



Cognitive Enterprise Data Platform (CEDP) Primer

January, 2016

Vision Statement

CEDP is a **world-class enterprise data infrastructure** that integrates strategic cognitive IBM offerings on a **cloud platform** with our enterprise environment to consolidate and make available **IBM's rich data resources** in an unprecedented way. This platform and its services will transform IBM into a **cognitive business enterprise** through **cognitive services** and providing means to **share** and **integrate data** in fundamentally new ways. The key offerings are Cognitive Enterprise Data Lake (**CoEDL**), cognition services (**CaaS**), data services (**DaaS**), and trust services (**TaaS**). Through these interfaces, **IBM Business** and **Functional Units** will interact with the platform as **producers** and **consumers**, i.e., both contributing to and using its capabilities.



CEDP: Offerings

Cognitive Enterprise Data Lake (CoEDL)

- **Repository** of Data (all types) **organized** to support applications and personas
- **Policy enforcement** on owning (define policies), writing (generate and overwrite) and viewing data
- **360-degree** view of data
- Single **enterprise-wide catalog**

Cognition-as-a-Service (CaaS)

- **Cognitive analytics** for **structured data** (SPSS, RStudio, Notebooks)
- **Annotators** for **unstructured data** (e.g. BigInsights Web Interface for SystemT, Watson Knowledge Studio for SIRE)
- **Deploy and execute cognitive analytics** and annotators
- **Natural language interfaces**
- Automate retraining (**Self-Learning**) of cognitive analytics

Data-as-a-Service (DaaS)

- High scale **data ingestion**
- **Accessing data** for analysis and manipulation
- **Self-service tools, Sandboxes, social and collaborative tool**
- Automated **ingestion of streaming and batch data**
- **Out-of-the-box tools** for querying, exploring, dashboarding, and executing pipelines over datasets

Trust-as-a-Service (TaaS)

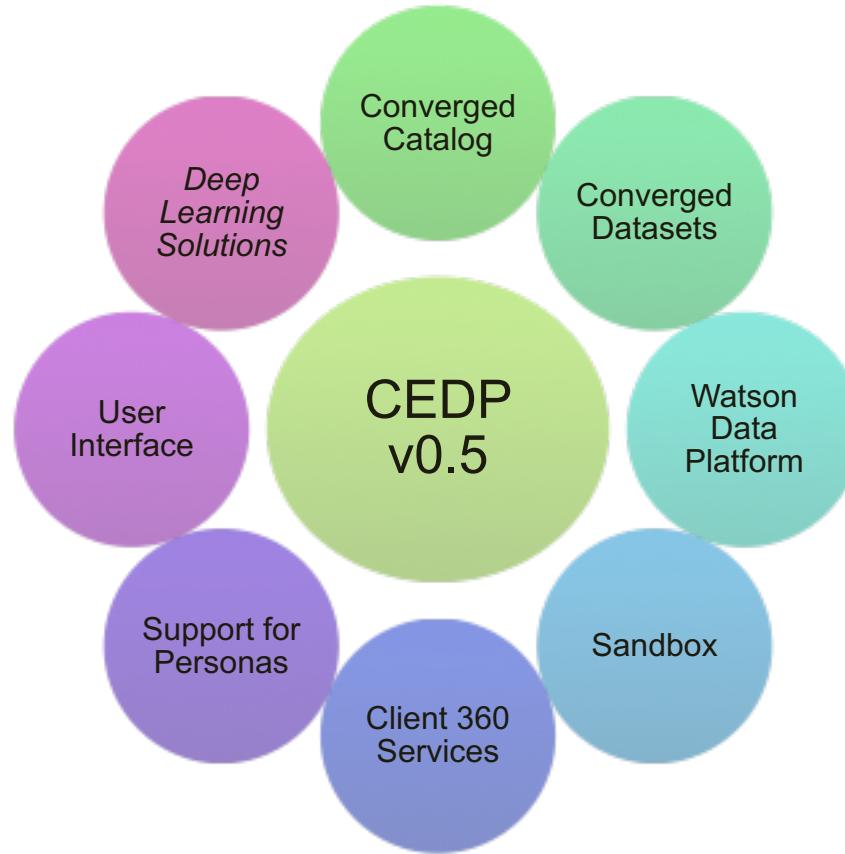
- **Access**
- **Governance**
- **Data lineage**
- **Social Signals**



CEDP v0.5 in Q42016

Producers

- ✓ Ingest data from Data Sources
- ✓ Ingest data for Reporting applications
- ✓ Develop and Deploy Analytics jobs
- ✓ Provide Technical Metadata
- ✓ Submit requests through Forms
- ✓ Provide access policies and authorize
- *Build reporting applications*
- *Deploy analytical models*
- *Ensure data quality*
- *Provide business questions for CaaS*
- *Real-time IOT data ingestion*



Consumers

- ✓ Enterprise-wide Catalog to browse data assets
- ✓ Obtain Client 360-degree views
- ✓ Perform Analytics with data
- ✓ Support Cognitive Process Initiatives
- ✓ Support Business Applications through APIs
- ✓ Submit requests through Forms
- Develop self-service applications
- Enterprise-wide governance
- Obtains answers to Business questions
- Share Trusted data within the enterprise
- Alerts from real-time IOT data ingestion

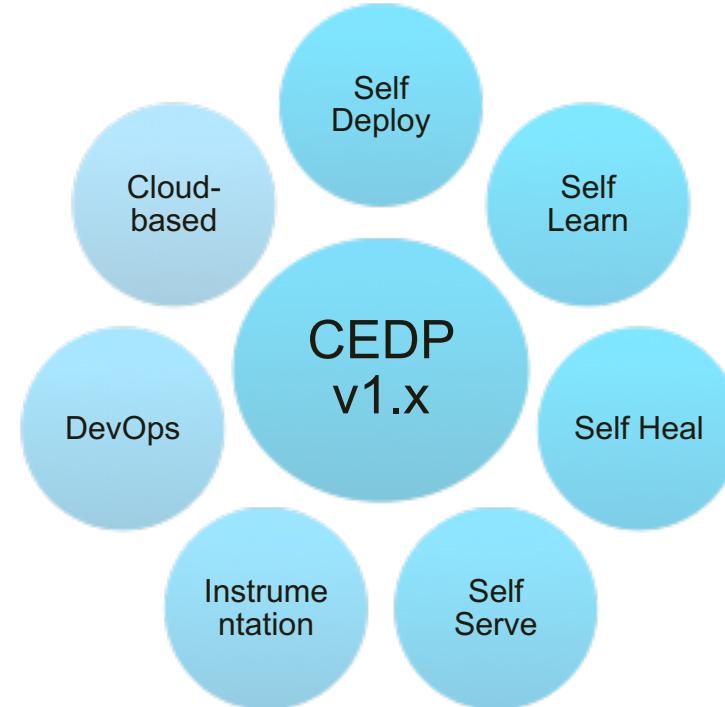
Governance (TaaS v0.5)



CEDP v1.0 in Q42017

Producers

- Ingests data for analytics
- Ingests infrequently used data
- Ingests data for sunsetting DW
- Builds reporting applications
- Deploys analytical models
- Provides technical metadata
- Ensures data quality
- Provides business questions for CaaS
- Real-time IOT data ingestion
- ...



Consumers

- Obtain 360-degree views
- Perform analytics with data
- Obtain reports
- Enterprise-wide catalog
- Enterprise-wide governance
- Obtains answers to Business questions
- Alerts from real-time IOT data ingestion
- Enables Cognitive Process Initiatives
- Share trusted data within the enterprise
- ...

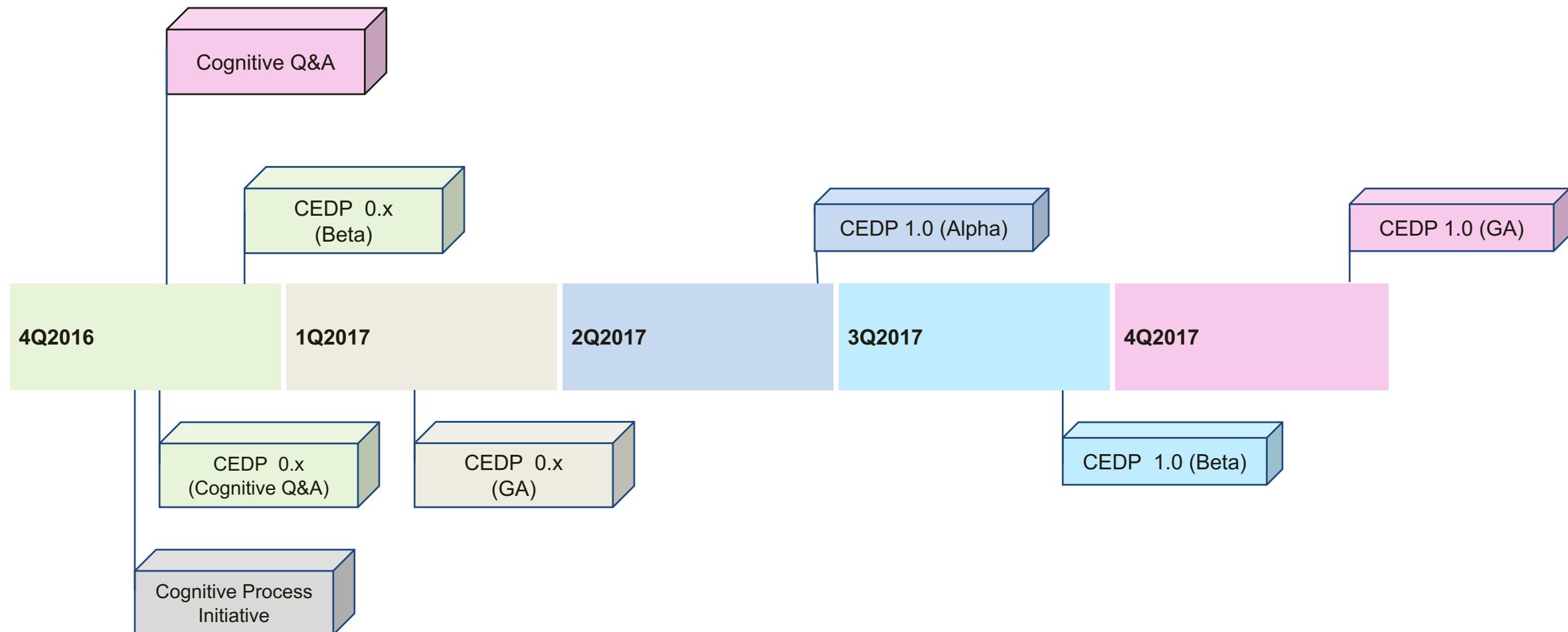


CEDP: Personas

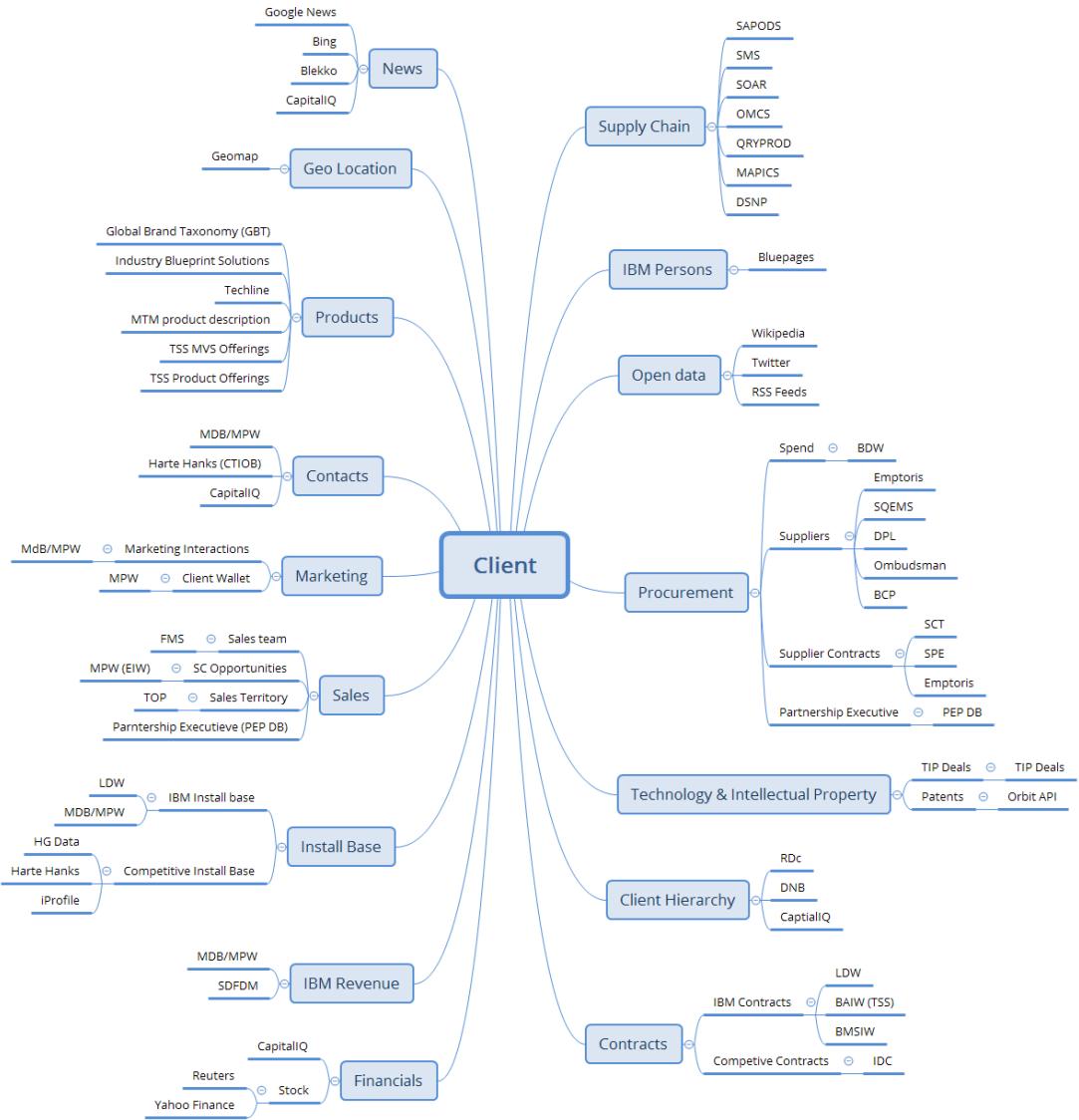
Data Engineer	Data Scientist	App Engineer	Business Analyst	Subject Matter Expert
<ul style="list-style-type: none">• Ingests data for analytics and reporting applications• Ensures Data Quality	<ul style="list-style-type: none">• Deploys analytical models• Support Cognitive Process Initiatives	<ul style="list-style-type: none">• Support Business Applications through APIs• Develop self-service applications	<ul style="list-style-type: none">• Perform Analytics with data• Obtain Client 360-degree views	<ul style="list-style-type: none">• Provide Technical Metadata• Ensure data quality• Obtains answers to Business questions



CEDP: Timeline



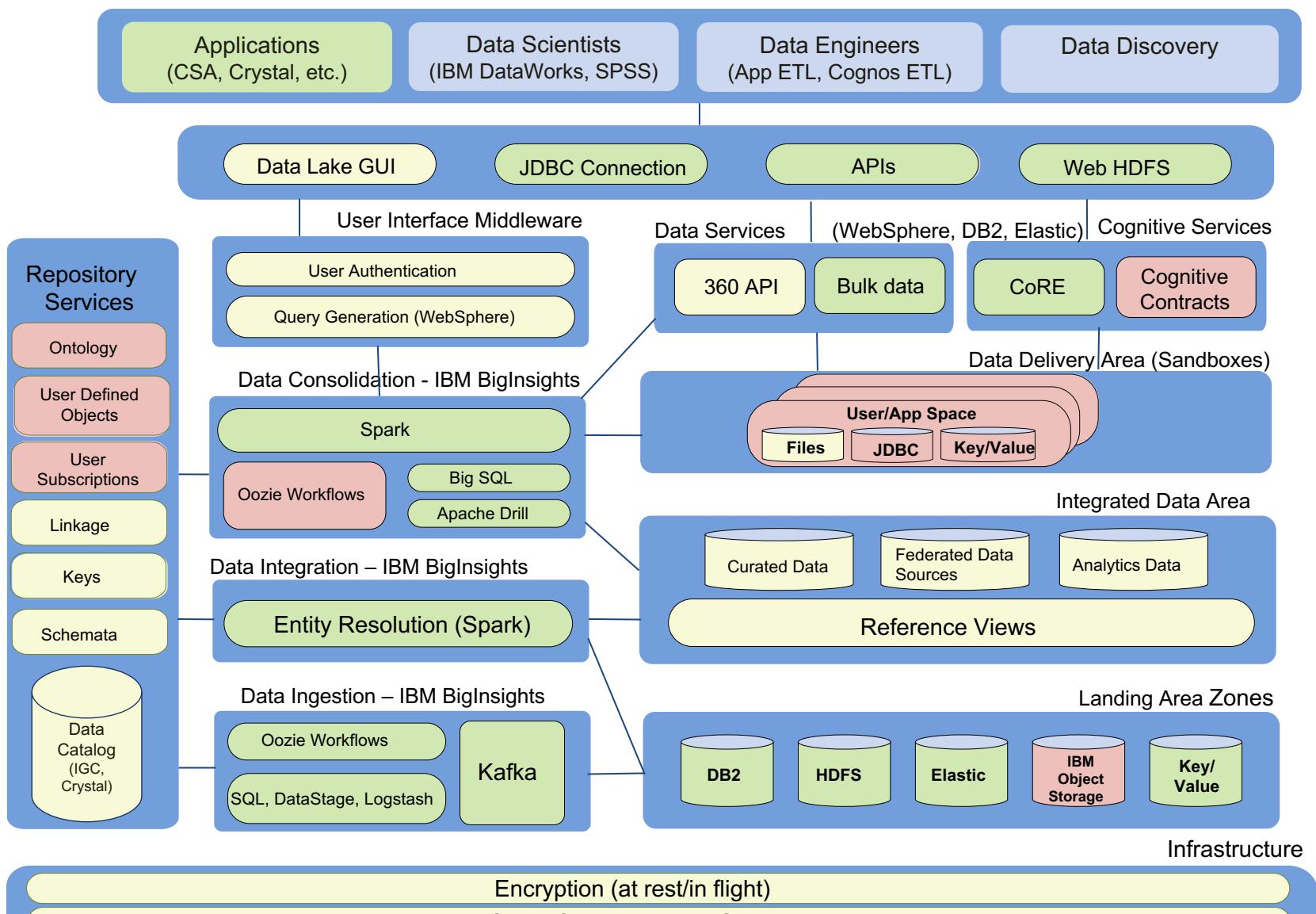
Data in CoEDL



Please note: Currently, not all datasets from the listed Data sources are ingested into the Lake.

CEDP v0.5 Architecture

DRAFT

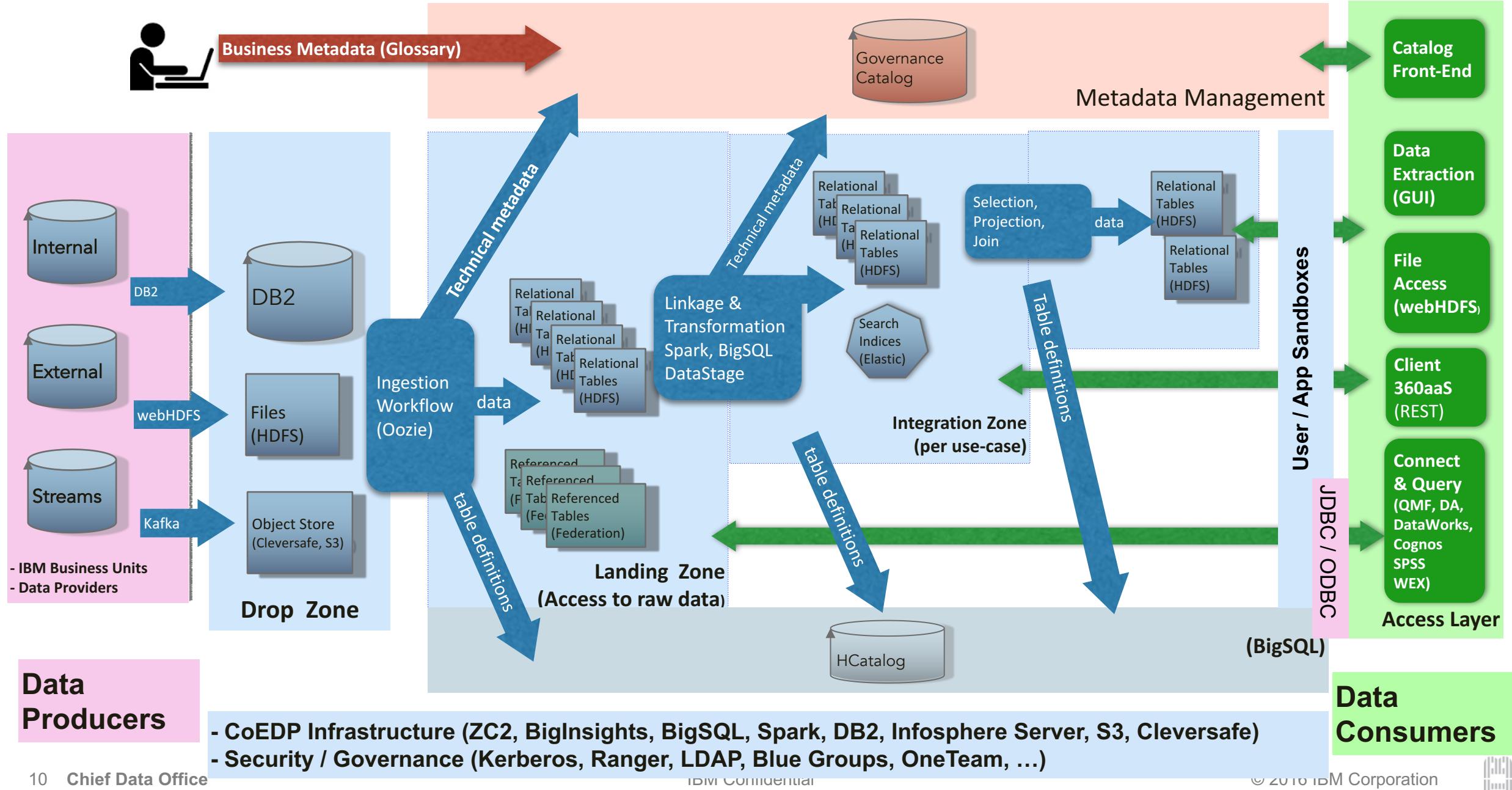


Authorization, Audit, Lineage
Data Governance (OneTeam, IBM security processes, etc.)

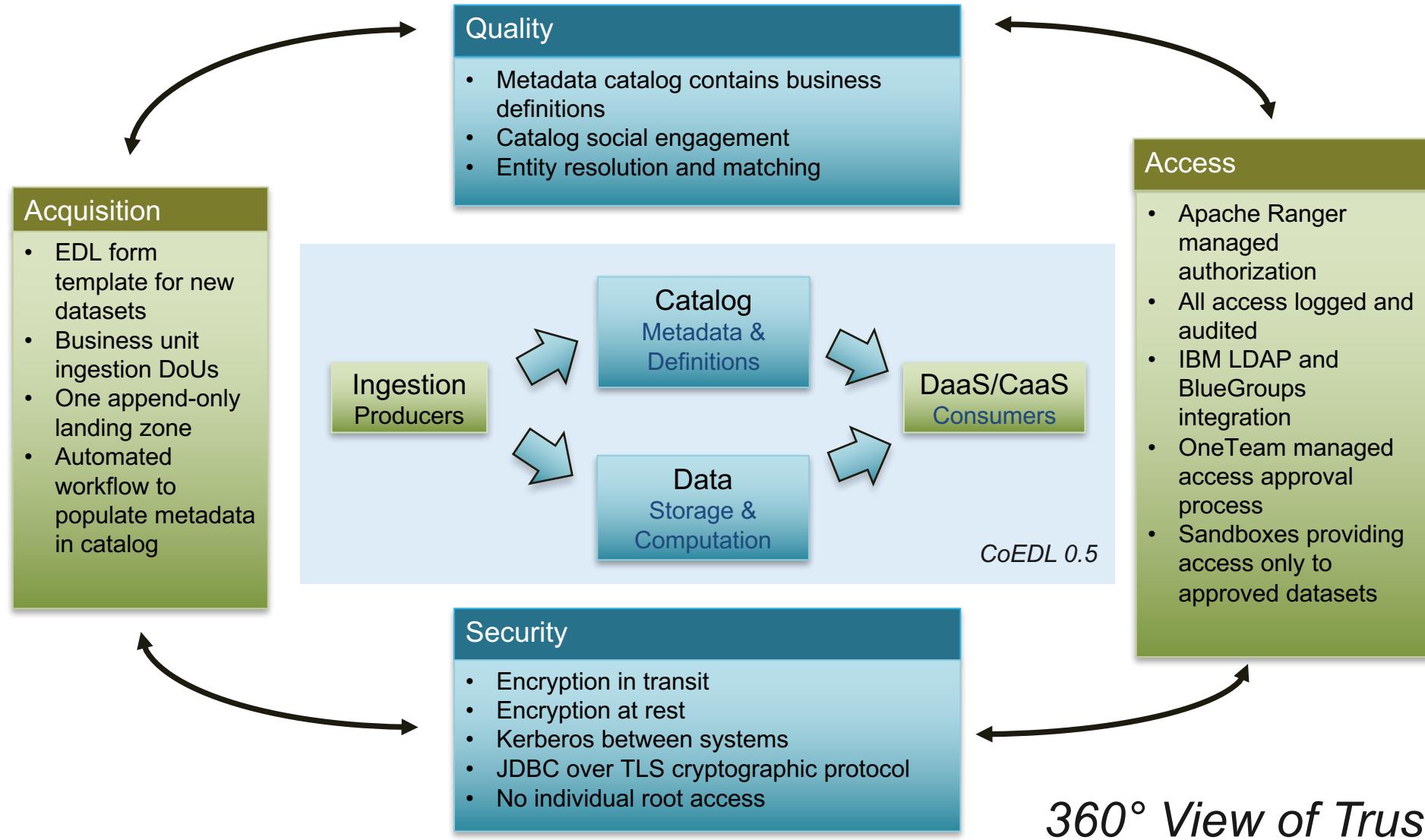


Data Flow & Services in CEDP v0.5

DRAFT



Trust as a Service (TaaS) in CEDP v0.5



360° View of Trust



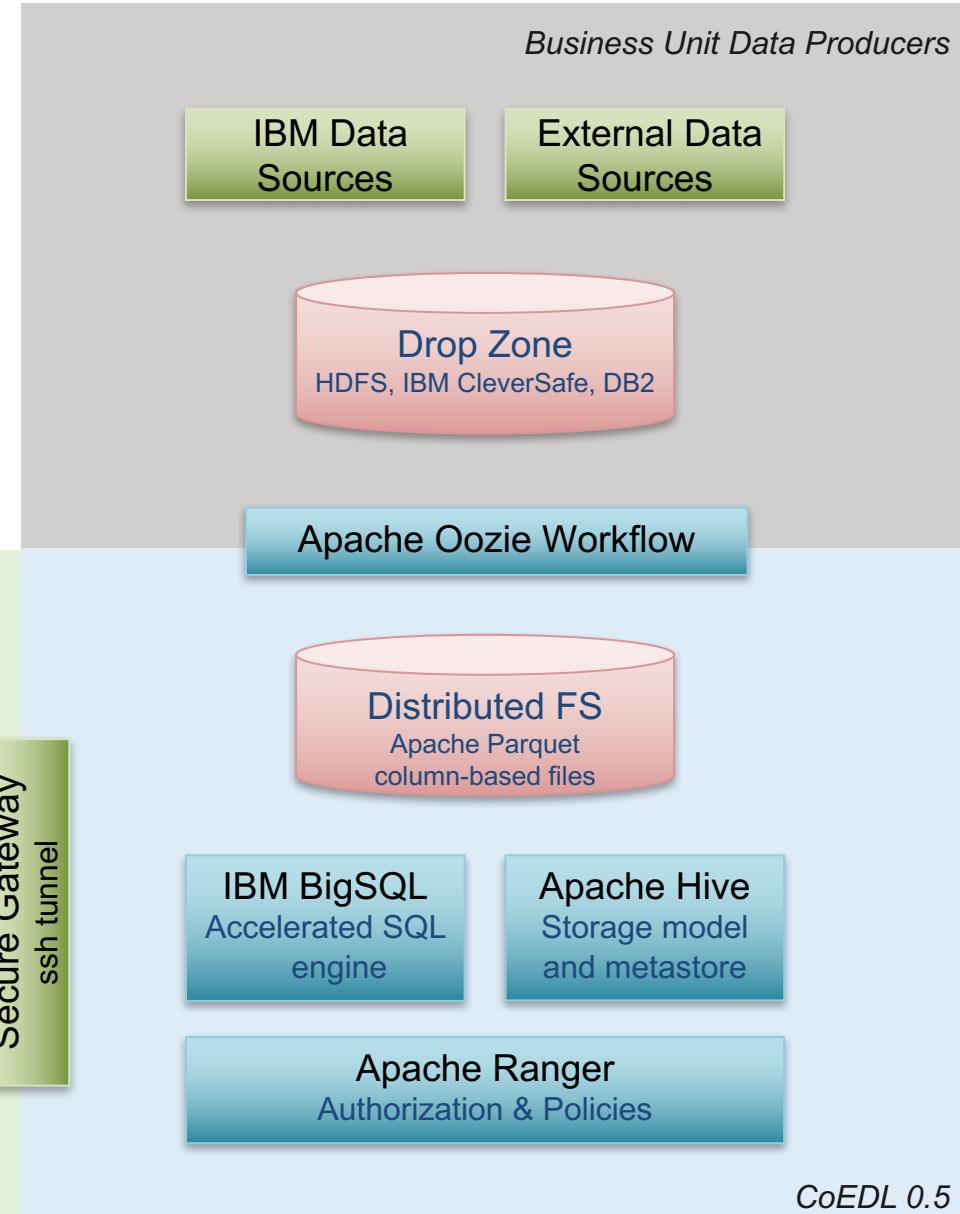
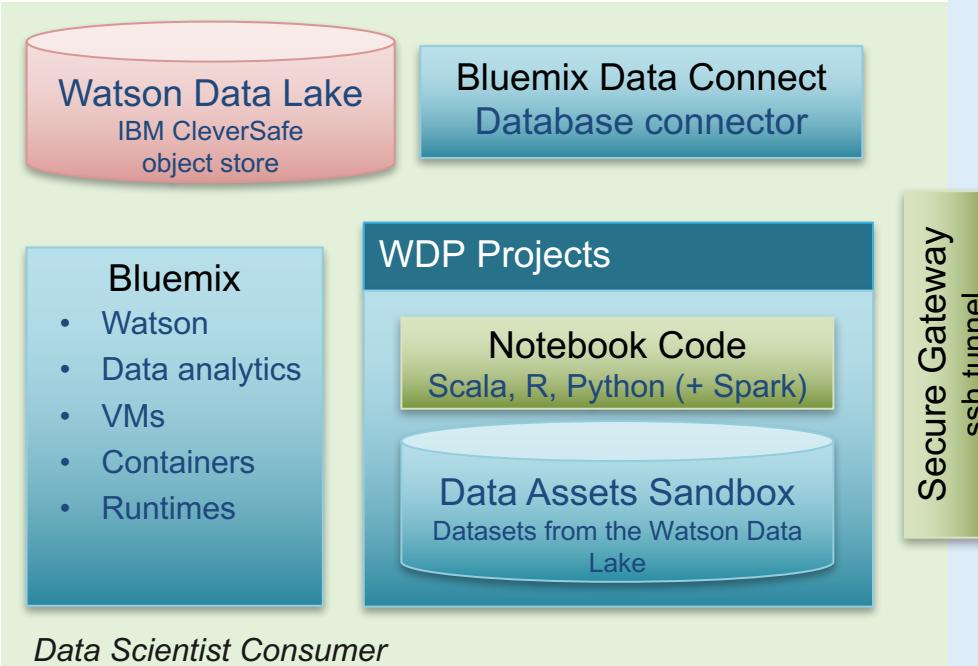
Appendix



DataWorks Integration in CEDP v0.5

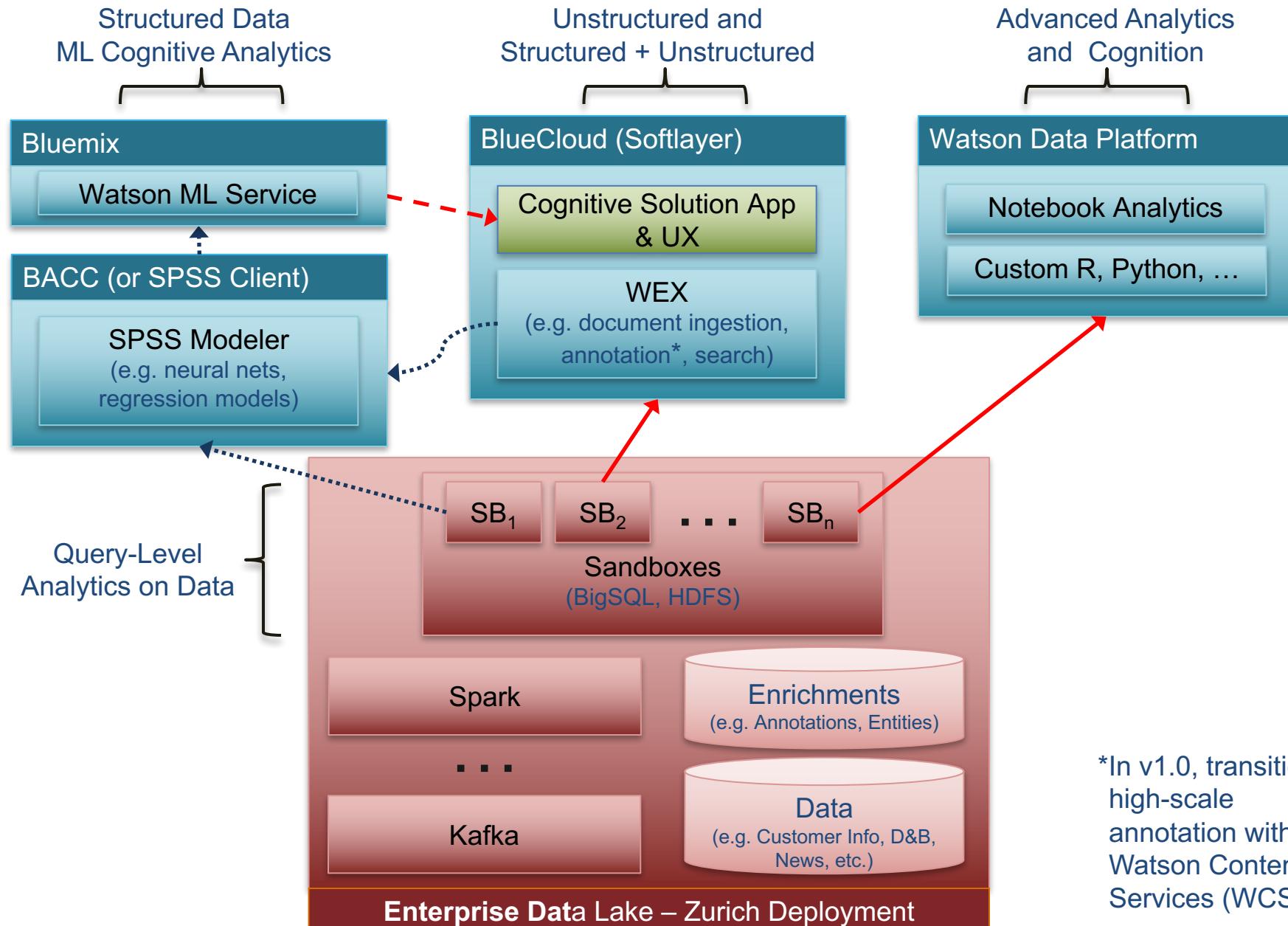
Watson Data Platform – CEDP 0.5 Features

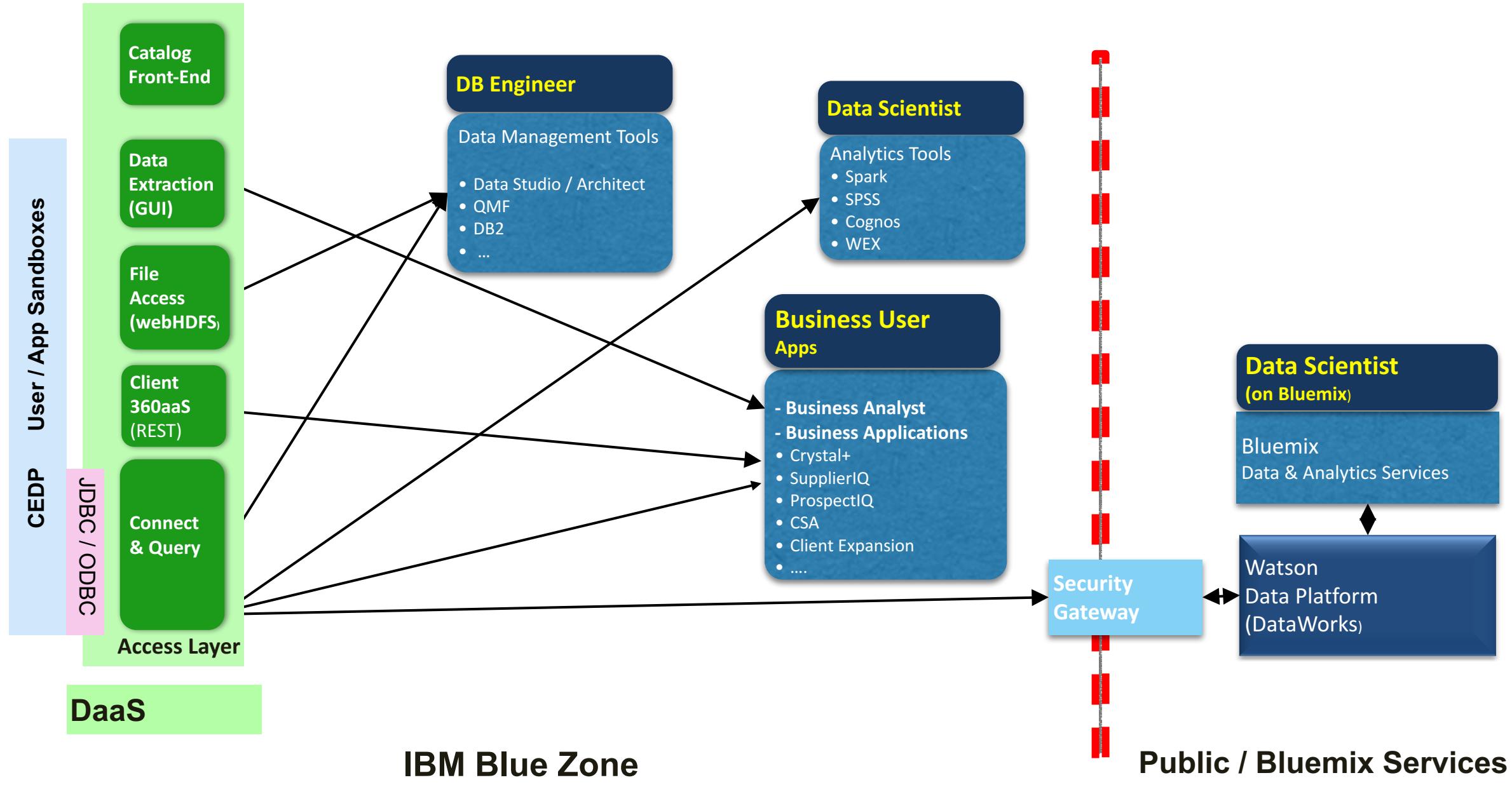
- Enterprise-wide CoEDL datasets (e.g. Client 360)
- User and team sandboxes managed in Bluemix
- Advanced analytics & cognition in notebooks
- Cognitive services from Bluemix Watson
- Access control policies mirrored from CoEDL
- Best-of-breed IBM and open source technologies



Cognition as a Service (CaaS) in CEDP v0.5

DRAFT





CEDP: Features and Terminology

Common Data Platform	Enterprise-wide Catalog	DaaS	CaaS	TaaS	SME Q&A
360-degree Views	Self-Service	Sandbox	Social / Collaborative	Trusted	Secure
Governed	Structured / Unstructured Data	Scale	Intuitive Interface	Cognitive Processes	Personas
	Migration of DataMarts		Deep Learning		



User Experience

DRAFT

IBM CEDP

Data Assets Data Services Catalog

Internal Sources

 RDC / CMR	 BMSIW
 T&P Database	 Social & Environment Management Database
 BDW	 Supplier Performance Evaluation
 Sales Play	 Client Reference

External Sources

 S&P Capital IQ	 D&B	 IDC	 PAOD
 Profondia (JPJ)	 HG Data	 Morningstar	 SEC
 First Research	 Northernlight	 Reuters	 Blekko
 Bing News Search	 Associated Press	 Gartner	 Orbit
 Wikipedia	 Twitter	 RAAD	

- Benefit your business with big data that is linked and up to date >>



What people say

AI Delo
Chief Executive Officer
HC Health and Human Resources

Read More

Reinvent what's possible

Let the data guide you products and experiences competition. Embrace the frontier.

List of actual Data Assets are to be obtained



Maxim Lysak
Developer

Cognitive Enterprise Data Platform



Visit the CEDP Community

Data Assets



- Overview variety of Data Sources that are integrated and linked in CEDP.
- Benefit your business with big data that is linked and up to date >>

Data Services



Ready to use services for:
Business Users, Data Scientists,



IBM CEDP



Maxim Lysak
Developer

Search service by name or description

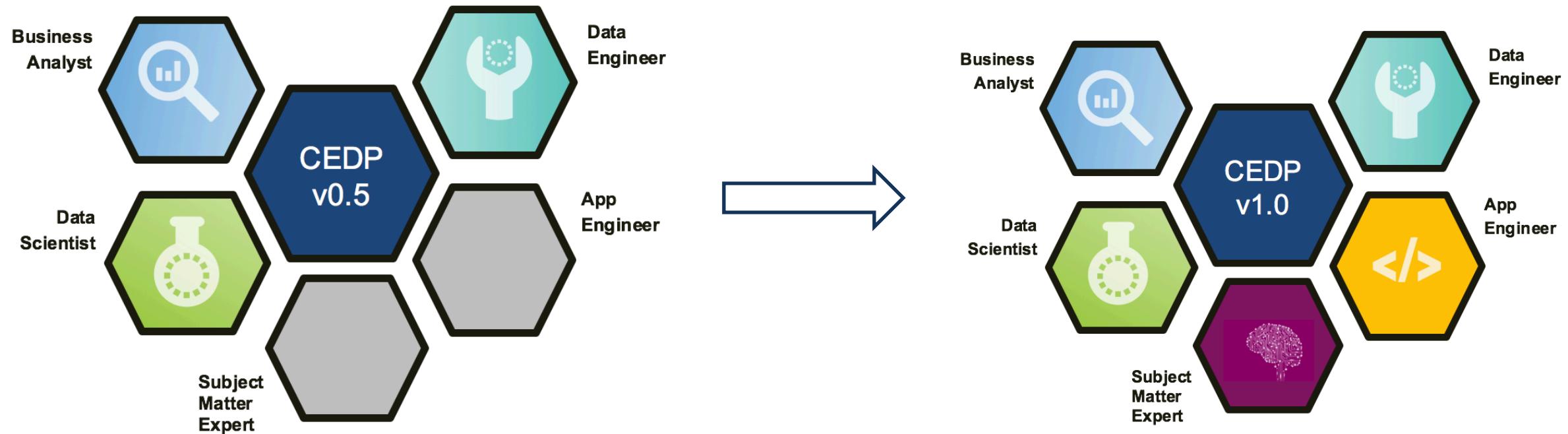
Data Assets

Data Services

Catalog

The screenshot shows the IBM CEDP Data Explorer interface. On the left, there is a sidebar titled "Filters" with dropdown menus for "All Account Levels", "All Countries", "All Industries", and a search bar. Below these are two buttons: "Export List Data" and "Export Company Data". A message indicates "Number of results: 17,334 (0 - 50)". To the right of the sidebar is a large network graph. The graph consists of various company logos (IBM T&IP, IDC, RAAD RESEARCH, G, PTS, D&B, JPJ, HG) connected by lines representing relationships. Labels for these connections include "TIP PARTNERSHIP_TEAM", "TIPLMIS AGREEMENTS", "CP.GARTNER_ESTIMATE", "CP.COMP_CONTRACTS_BY_SODB_V2", "CP.COMP_INSTALL_BY_SEO", "DIL.DUNS_REFERENCE_BY_SEO", "CP.REUTERS_SYMBOL", "CP.COMP_INSTALL_BY_AMC", "CP.COMP_INSTALL_BY_HG", and "CP.COMP_INSTALL_BY_HG". A legend on the left side of the graph area identifies the colors: red for "CPTAGS", blue for "CPGEO", green for "CP.DUNS_REFERENCE_BY_SEO", purple for "ALL_BASE_SCORINGS", and orange for "CP.COMPETITIVE_INSTALL_BASE_BY_IBM".

CEDP support for Personas



CEDP v0.5
(Q4 2016)

CEDP v1.0
(Q4 2017)



CEDP Timeline

