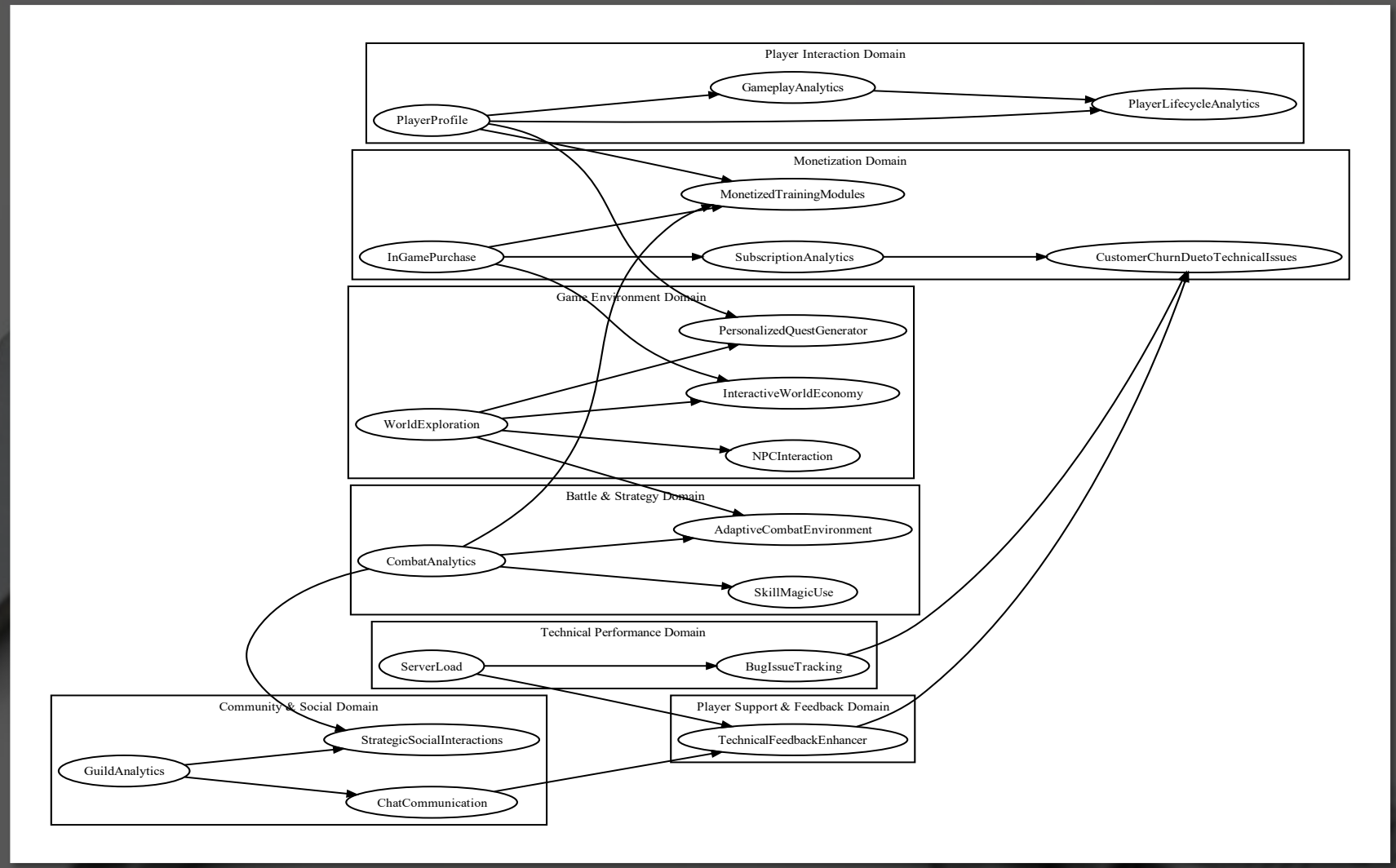
DATA PRODUCT (MANAGEMENT)

WAS IST DATA PRODUCT MANAGEMENT UND WARUM IST ES WICHTIG?

- **Definition**:
  Ein „Data Product“ ist eine selbsterhaltende, autonome Datenlösung, die eine spezifische Geschäfts- oder Anwendungsfunktion erfüllt und einen klaren Mehrwert bietet.

- **Merkmale eines Data Products**
  - **Selbstbeschreibend**:
    Daten und deren Bedeutung sind klar und ohne zusätzliche (externe) Dokumentation verständlich.
  - **Qualitätsgesichert**:
    Ständige Überwachung und Qualitätssicherung zur Gewährleistung der Datenintegrität.
  - **Zugänglichkeit**:
    Daten können leicht von anderen Teams oder Systemen abgerufen und genutzt werden.

DOMAIN-DRIVEN DESIGN IM DATA MESH

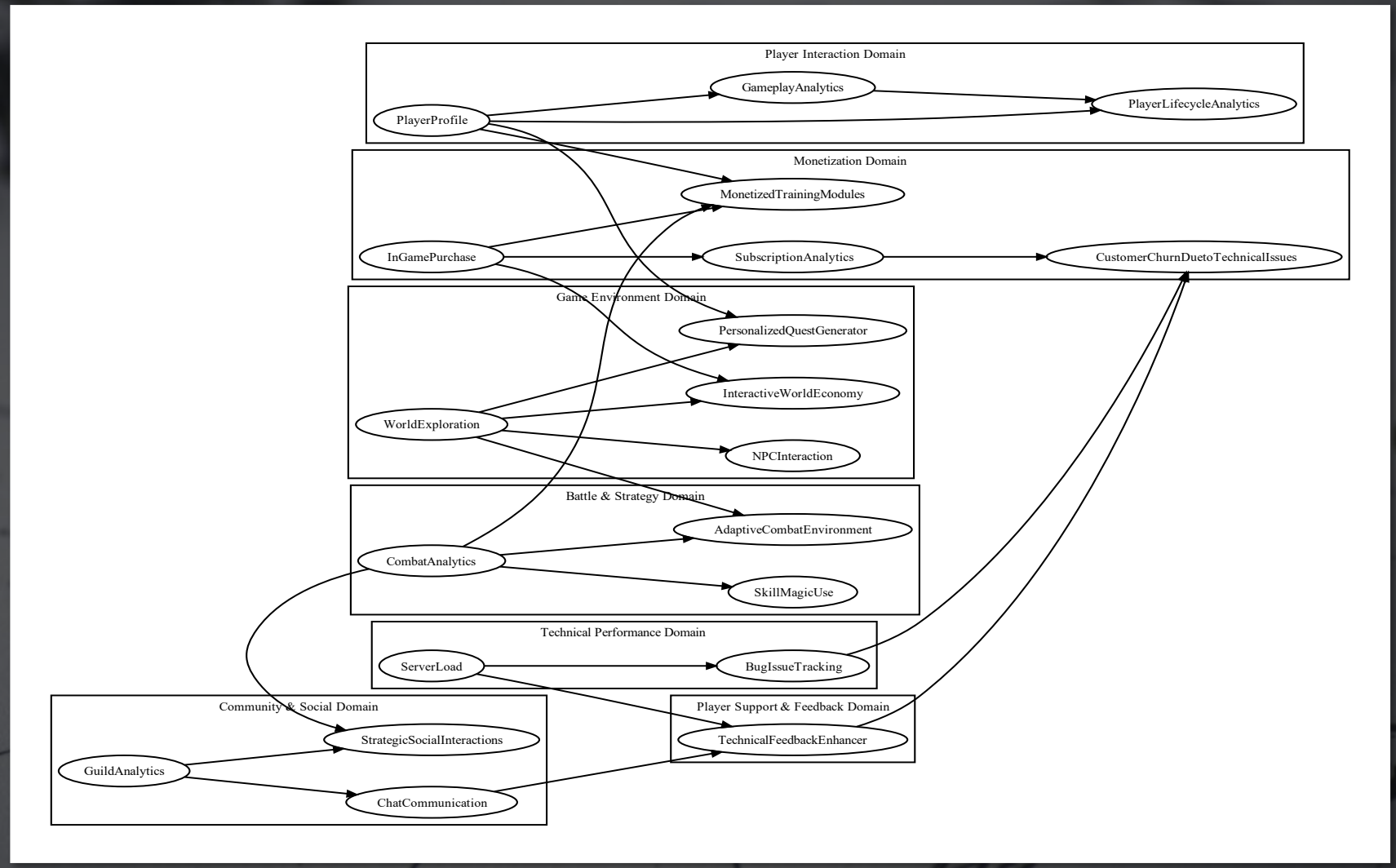


DATA PRODUCT (MANAGEMENT)

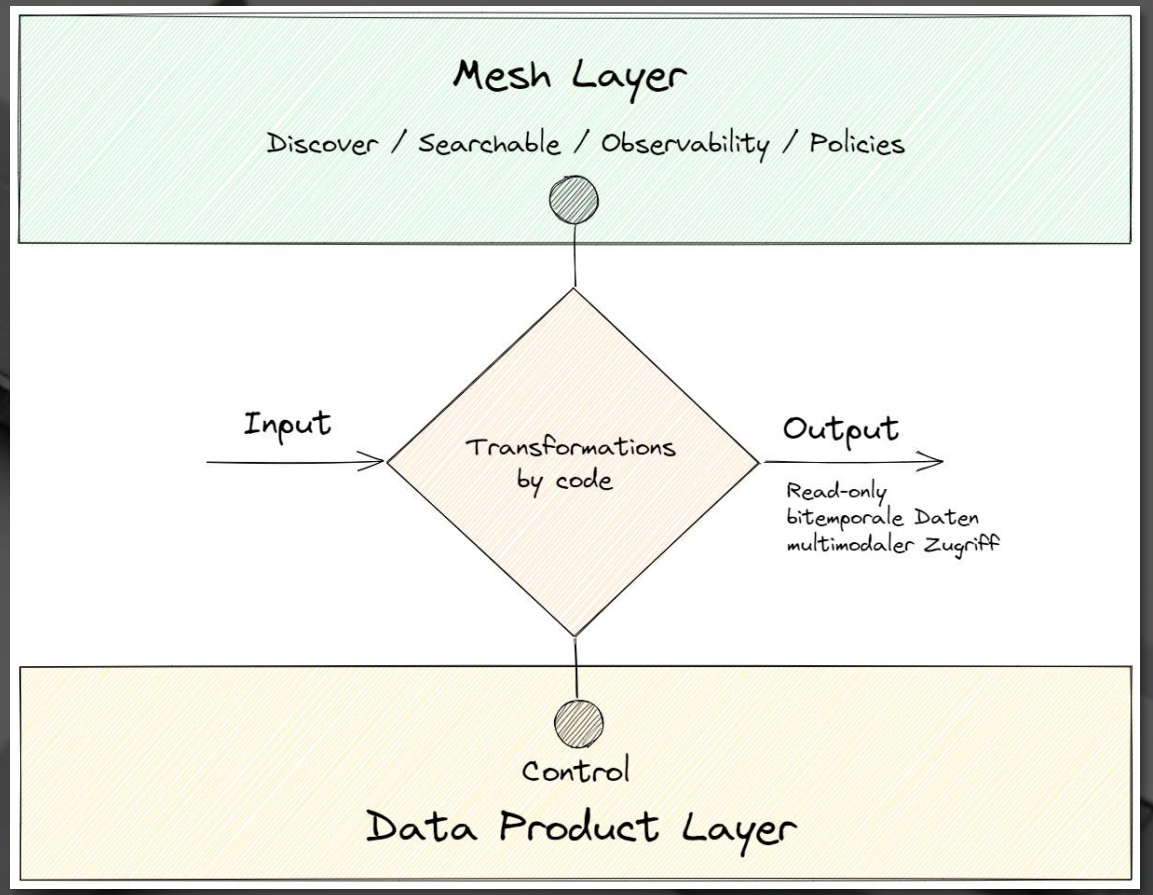
„DATENARTEN“

- **Source-aligned domain data:**
  Dieser Begriff bezieht sich auf Daten, die direkt aus ihrer ursprünglichen Quelle stammen und im Kontext dieser Quelle am relevantesten sind. In einem Data Mesh werden diese Daten in der Regel von den Teams verwaltet, die für die Quellsysteme verantwortlich sind. Das Ziel ist es, eine Grundlage für die Datennutzung zu schaffen, die die Komplexität und den Kontext der Datenquelle berücksichtigt.
- **Aggregate domain data:**
  Hierbei handelt es sich um Daten, die aus verschiedenen Quellen aggregiert werden, um eine breitere Sicht auf ein bestimmtes Geschäfts- oder Anwendungsdomäne zu bieten. Das Ziel ist, interdisziplinäre Erkenntnisse zu fördern und komplexere Fragen zu beantworten, die über die Grenzen einzelner Quellsysteme hinausgehen.
- **Consumer-aligned domain data:**
  Diese Daten sind speziell auf die Bedürfnisse eines bestimmten Verbrauchers oder einer Gruppe von Verbrauchern innerhalb der Organisation ausgerichtet. Sie sind oft eine spezialisierte Ansicht oder Transformation von Quell- oder aggregierten Daten, die speziell dafür entwickelt wurde, bestimmte Anwendungsfälle oder Geschäftsprozesse zu unterstützen. Das Ziel ist, den Datenverbrauch so einfach und effizient wie möglich zu gestalten, indem Daten in einer Form bereitgestellt werden, die direkt für den Endverbraucher nützlich ist.

DATA PRODUCT „CONTAINERS“



DATA PRODUCT MANAGEMENT



DATA PRODUCT MANAGEMENT

EIGENSCHAFTEN DES „OUTPUTS“ EINES DATEN PRODUKTES

- **Read-only**
  - **Beschreibung**: Der Output des Data Products ist ausschließlich lesbar und nicht veränderlich.
  - **Vorteil**: Sicherstellt Datenintegrität und verhindert unbeabsichtigte Manipulationen oder Veränderungen.
- **Bitemporale Daten**
  - **Beschreibung**: Daten, die zwei Zeitdimensionen besitzen – eine für den tatsächlichen Zeitpunkt des Geschehens (Gültigkeitszeit) und eine für den Zeitpunkt der Erfassung im System (Transaktionszeit).
  - **Vorteil**: Ermöglicht eine präzise Historie sowohl der realen Weltveränderungen als auch der Systemveränderungen.
- **Multimodaler Zugriff**
  - **Beschreibung**: Daten können über verschiedene Modalitäten und Schnittstellen abgerufen werden, zum Beispiel als RestAPI, GraphQL, EventStream, Parquet Files etc. (und, nicht entweder!)
  - **Vorteil**: Erhöht die Flexibilität und Zugänglichkeit der Daten für unterschiedliche Anwendungen und Benutzerprofile.

FEDERATED DATA GOVERNANCE

WAS IST FEDERATED DATA GOVERNANCE?

- **Definition von Federated Data Governance:**
  Federated Data Governance ist ein Ansatz zur Datenverwaltung, bei dem die Verantwortung für Datenqualität, Sicherheit und Zugriffsberechtigungen über verschiedene Abteilungen, Domänen oder Geschäftseinheiten innerhalb einer Organisation verteilt wird. Dies fördert die dezentrale Entscheidungsfindung und ermöglicht es den einzelnen Einheiten, spezifische Datenanforderungen und -richtlinien zu definieren und zu verwalten.

- **Gegenüberstellung zu zentralisierten Modellen:**
  In einem zentralisierten Daten-Governance-Modell liegt die Verantwortung für alle Aspekte der Datenverwaltung bei einer zentralen Einheit oder Abteilung. Während dieses Modell eine konsistente Verwaltung und Kontrolle ermöglichen kann, kann es auch zu Engpässen, mangelnder Agilität und einer langsamen Reaktion auf geschäftliche oder technologische Veränderungen führen.

FEDERATED DATA GOVERNANCE

VORTEILE DER FEDERATED DATA GOVERNANCE

- **Skalierbarkeit:**
  Federated Data Governance ermöglicht eine einfache Skalierung, da sie die Verwaltung einer wachsenden Anzahl von Datenprodukten und -domänen unterstützt. Durch die Dezentralisierung der Datenverantwortung kann das System problemlos erweitert werden, ohne dass ein Flaschenhals bei der Datenverwaltung entsteht.

- **Agilität:**
  Dieser Ansatz ermöglicht eine höhere Agilität, da einzelne Teams die Flexibilität haben, unabhängig voneinander zu arbeiten und schnelle Entscheidungen zu treffen. Statt sich durch eine zentrale Steuerung aufzuhalten, können Teams lokal auf Veränderungen reagieren und Innovationen schneller umsetzen.

- **Eigenverantwortung:**
  Durch die Delegierung der Verantwortung für Datenqualität und -sicherheit an die Teams selbst wird ein höheres Maß an Datenpflege erreicht. Teams haben einen direkten Anreiz, die Qualität und Sicherheit der Daten zu gewährleisten, da sie direkt für die Ergebnisse ihrer Arbeit verantwortlich sind.

FEDERATED DATA GOVERNANCE

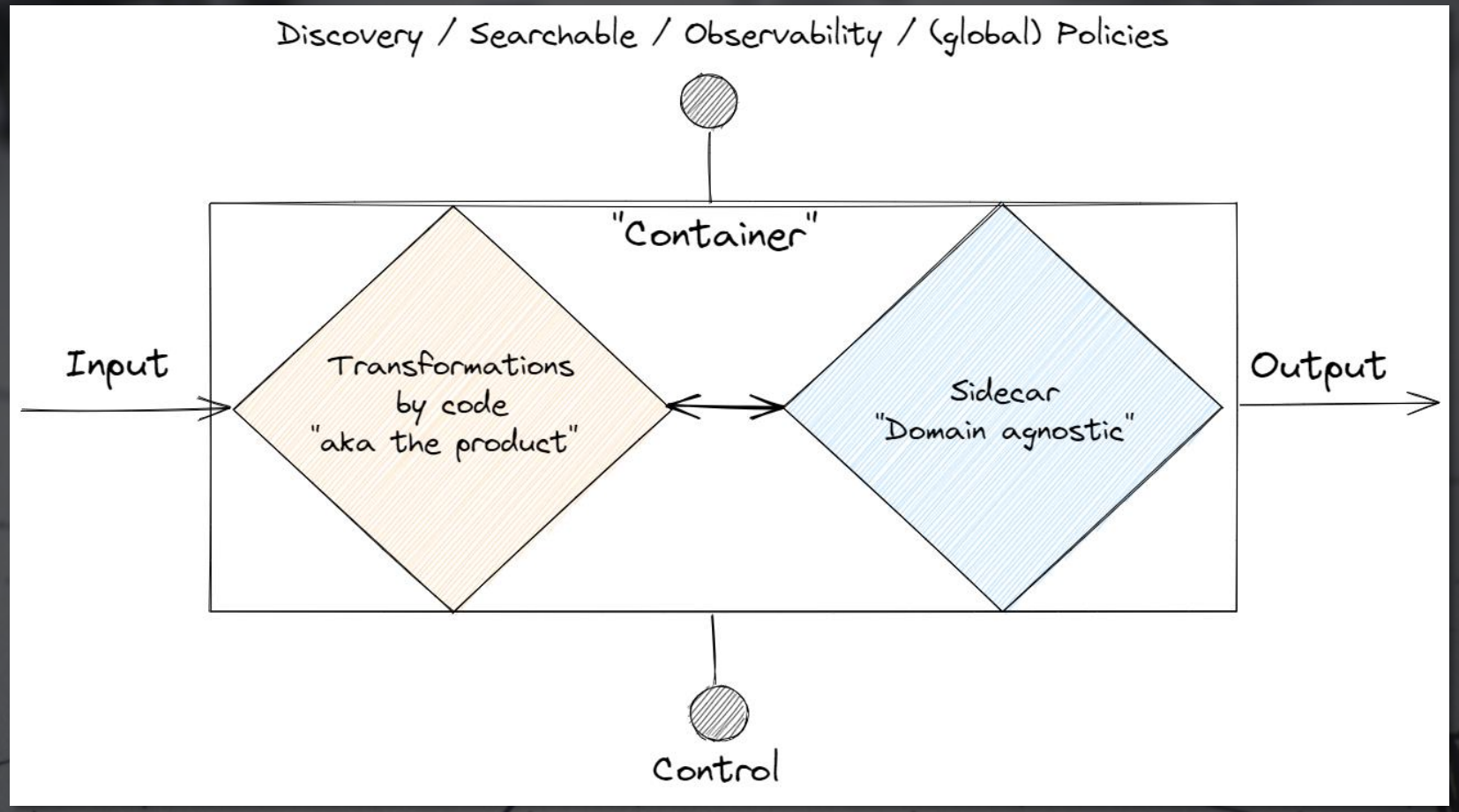
HERAUSFORDERUNGEN UND LÖSUNGSANSÄTZE

- **Komplexität:**
  In federierten Systemen mit mehreren Einheiten ergibt sich eine inhärente Komplexität bei der Koordination und Kommunikation. Eine zentrale Steuerung ist oft nicht praktikabel, daher müssen Mechanismen entwickelt werden, die eine effiziente, dezentrale Koordination ermöglichen.

- **Datensicherheit:**
  In einem föderierten Umfeld ist es unerlässlich, klare Verantwortlichkeiten für die Datensicherheit zu definieren. Ohne klare Zuweisung von Zuständigkeiten könnten Sicherheitslücken entstehen, die das gesamte System gefährden könnten.

- **Qualität:**
  Die Qualität der Daten muss über alle Domänen und Einheiten hinweg gewährleistet werden. Es bedarf daher einer übergreifenden Qualitätskontrolle, die sicherstellt, dass die Daten in allen Teilbereichen den gleichen Qualitätsstandards genügen.

DATA PRODUCT CONTAINER



FEDERATED DATA GOVERNANCE

VERWENDUNG DES SIDECAR ANSATZES

- **Ko-Deployment:**
  Das „Sidecar“ wird gemeinsam mit dem Hauptcontainer (z.B. Data Product Container) innerhalb derselben Orchestrierungseinheit (z.B. Docker Container, Kubernetes-Pod) deployed. Beide teilen Ressourcen und Lebenszyklus, sind jedoch modular voneinander getrennt.
- **Spezialisierte Funktionen:**
  Das Sidecar übernimmt spezielle Aufgaben wie Logging, Monitoring, Datenqualitätsprüfungen oder Sicherheitsmaßnahmen, die nicht direkt in den Hauptcontainer integriert sein müssen.
- **Modularität und Isolation:**
  Durch die Trennung der Verantwortlichkeiten in Sidecar und Hauptcontainer wird das System modularer und leichter wartbar. Teams können sich auf ihre spezifischen Fachgebiete konzentrieren.
- **Skalierbarkeit und Flexibilität:**
  In einer föderierten Data Mesh-Architektur können mehrere Data Product Container mit ihren jeweiligen Sidecars bereitgestellt werden, was die Skalierbarkeit und Anpassungsfähigkeit des Systems erhöht.

PRAKTISCHE ASPEKTE UND HERAUSFORDERUNGEN

RAHMENBEDINGUNGEN FÜR DIE IMPLEMENTIERUNG

- **Unternehmensgröße und –struktur**
  - Kleine vs. Mittelständische vs. Große Unternehmen
  - Zentralisierte vs. Dezentralisierte Strukturen
  - Spezifische Herausforderungen bei der Implementierung in verschiedenen Unternehmensstrukturen
- **Technologische Infrastruktur & Entwicklererfahrung**
  - Bestehende Systeme
  - Notwendige Technologien und Fähigkeiten für Data Mesh (Cloud-Infrastruktur, Microservices etc.)
- **Organisationskultur und –strategie**
  - Bedeutung der Datenkultur und datengetriebenen Entscheidungsfindung
  - Strategische Ausrichtung des Unternehmens (Wachstum, Konsolidierung etc.)
  - Notwendige Veränderungsmanagement-Strategien

IMPLIKATIONEN UND ZUKUNFTSAUSSICHTEN

IMPLIKATIONEN

- **Dezentralisierte Datenverantwortung:**
  Anstatt eines zentralisierten Data-Teams sind nun dezentrale Teams verantwortlich für die Erstellung, Wartung(, Vermarktung) und Qualität ihrer Daten Produkte.
- **Kultureller Wandel:**
  Unternehmen müssen den Wandel von datenzentrierten zu domainzentrierten Ansätzen fördern, wodurch die Datenverantwortung dezentralisiert wird.
- **Federated Governance:**
  Der Umgang mit Daten wird dezentralisiert, wobei jeder „Data Product Owner" für die Datenqualität, -sicherheit und -zugänglichkeit verantwortlich ist.
- **Technische Herausforderungen:**
  Eine verstärkte Betonung der Automatisierung, Skalierbarkeit und Self-Service-Tools wird notwendig.
- **Zusammenarbeit und Kommunikation:**
  Der Erfolg des Data Mesh hängt von der engen Zusammenarbeit zwischen Daten Produkt-Anbietern und anderen Geschäftseinheiten ab.

IMPLIKATIONEN UND ZUKUNFTSAUSSICHTEN

ZUKUNFTSAUSSICHTEN VON DATA MESH

- **Standardisierung und Best Practices:**
  Mit der Verbreitung von Data Mesh werden Industriestandards und Best Practices für dessen Implementierung und Management entstehen.
- **Erweiterte Tool-Landschaft:**
  Die wachsende Popularität von Data Mesh wird die Entwicklung neuer Tools und Plattformen für das Management und die Überwachung von Datenprodukten anregen. (nextdata.com?)
- **Integration von KI und maschinellem Lernen:**
  Data Mesh wird mit KI-Plattformen integriert werden, um datengetriebene Entscheidungsfindung zu unterstützen.
- **Breitere Akzeptanz:**
  Mit dem Erfolg früher Implementierungen wird Data Mesh in einer Vielzahl von Branchen und Unternehmensgrößen bzw. -strukturen angenommen werden.
- **Weiterentwicklung des Konzepts:**
  Wie bei allen „Technologien“ wird es Iterationen und Weiterentwicklungen des Data Mesh-Konzepts geben, um den sich ändernden Geschäftsanforderungen gerecht zu werden.

IMPLIKATIONEN UND ZUKUNFTSAUSSICHTEN

KRITIK UND EINSCHRÄNKUNGEN 1/2

- **Hohe Einstiegshürden:**
  Die Notwendigkeit, existierende zentralisierte Datenarchitekturen zu dezentralisieren, kann zu Anfang überwältigend sein und erfordert erhebliche Anfangsinvestitionen (und Nerven).
- **Risiko von Datensilos:**
  Wenn Teams nicht effektiv kommunizieren und Daten teilen, können Datensilos entstehen, die den Wert des gesamten Data Mesh mindern.
- **Fehlende Reife von Werkzeugen:**
  Viele bestehende Datenwerkzeuge sind nicht für dezentralisierte Datenarchitekturen ausgelegt. Dies kann zu Schwierigkeiten bei der Implementierung und Skalierung führen.
- **Kulturelle Herausforderungen:**
  Unternehmen, die traditionell zentralisierte Datenansätze verwendet haben, könnten Schwierigkeiten haben, den kulturellen Wandel zu einem dezentralisierten Modell zu vollziehen.

IMPLIKATIONEN UND ZUKUNFTSAUSSICHTEN

KRITIK UND EINSCHRÄNKUNGEN 2/2

- **Standardisierungsprobleme:**
  Ohne klare Standards und Best Practices können Teams unterschiedliche Ansätze zur Datenverwaltung verfolgen, was zu Inkonsistenzen führt. (Hotspot Security!)
- **Schwierigkeiten bei der Überwachung:**
  Mit vielen verteilen Datenquellen und -produkten kann die Überwachung und das Qualitätsmanagement eine Herausforderung darstellen. (Management Layer!)
- **Unklare Verantwortlichkeiten:**
  Obwohl das Konzept des Datenproduktbesitzers im Data Mesh zentral ist, kann es in der Praxis zu Unklarheiten darüber kommen, wer für was verantwortlich ist, insbesondere wenn es zu Problemen kommt.

HERAUSFORDERUNGEN & BEST PRACTICES

HERAUSFORDERUNGEN

- **Technologieabhängigkeit:**
  Technologieorientierte Unternehmen haben oft bereits umfangreiche IT-Infrastrukturen. Das Hinzufügen von Data Mesh kann Komplexität und Abhängigkeiten erhöhen.
- **Recruiting:**
  Entwickler, die sich mit aktuellen Technologien auskennen, sind schwer zu finden und werden oft nicht von Konzepten wie Data Mesh angezogen, da sie möglicherweise nicht "spannend" genug sind.
- **Integrationsaufwand:**
  Bestehende Daten- und IT-Systeme müssen sorgfältig mit dem Data Mesh Ansatz integriert werden.
- **Unterschiedliche Prioritäten:**
  Technologieorientierte Teams könnten sich auf Produktentwicklung und -innovation konzentrieren und weniger auf Dateninfrastruktur.
- **Nicht schlüsselfertige Produkte:**
  Im Gegensatz zu "fertigen" Cloud DWH/Lake Lösungen erfordert Data Mesh individuelle Entwicklung und Anpassung für jedes Datenprodukt, wobei ein (erheblicher) Eigenaufwand notwendig sein kann.

HERAUSFORDERUNGEN & BEST PRACTICES

HERAUSFORDERUNGEN/LEARNINGS - FEDERATED DATA GOVERNANCE

- **Zentrale vs. Dezentrale Governance:**
  Unternehmen berichten von Schwierigkeiten, von einer zentral kontrollierten und durchgesetzten Daten-Governance zu einer föderierten Governance-Struktur zu wechseln.
- **Sicherheit, Datenschutz & Regulation:**
  Besonders bei Themen wie Sicherheit, Datenschutz und regulatorischen Vorgaben ist die Herausforderung groß. Wenn diese Aspekte zuvor zentralisiert verwaltet wurden, kann ihre Dezentralisierung zu Unklarheiten und potenziellen Risiken führen.
- **Zuständigkeiten:**
  Die föderierte Struktur kann zu Verwirrung darüber führen, wer für welche Daten oder Prozesse verantwortlich ist. Dies kann zu Verzögerungen und Fehlkommunikationen führen.
- **Regel-Implementierung:**
  Während Regeln zentral gespeichert werden sollten, müssen sie dezentral gepflegt und angewendet werden. Ansätze wie ein Sidecar können helfen, Regeln direkt in Datenprodukte zu integrieren.
- **Zentrales Regelrepository:**
  Ein zentrales Regelrepository ermöglicht es, dass trotz dezentraler Pflege von Datenprodukten konsistente und einheitliche Regeln angewendet werden.

HERAUSFORDERUNGEN & BEST PRACTICES

BEST PRACTICES

- **Interne Schulungen:**
  Bieten Sie Schulungen für bestehende (Dev-)Teams an, um das Wissen über Data Mesh zu vertiefen und seine Vorteile zu demonstrieren.
- **Brücken bauen:**
  Arbeiten Sie eng mit Entwicklern zusammen, um herauszufinden, welche Anreize und Tools sie benötigen, um sich für Data Mesh zu begeistern.
- **Modulare Implementierung:**
  Beginnen Sie mit einer modularen Implementierung von Data Mesh, um Komplexität zu reduzieren und schnelle Erfolge zu erzielen.
- **Klare Kommunikation:**
  Stellen Sie sicher, dass die Vorteile und Ziele von Data Mesh klar und regelmäßig an alle Stakeholder kommuniziert werden, um ein gemeinsames Verständnis und Engagement zu fördern.

LEARNINGS

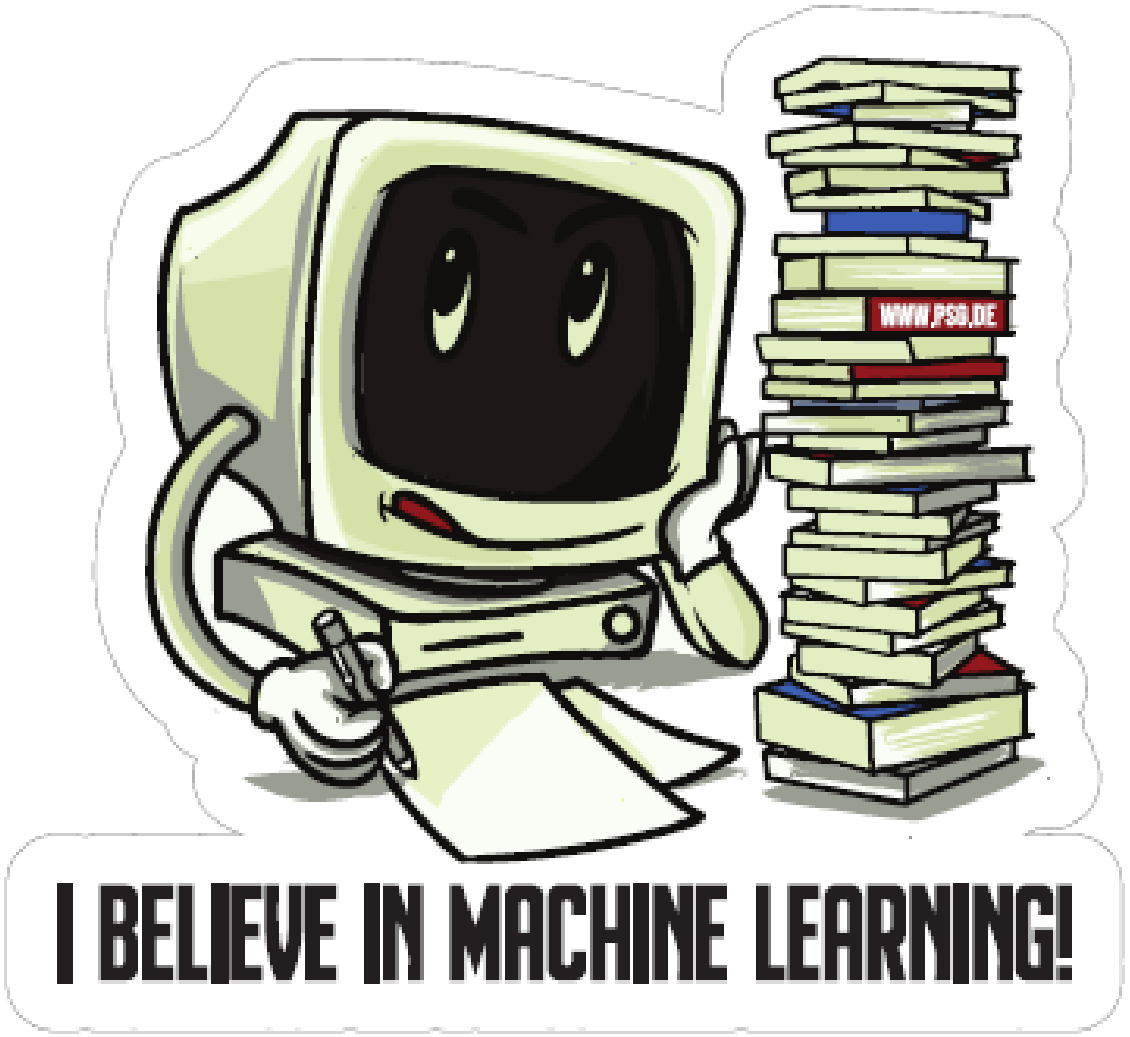
VERSTEHEN DER SPEZIFISCHEN ANFORDERUNGEN

- **Branche und Geschäftsmodell:**
  Verschiedene Branchen haben unterschiedliche Datenbedürfnisse und -prioritäten. Ein Finanzunternehmen wird andere Anforderungen an Datensicherheit und -regulierung haben als ein Einzelhandelsunternehmen.
- **Datenreife und -qualität:**
  Bevor Sie eine Data Mesh-Architektur einführen, müssen Sie verstehen, wie fortgeschritten Ihr Datenökosystem ist und welche Datenqualitätsherausforderungen bestehen. (It‘s a trap!)
- **Technologische Infrastruktur:**
  Beurteilung der aktuellen technologischen Fähigkeiten und Integration mit neuen Tools und Plattformen, die für eine Data Mesh erforderlich sind. (Quasi jeder…)
- **Kulturelle und organisatorische Faktoren:**
  Welche kulturellen Veränderungen sind erforderlich, um den Übergang zu einer Data Mesh-Architektur zu erleichtern? Wie kann Ihr Unternehmen diese Änderungen unterstützen und vorantreiben?

LEARNINGS

PRAKTISCHE SCHRITTE ZUR IMPLEMENTIERUNG

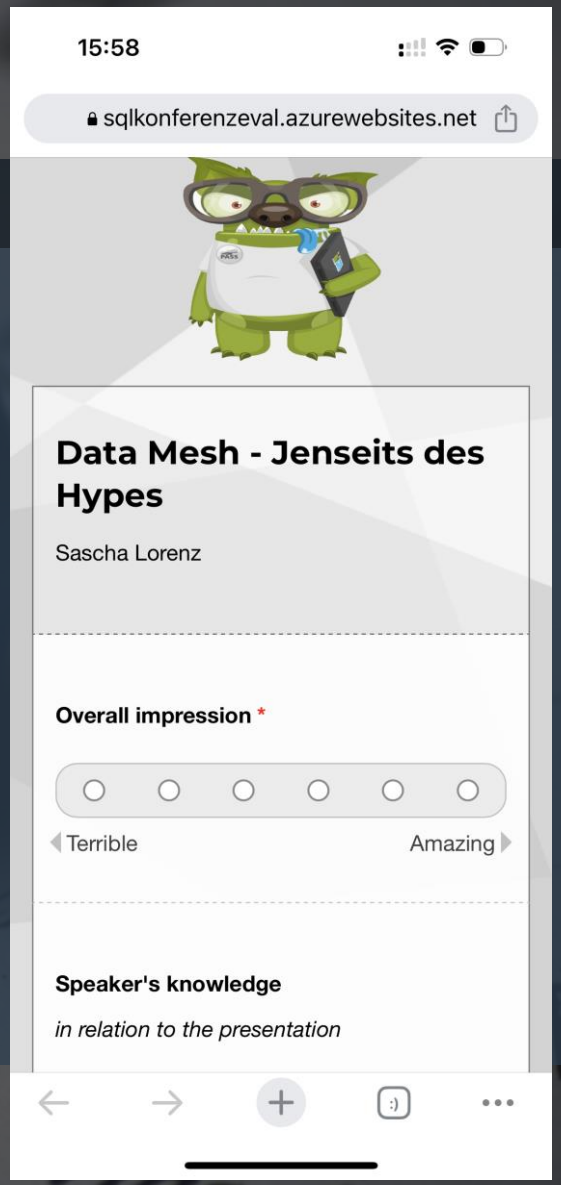
- **Domain-Driven Design:**
  Dieser Ansatz hilft, den Datenbedarf anhand spezifischer Geschäftsanforderungen zu identifizieren. Durch die Aufteilung in verschiedene "Domains" können Datenquellen effektiv und zweckmäßig organisiert werden.
- **Federated Data Governance:**
  Erstellen Sie ein System, bei dem verschiedene Geschäftsbereiche oder "Domains" ihre eigenen Daten verwalten, aber gleichzeitig den Unternehmensstandards folgen.
- **Data Product Management:**
  Betrachten Sie Daten bzw. Informationen als Produkte. Dies erfordert eine klare Strategie, Verantwortung und Metriken, um den Wert der Daten zu messen und zu maximieren.
- **Technische Herausforderungen:**
  Identifizieren und planen Sie für technische Hürden wie Datenintegration, Sicherheit, Datenqualität und Skalierbarkeit.



- If you are interested in this one, and there is no more available here?
- Just ask me!

FEEDBACK




15:58
sqlkonferenzeval.azurewebsites.net
Data Mesh - Jenseits des Hypes
Sascha Lorenz
Overall impression *
Terrible
Amazing
Speaker's knowledge
in relation to the presentation

PSG Projekt Service GmbH

The SQL Server Company



SASCHA LORENZ

Geschäftsführer

Neuer Wall 80
20354 Hamburg

+49 40 39 88 28-62

lorenz@psg.de

www.psg.de