

Data-as-a-Service with Snowflake and KNIME

Marcel Meyer – SIEMENS Healthineers

Ripu Jain – Snowflake

Stephen Rauner - KNIME

November 15, 2022



Speakers



Marcel Meyer

Head of Data Governance and Management



Ripu Jain

Senior Partner Engineer

Superpower Snowflake Data

- KNIME is now a Snowflake Premier Partner
- Access and execute your cloud data
- Offering a full data analytics stack
- Americas and Europe as targeted Geos



+



Safe Harbor and Disclaimers

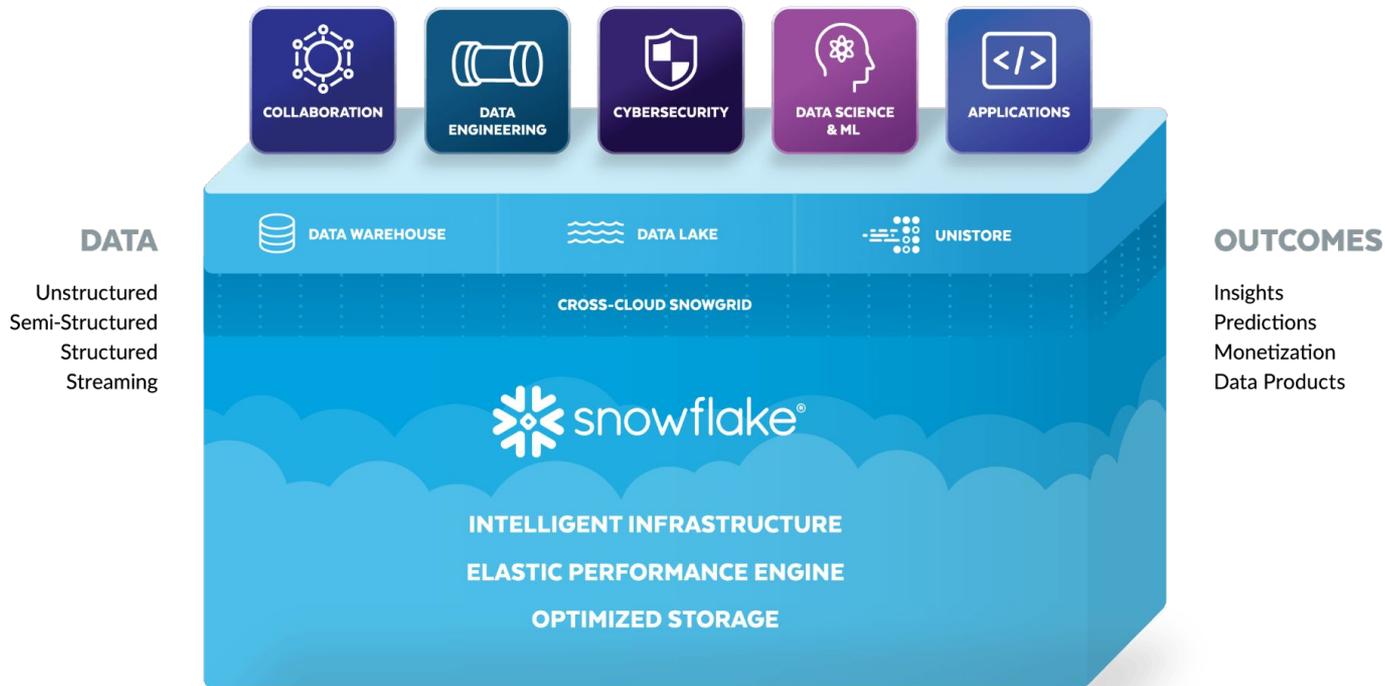
Other than statements of historical fact, all information contained in these materials and any accompanying oral commentary (collectively, the “Materials”), including statements regarding (i) Snowflake’s business strategy and plans, (ii) Snowflake’s new or enhanced products, services, and technology offerings, including those that are under development or not generally available, (iii) market growth, trends, and competitive considerations, and (iv) the integration, interoperability, and availability of Snowflake’s products, services, or technology offerings with or on third-party platforms or products, are forward-looking statements. These forward-looking statements are subject to a number of risks, uncertainties and assumptions, including those described under the heading “Risk Factors” and elsewhere in the Annual Reports on Form 10-K and the Quarterly Reports on Form 10-Q that Snowflake files with the Securities and Exchange Commission. In light of these risks, uncertainties, and assumptions, the future events and trends discussed in the Materials may not occur, and actual results could differ materially and adversely from those anticipated or implied in the forward-looking statements. As a result, you should not rely on any forwarding-looking statements as predictions of future events.

Any future product or roadmap information (collectively, the “Roadmap”) is intended to outline general product direction; is not a commitment, promise, or legal obligation for Snowflake to deliver any future products, features, or functionality; and is not intended to be, and shall not be deemed to be, incorporated into any contract. The actual timing of any product, feature, or functionality that is ultimately made available may be different from what is presented in the Roadmap. The Roadmap information should not be used when making a purchasing decision. In case of conflict between the information contained in the Materials and official Snowflake documentation, official Snowflake documentation should take precedence over these Materials. Further, note that Snowflake has made no determination as to whether separate fees will be charged for any future products, features, and/or functionality which may ultimately be made available. Snowflake may, in its own discretion, choose to charge separate fees for the delivery of any future products, features, and/or functionality which are ultimately made available.

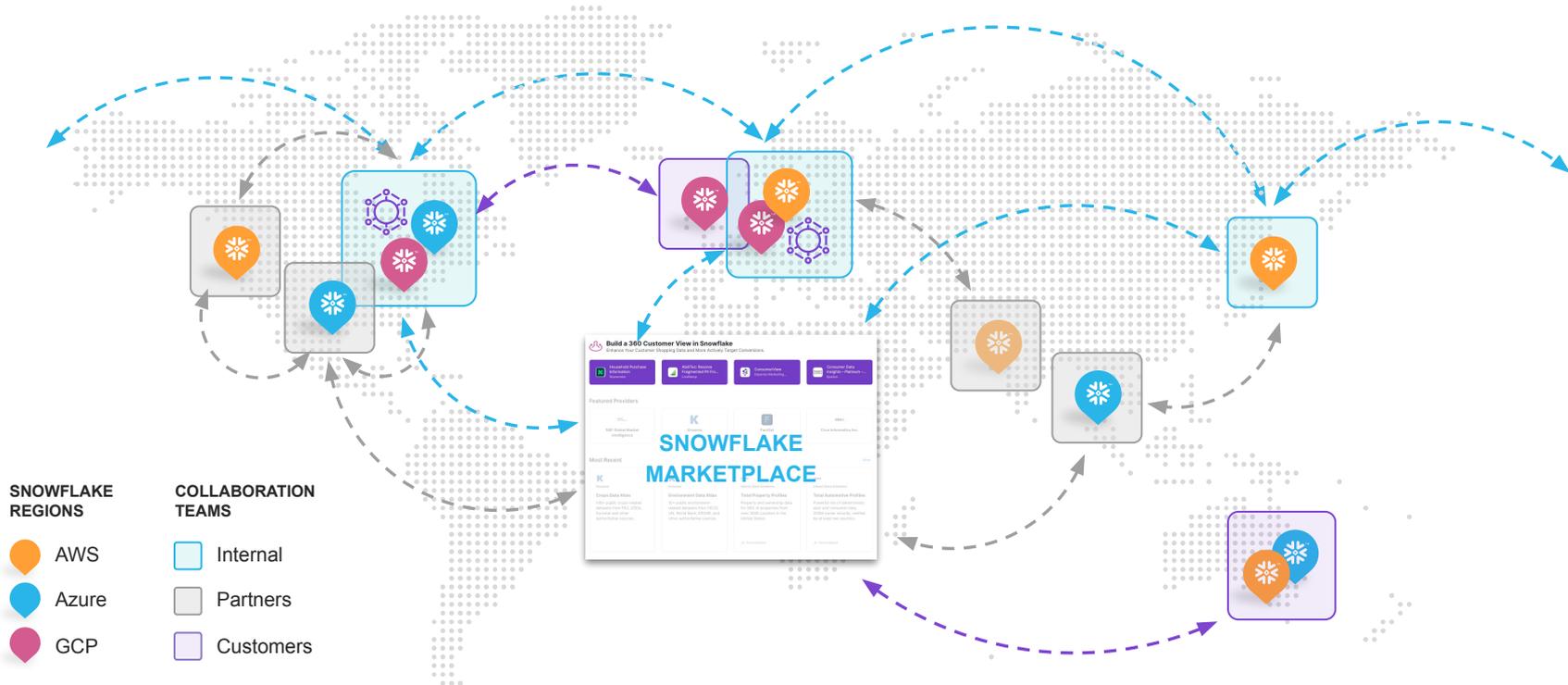
© 2022 Snowflake Inc. All rights reserved. Snowflake, the Snowflake logo, and all other Snowflake product, feature and service names mentioned in the Materials are registered trademarks or trademarks of Snowflake Inc. in the United States and other countries. All other brand names or logos mentioned or used in the Materials are for identification purposes only and may be the trademarks of their respective holder(s). Snowflake may not be associated with, or be sponsored or endorsed by, any such holder(s).



One Platform. Many Workloads.



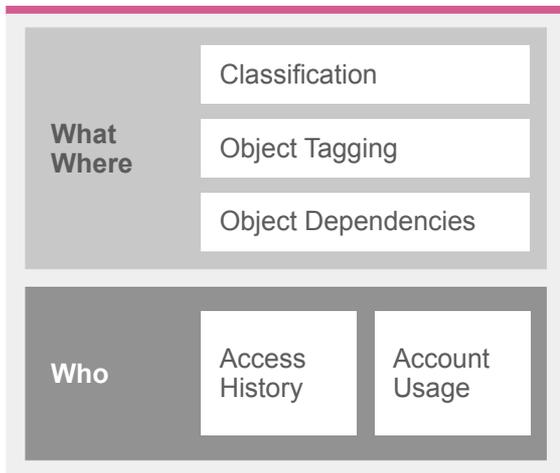
Collaborating in the Data Cloud



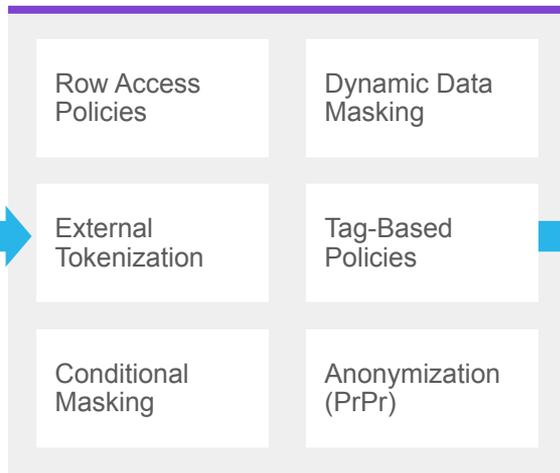
Unified Governance



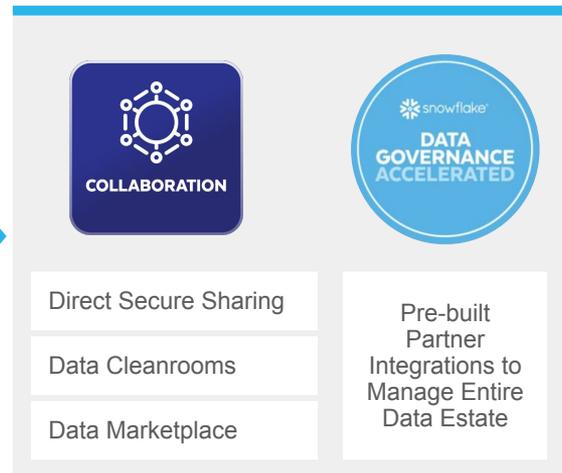
Know Your Data



Protect Your Data



Connect Your Ecosystem



Snowpark for Python



Familiar Programming Constructs

Use familiar syntax with DataFrame abstraction



Rich Ecosystem

Easy access to hundreds of packages with automated dependency management



Secure Processing

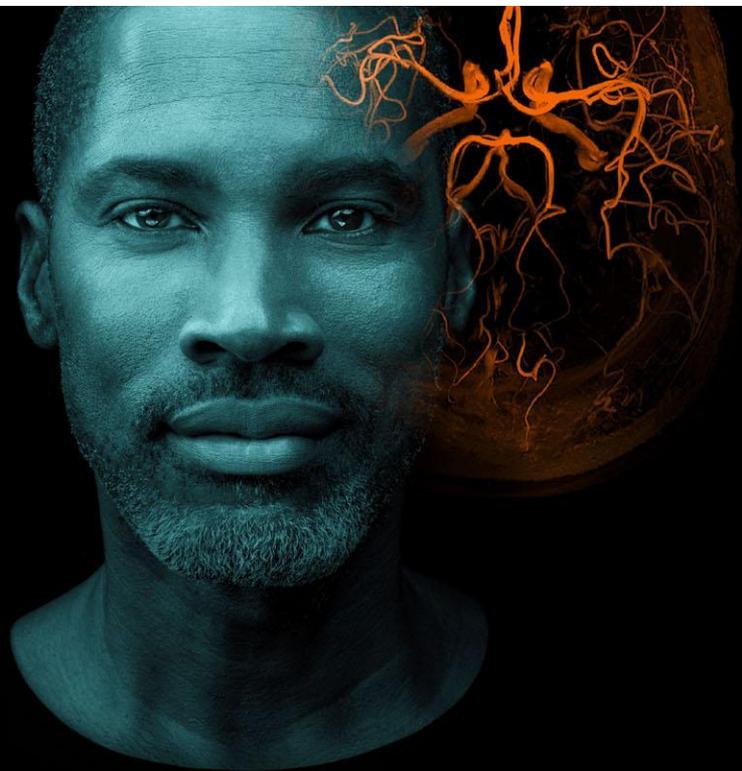
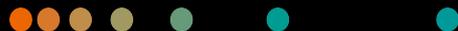
Build with confidence in a highly secure, sandboxed environment



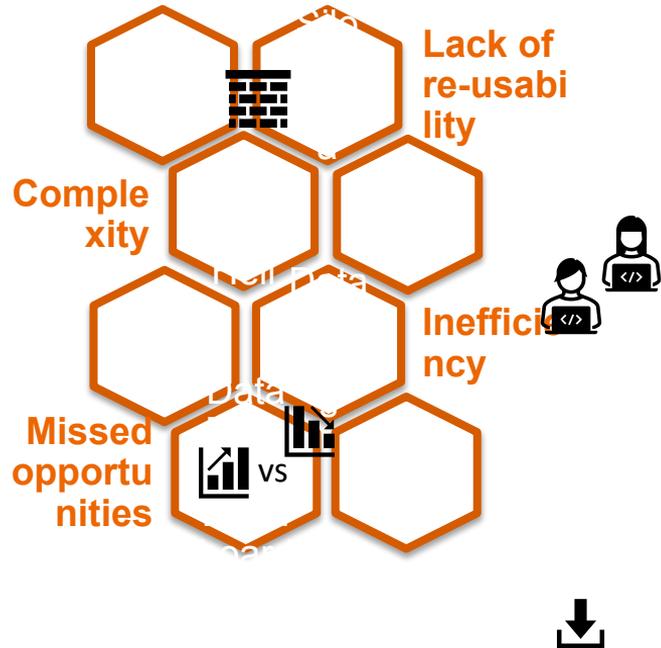
Data-as-a-service with Snowflake and KNIME

Siemens Healthineers, Marcel Meyer
CRM Excellence –
Head of Data Governance & Management

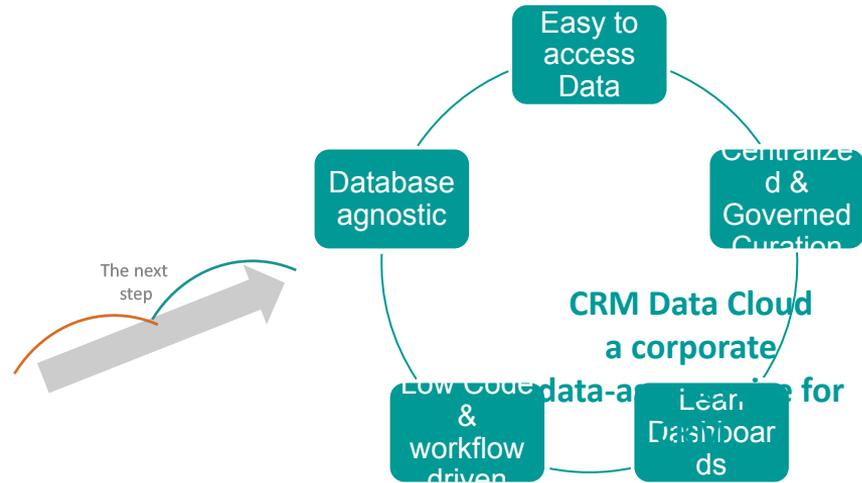
KNIME SUMMIT Fall 2022, 11/15/2022



In 2020 we decided our current approach for CRM analytics needs a change!



Identified problems



Improvements aimed

With the CRM Data Cloud we pioneered data management within Siemens Healthineers

30+ complex analytics
use cases empowered



User size doubled
in last 12 month

30k+ data
downloads as
day



70%+ users access
data multiple times
per week

Flexible & easy
connectivity
(Excel or dashboarding/analytics)



700+ colleagues
onboarded

The CRM Data Cloud supports a broad range of use cases



CRM Data Data Cloud is THE globally accepted governed and curated data source for CRM data at Siemens Healthineers

Three examples out of 30+ use cases owned by Regions & Business Areas

Latin America

Commercial Excellence

The use case

The commercial excellence dashboard in Latin America is the one stop shop for all sales leaders helping them to understand business trends and performance

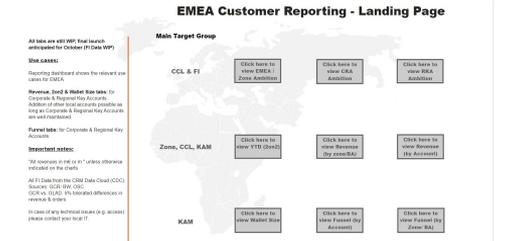


Europe

EMEA "Cure" Dashboard

The use case

The "Cure" dashboard was developed by the EMEA Regional Office to support various management level conducting business reviews on customer level

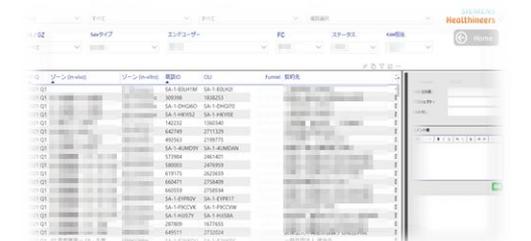


Asia

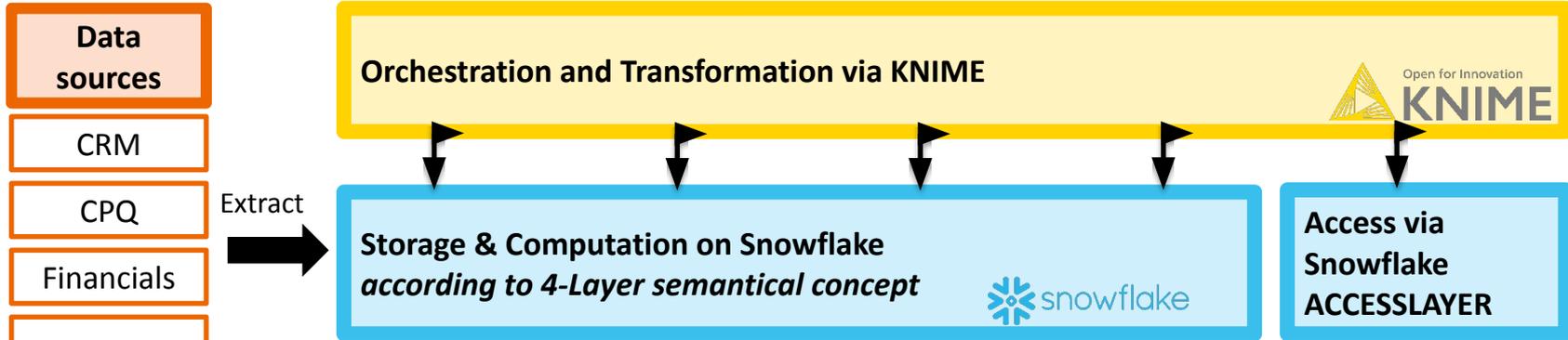
Japan – Key Account Funnel Dashboard

The use case

Funnel transparency for Key account management connecting multiple data marts in one dashboard. Fully automated data synchronization with CDC.



The overarching architecture of the CRM Data Cloud



Why these Tools?



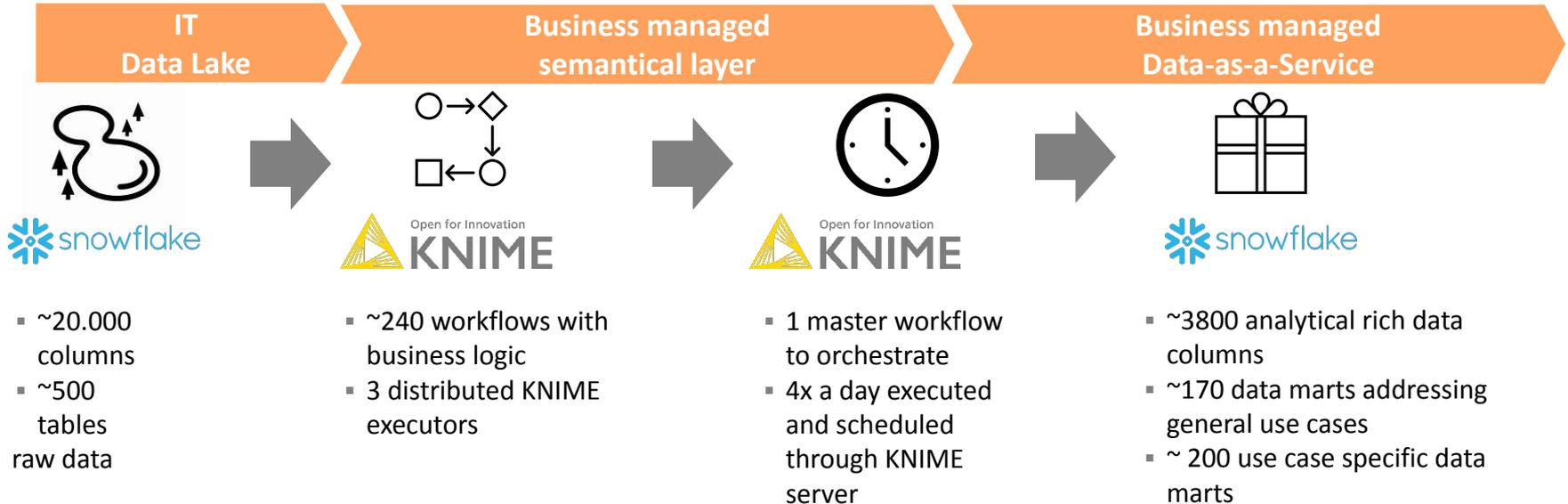
Low-Code creation of transformation Workflows



DWH and computation performance for huge datasets. In line with global IT Strategy

Row Level Security and integrated with Microsoft environment

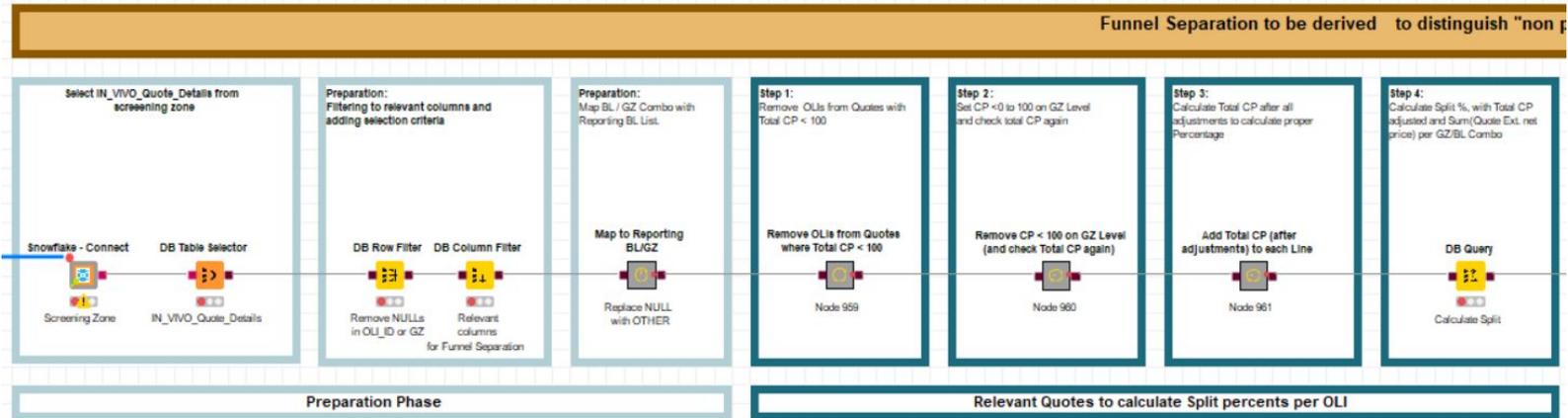
A highly automated data management delivers consistent data on large scale



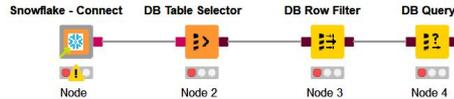
Examples of Workflow:

Transform data in a visual and low code way

- Nodes with dedicated behavior (e.g.: Table Selector performs a select on the Database)
- Annotations are possible to explain transformation steps in detail
- Metanodes (groups of Nodes) to cluster logics that require multiple nodes
- Visually representation of data flow
- Reusable logics bundled in Components and shared across workflows

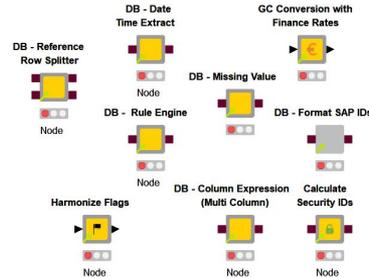


Low Code



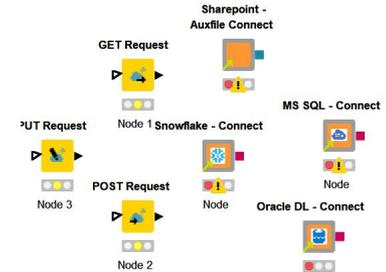
- It's easy to handover workflows within the team
- Visualization helps to understand the general idea of the workflow
- Re-wiring is easily possible
- Intermediate results can be examined during bug-fixing/error-analysis
- You can still code when needed

Re-usable components



- Enriching KNIME native capabilities with Snowflake functionality
- Encapsulating & govern business logics
- Reduce time-to-solution by re-using a proven solution to a data problem

Interoperability



- We managed to migrate our data processing from Oracle to Snowflake within 4 weeks
- External data outside our data lake (e.g. mappings, etc.) can easily be brought in
- Interaction with other APIs possible during data load (e.g. data quality cleansing)

We have built a rich repository of components enabling Snowflake In-DB calculations, business logic, and automation



- ~40 components for in DB data manipulation
- Components to check certain columns out of Snowflake, manipulate with KNIME and join back to snowflake
- Adjust DDL of tables/view incl. metadata while writing data
- ...

- Date to FY conversion
- Currency translations
- Organizational mappings
- ...

- Master Workflow orchestration incl. logging
- Automated Row Level Access Policy & Data Masking on Snowflake the
- Workflow meta data extraction to Data Catalog
- ...

Some valuable lessons learned to share with you



Use meta data & annotations to describe/document



Organize components in a similar/same structure than KNIME nodes



Build simple example workflow (train, test, improve)



Keep components updated & efficient by using new KNIME features, plan for maintenance!



Shared components are linked with absolute links!
Think twice about location and name before using!



Generalize scope rather being too narrow!



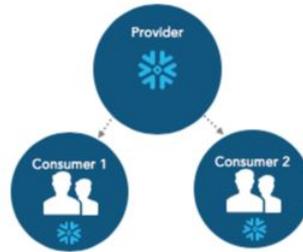
Use Configuration Layout feature to optimize usability

Warehouse scaling



- You can switch warehouses without re-connect -> You can for instance switch to a better warehouse only if large data is processed (we added this to our components!)
- You may want to use different warehouses during development and operational workflows (we added this automatic switch into our components)

Data sharing



- Share data amongst different snowflake instances -> We managed data integration within 24h in a recent acquisition
- Zero-Copy cloning is a very powerful function that can be used throughout development/testing

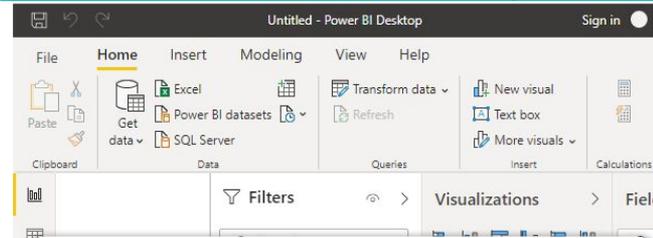
Scalable security



- Leverage row-access policies (can be automated easily with KNIME)
- Tag based data masking allows to master DP challenges (meta data can be forced through KNIME components if needed)

Honestly,...our users don't care about how we manage it...as long as they have easy access to the data with their tool of choice

MS Power BI standard connectivity



A model refreshed with a push of a button

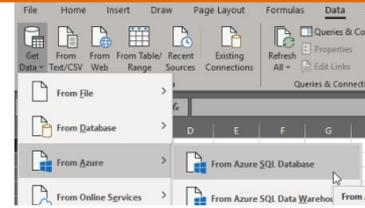
Apply query changes

datacloud.012b25348e9e.database.windows.net/Topic_Areas.

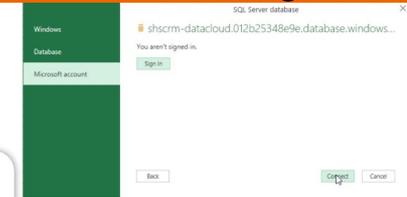
- 999_RLS_GAMA_Users - Countries
239,426 rows from shscrm-datacloud.012b25348e9e.database.windows.net/Topic_Areas.
- 999_Funnel
1,177,734 rows from shscrm-datacloud.012b25348e9e.database.windows.net/Topic_Areas.
- 999_OLIs
355,493 rows from shscrm-datacloud.012b25348e9e.database.windows.net/Topic_Areas.

Cancel

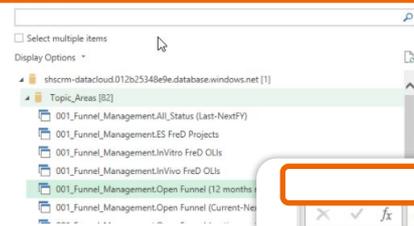
Excel standard function



Office 365 Login



Select data mart



Get your data into Excel

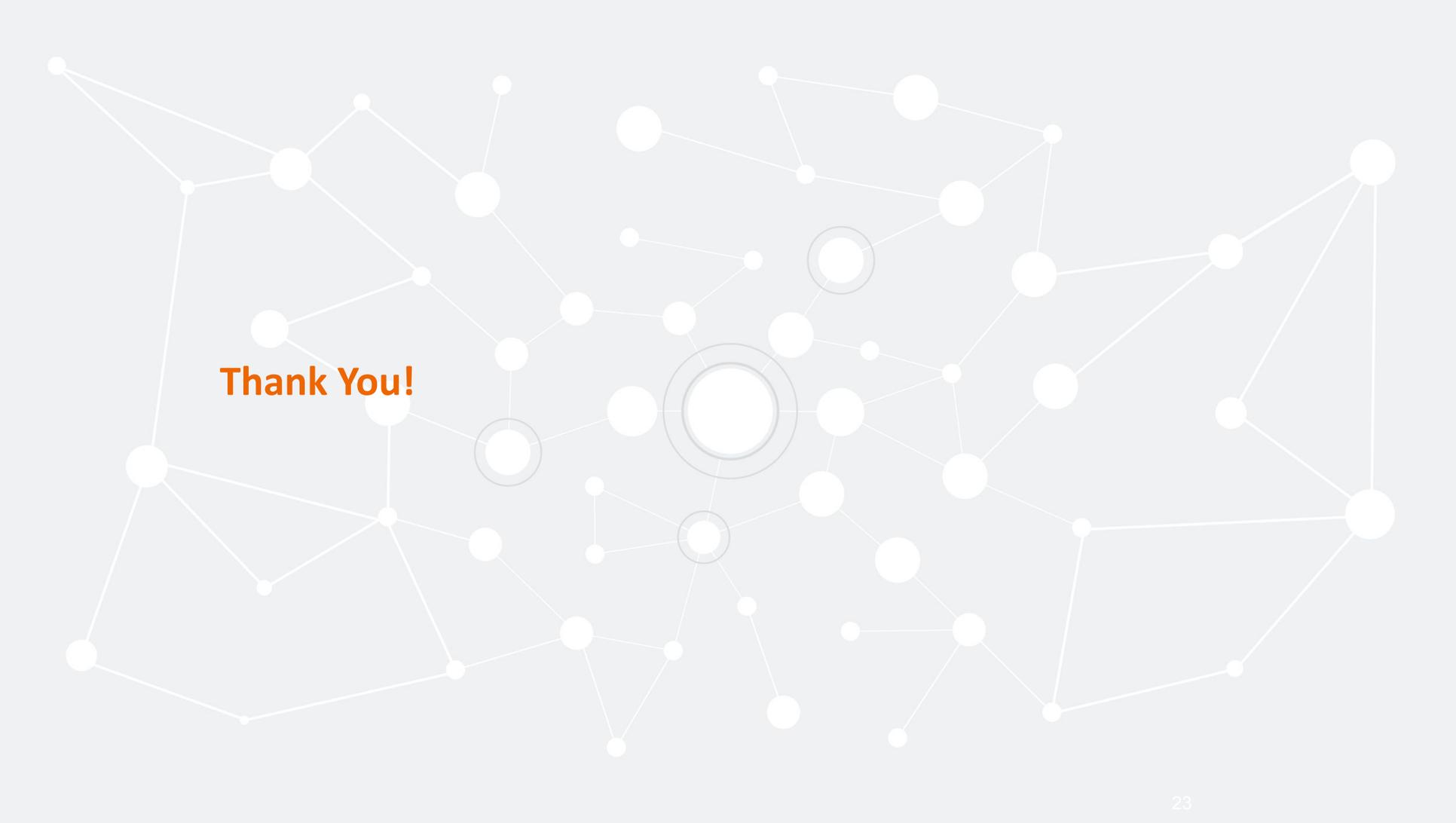
= Topic_Areas[Schema="001_Funnel_Management",Item="Open Funnel (12 months rolling)"][Data]

	Account City	Account Class In Vitro (DX)	Account Class In Vivo (EQ)	Account Country
1	+I4 São Cosmes-Gondomar	NULL	B3	PT
2	null Vaduz	C1	TA	GR
3	null Vaduz	C1	TA	GR
4	+I4 San Cristóbal de La Laguna	A1	A1	ES
5	null LAS PALMAS DE GRAN CANARIAS	B1	C1	ES
6	null Las Palmas De Gran Canaria	B1	A2	ES
7	null Las Palmas De Gran Canaria	B1	A2	ES
8	+I4 Las Palmas De Gran Canaria	B1	A2	ES
9	null Las Palmas De Gran Canaria	B1	A2	ES

Summary & Thank you!

- KNIME is more than just a tool for data science or to stitch some Excel or CSV files together (it can do so much more)
- KNIME server and distributed executors open the universe for an enterprise usage
- KNIME and Snowflake work great together
- Make use of the strengths of both tools
 - KNIME: Low-Code, Re-usable components, meta data for workflows, endless options to solve a data problem
 - Snowflake: The power of a database in the cloud, strong security, scaling (to keep credit usage under control)
- Have a game plan by having a good data architecture!
- Stay creative, innovative and solution oriented!





Thank You!