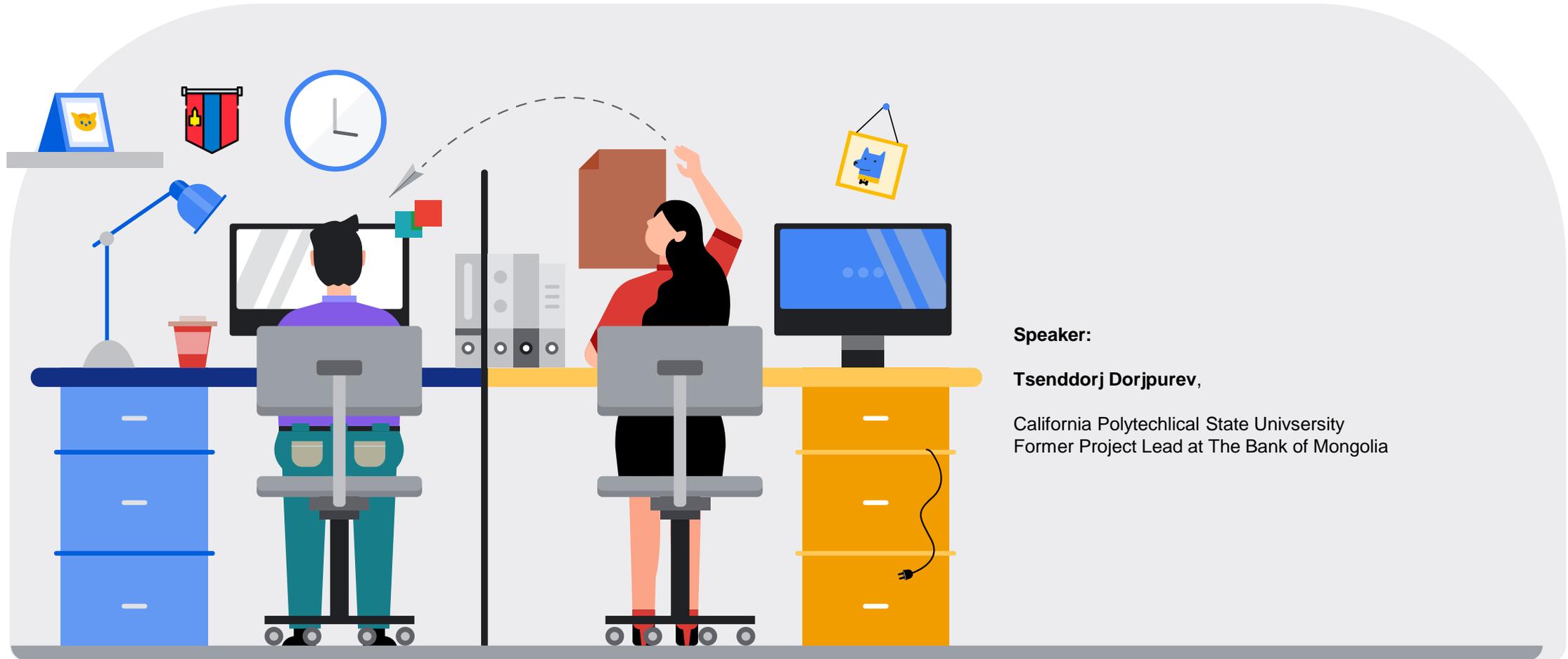




Big Data Platform (FinPulse) Initiative



“Reimagining regulatory data collection and analysis”



Speaker:

Tsenddorj Dorjpurev,

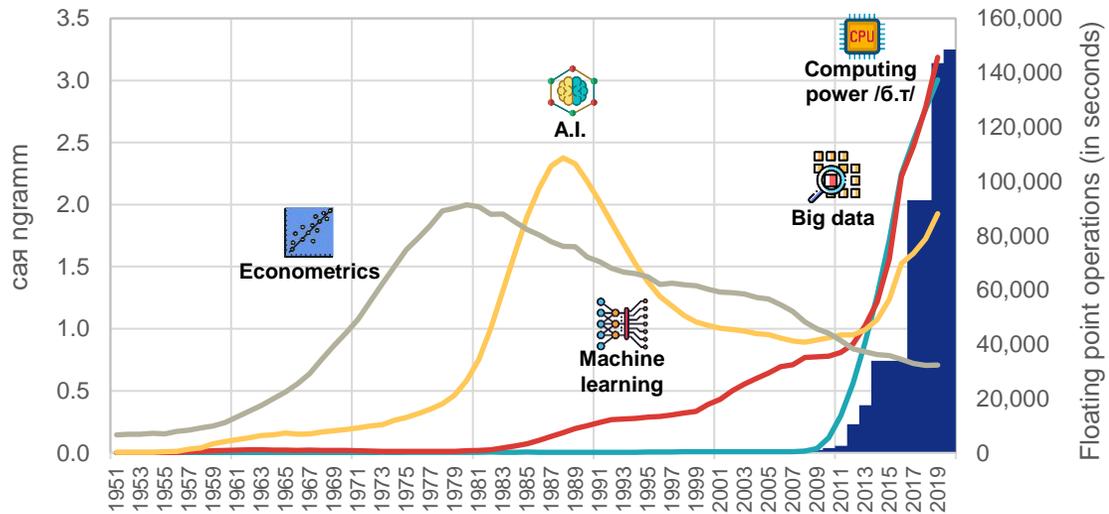
California Polytechnical State University
Former Project Lead at The Bank of Mongolia

17 OCT 2023



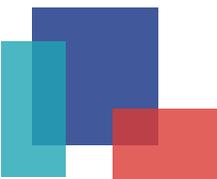
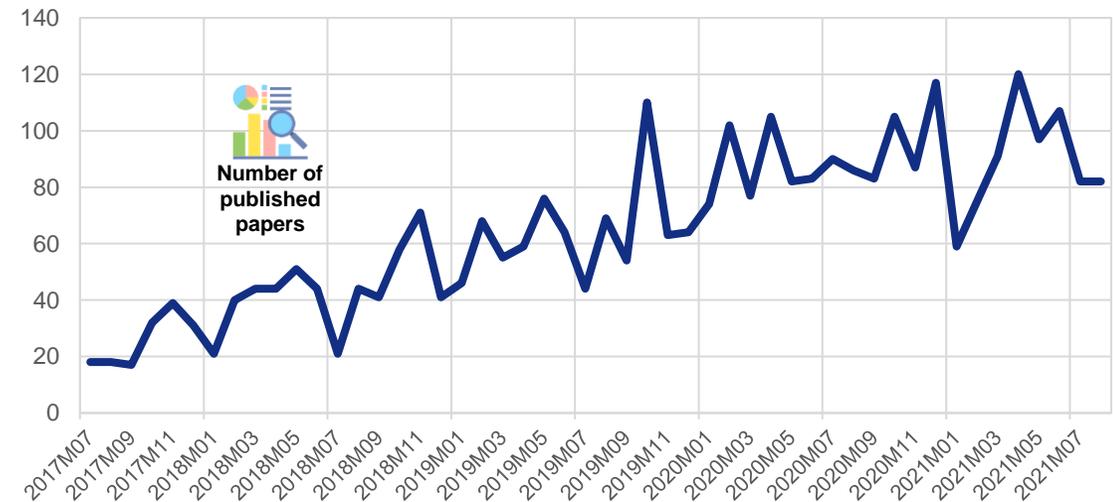
General interest

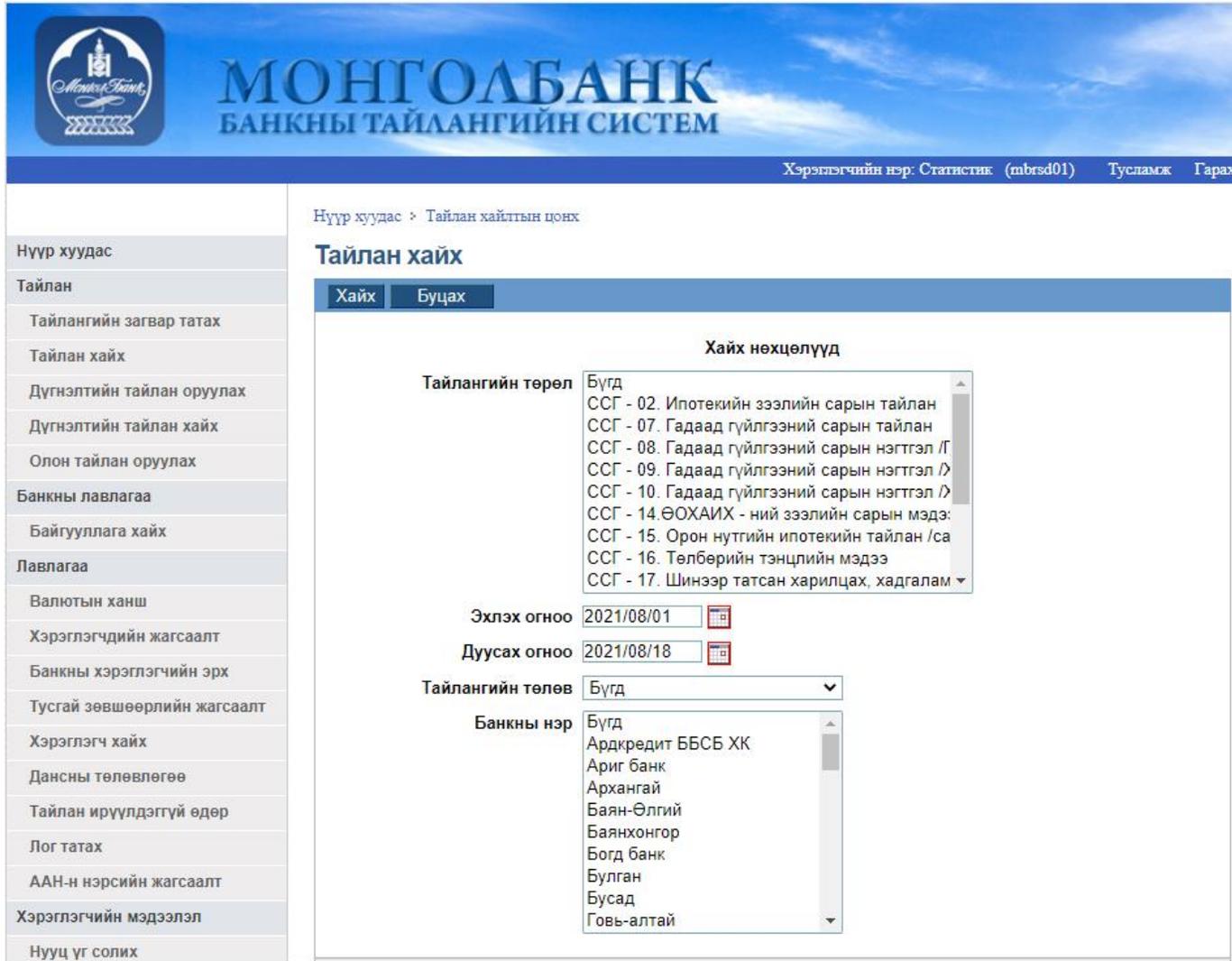
A strong interest in big data can be seen after 2010, where the term machine learning took up in the second half of the '80s and artificial intelligence has measurable volume as early as the '60s, and it has shown a decline through the '90s 2000s and the first half of 2010s as well



Academic interest

The field of Big Data has **3,315 papers (310 authors)** compared with 22,500 articles published in the field of central banking, from a total of 1,001,932 articles in the archive. In the following figure, we can see that **around 100 articles** are being published in this field every month.





МОНГОЛБАНК
БАНКНЫ ТАЙЛАНГИЙН СИСТЕМ

Хэрэглэгчийн нэр: Статистик (mbrsd01) Тусламж Гаргах

Нүүр хуудас > Тайлан хайлтын цонх

Тайлан хайх

Хайх Буцах

Хайх нөхцөлүүд

Тайлангийн төрөл: Бүгд
 ССГ - 02. Ипотекийн зээлийн сарын тайлан
 ССГ - 07. Гадаад гүйлгээний сарын тайлан
 ССГ - 08. Гадаад гүйлгээний сарын нэгтгэл /Г
 ССГ - 09. Гадаад гүйлгээний сарын нэгтгэл /Д
 ССГ - 10. Гадаад гүйлгээний сарын нэгтгэл /И
 ССГ - 14. ӨОХАИХ - ний зээлийн сарын мэдэ:
 ССГ - 15. Орон нутгийн ипотекийн тайлан /са
 ССГ - 16. Төлбөрийн тэнцлийн мэдээ
 ССГ - 17. Шинээр татсан харилцах, хадгалам

Эхлэх огноо: 2021/08/01

Дуусах огноо: 2021/08/18

Тайлангийн төлөв: Бүгд

Банкны нэр: Бүгд
 Ардкредит ББСБ ХК
 Ариг банк
 Архангай
 Баян-Өлгий
 Баянхонгор
 Богд банк
 Булган
 Бусад
 Говь-алтай

+250



+70 templates are received by 5 departments:

RSD

PSD

RMFMD

BRPD

SD

- Duplicated information => Cross check validation
- Increasing frequency of changes in the templates and ad hoc data collection
- Increased fatigue at bank employees due to increasing regulatory data demand
- Demand for higher frequency, cohesive data collection

Project Impact



Creation of a standard, unified database



Reduced time-lag of official statistics



Data collection frequency



Data volume



Increased labor efficiency
(cost reduction and human error)



In Numbers

Condensed **70+ reports** and **250+ sheets** of data templates into a **single standardized data acquisition model**.



Reduced the time between data collection and data dissemination from **t+17 days** to **t+1 day**.



Increased average data collection frequency from **monthly** to **daily**.



Increased recorded annual data volume **x425**, from **40GB** to **17TB** (26GB when data mirroring and more with backup).

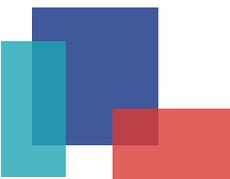


Enabled **200+ workers** in the banking sector to focus on higher-value tasks such as problem solving and analytics.



Improvements of the Entire Central Bank and Partners

- Increased analytical capabilities that supervisors can use to provide policymakers and upper management with instant access to reports and key indicators.
- The standardized data model and staging database setup allow banks to change systems without disrupting the data-sharing connection.



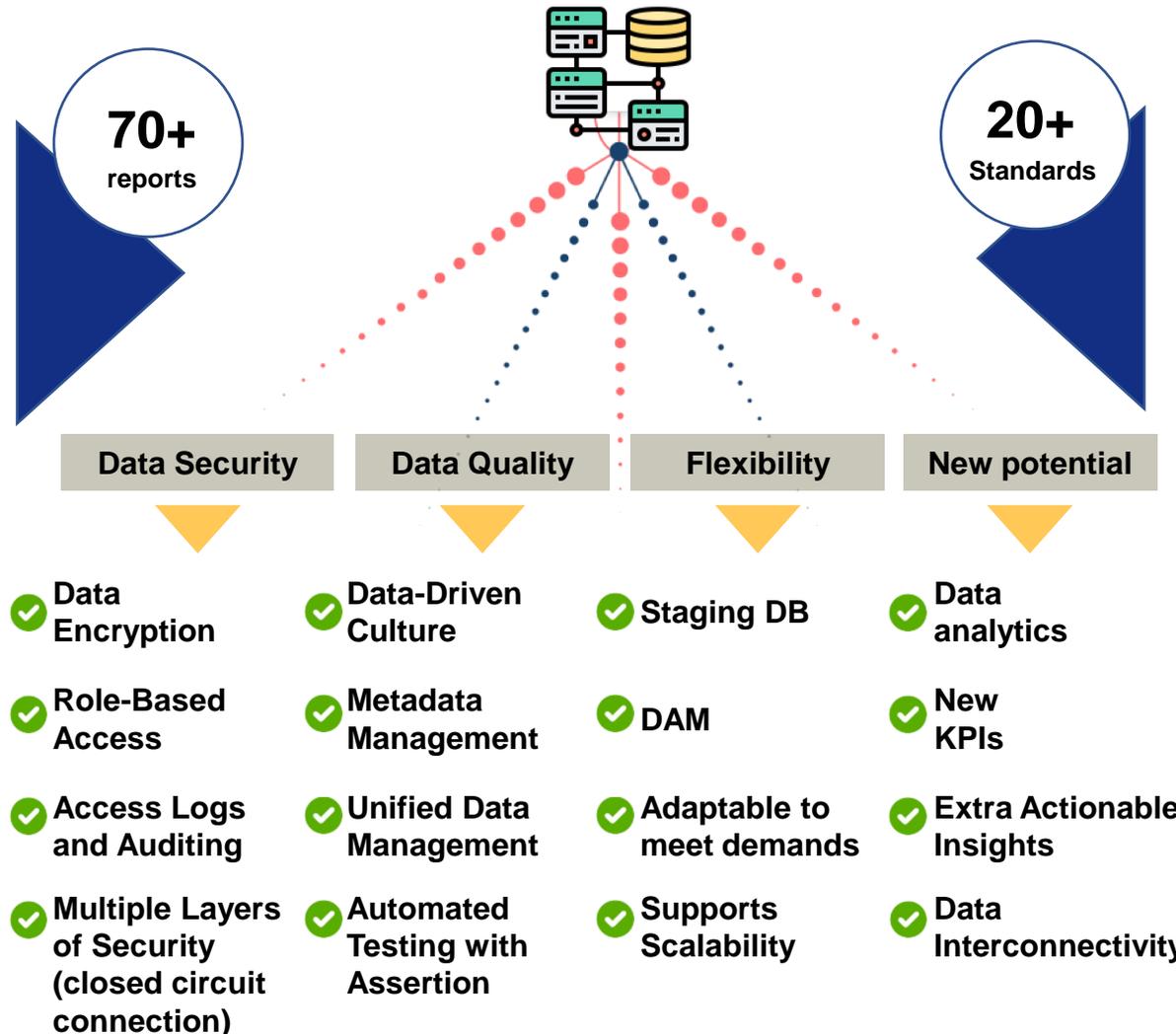

Regulatory reports

- Loan Report of Banks
- SME Loan Report
- Rural-Area Loan Report
- Mortgage Loan Report
- Consolidated Balance Sheet
- Interest Rate Lending Report
- Balance of Payments
- Liquidity Ratio Report

...

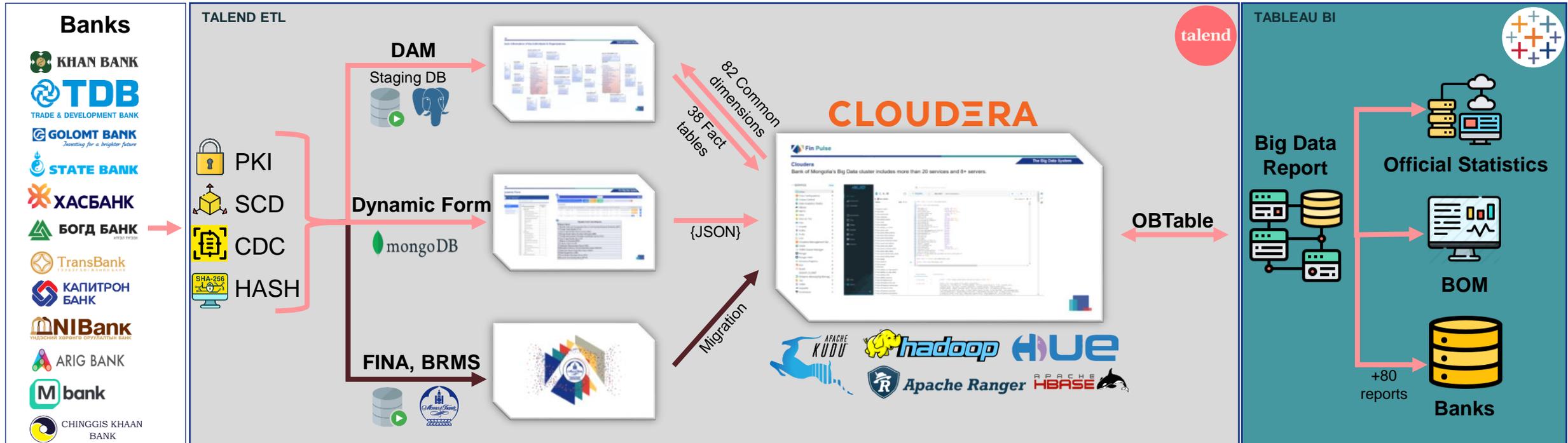


Data Acquisition Model of Big Data (FinPulse) System


Official statistics standards & guidelines

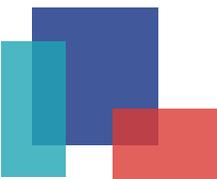

- IFRS
- IMF SDDS
 - Monetary and Financial Statistics Guidelines
 - Balance of Payments Manual
- The System of National Accounts
- Green Taxonomy
- AFI Sex-disaggregated data guidelines
- The International standard industrial classification of all economic activities ISIC
- Merchant Category codes
- Zip postal code

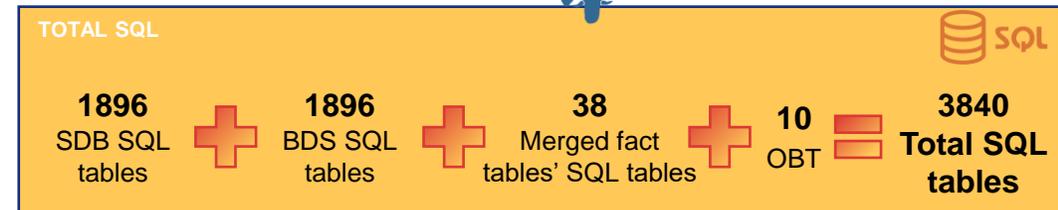
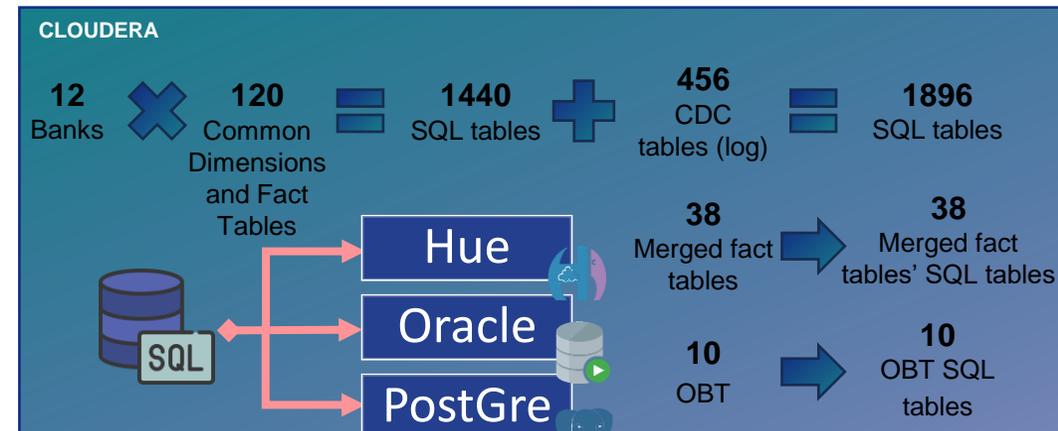
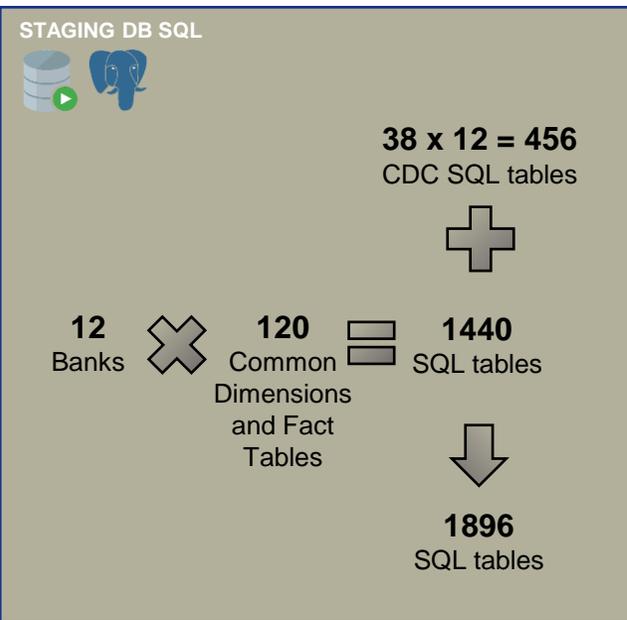
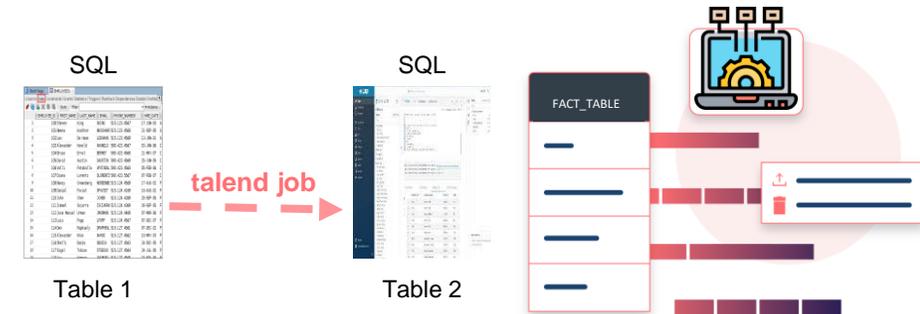
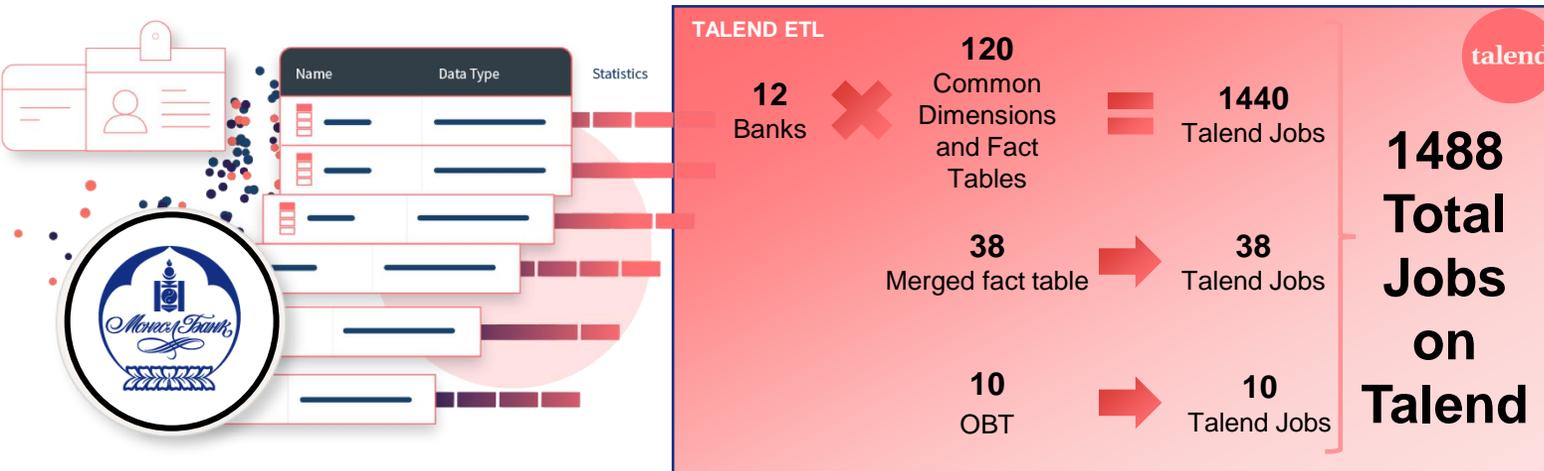
...



The Big Data platform uses staging databases (i.e., parallel warehouse) that are kept on the bank's premises to hold raw, granular-level data using the BoM's Data Acquisition Model (DAM) format which is encrypted and transmitted to the BoM through a closed-circuit connection, and offers a single, unified source of information for both supervisors and banks to access in near real-time. Besides DAM, the platform also collects data from commercial banks and licensed institutions through manually filled Dynamic form.

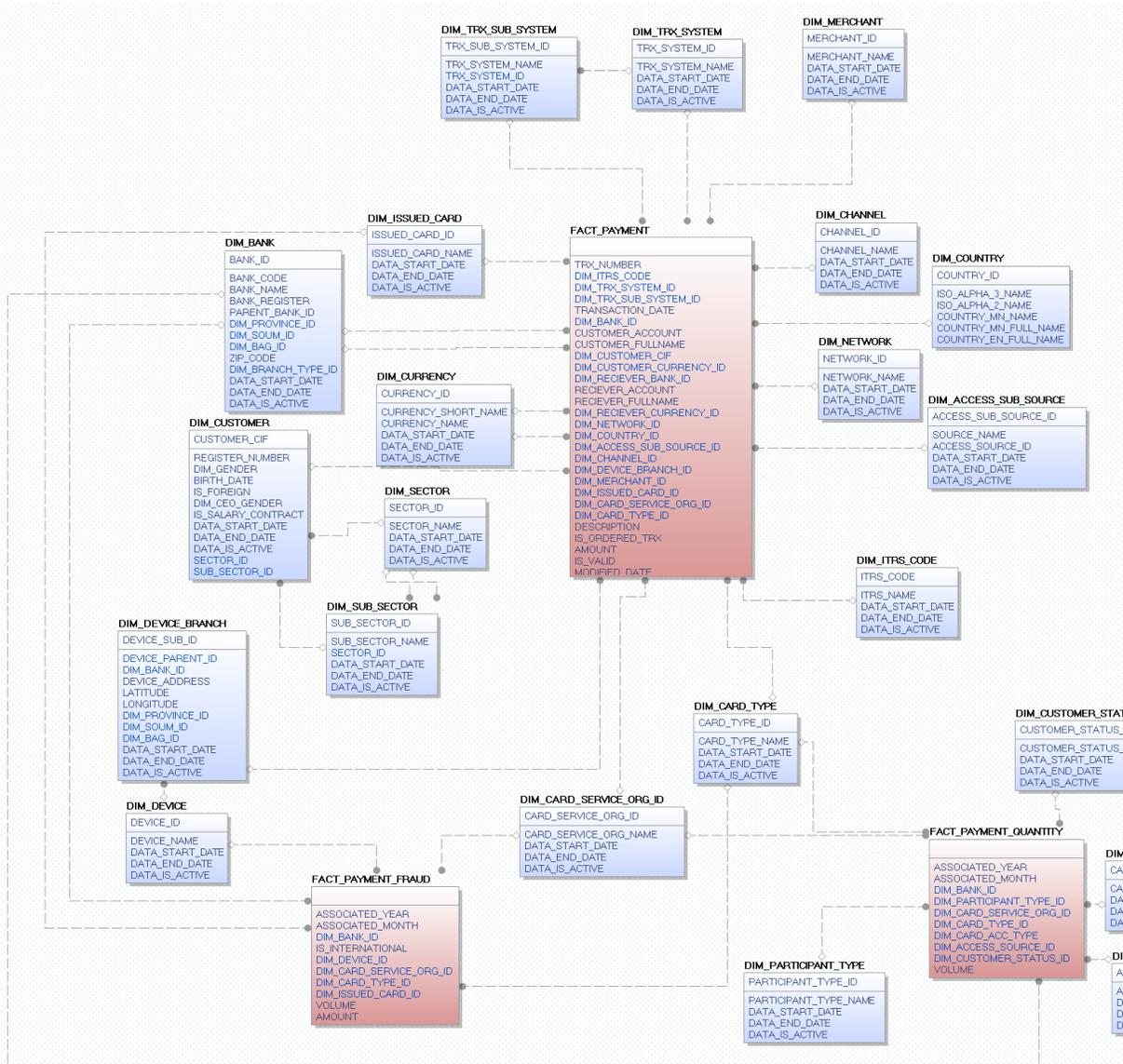
1. **Data Acquisition Model (DAM):** The DAM generates an automated report daily with access in near real-time, eliminates data redundancy and inconsistency both interdepartmentally and system-wide, and can acquire and assimilate more forms of data.
2. **Dynamic Form:** Data that is not readily available through DAM will be collected using a web-based dynamic questionnaire management solution connected to the big data system. Dynamic form uses customizable data templates at a fixed frequency and is easy to make additional changes and updates to the reports.





Payment Information

Data Acquisition Model



FACT_PAYMENT table	COLUMN_NAME	NAME	DESCRIPTION
	TRX_NUMBER	Гүйлгээний дугаар	Банкуудын гүйлгээ бүрийг тодорхойлсон давтагдашгүй ID дугаар
	DIM_ITRS_CODE	ITRS дугаар	Гадаад гүйлгээ хийсэн тохиолдолд DIM_ITRS table-н харгалзах ITRS_CODE оруулна.
	DIM_TRX_SYSTEM_ID	Төлбөр тооцооны системийн төрөл	Банк доторхи гүйлгээ бол банкууд Inner-bank, Гадаад гүйлгээ Foreign transaction төрлөөр тус тус оруулна.
	DIM_TRX_SUB_SYSTEM_ID	Төлбөр тооцооны системийн дэд төрөл	Банк доторхи гүйлгээ бол банкууд Inner-bank, Гадаад гүйлгээ Foreign transaction төрлөөр тус тус оруулна.
	TRANSACTION_DATE	Гүйлгээ хийгдсэн өдөр, цаг, минут	YYYY-MM-DD HHMMSS форматаар байна
	DIM_BANK_ID	Банк, салбар	Харицагч салбар дээр хийгдсэн гүйлгээ бол салбарын кодоор авна.
	CUSTOMER_ACCOUNT	Харицагчийн дансны дугаар	Илгээчийн данс ихдээ 12 оронтой тоо оруулна.
	CUSTOMER_FULLNAME	Харицагчийн нэр	Гадаад гүйлгээ хийсэн тохиолдолд илгээчийн дансны нэрийг текстээр оруулна.
	CUSTOMER_CIF	Харицагчийн CIF	Банканд бүртгэлтэй харицагчийн дахин давтагшгүй CIF дугаарыг оруулна.
	DIM_CUSTOMER_CURRENCY_ID	Харицагчийн дансны валют	DIM_CURRENCY table-н харгалзах CURRENCY_ID оруулна
	EXCHANGE_RATE	Харицагчийн валютын ханш	Валютын ханшын дүнг оруулна
	DIM_RECEIVER_BANK_ID	Хүлээн авагчийн банк	DIM_BANK table-н харгалзах BANK_ID оруулна
	RECEIVER_ACCOUNT	Хүлээн авагчийн дансны дугаар	Хүлээн авагчийн данс ихдээ 12 оронтой тоо оруулна.
	RECEIVER_FULLNAME	Хүлээн авагчийн нэр	Гадаад гүйлгээ хийсэн тохиолдолд илгээчийн дансны нэрийг текстээр оруулна.
	DIM_RECEIVER_CURRENCY_ID	Хүлээн авагчийн дансны валют	DIM_CURRENCY table-н харгалзах CURRENCY_ID оруулна
	DIM_NETWORK_ID	Нэгдсэн болон гадаад сүлжээний төрөл	DIM_NETWORK table-н харгалзах NETWORK_ID оруулна
	DIM_COUNTRY_ID	Хүлээн авагчийн улсын дугаар	DIM_COUNTRY table-ийн COUNTRY_ID оруулна.
	DIM_ACCESS_SUB_SOURCE_ID	Хандалтын эх үүсвэр	DIM_ACCESS_SUB_SOURCE table-н харгалзах ACCESS_SUB_SOURCE_ID оруулна
	DIM_CHANNEL_ID	Гүйлгээний төрөл	DIM_CHANNEL table-н харгалзах CHANNEL_ID оруулна
	DIM_DEVICE_ID	Төхөөрөмжийн төрөл	DIM_DEVICE table-ийн харгалзах DEVICE_ID оруулна
	DIM_TERMINAL_ID	Төхөөрөмжийн терминал дугаар	DIM_DEVICE_BRANCH table-ийн харгалзах TERMINAL_ID оруулна
	DIM_MERCHANT_CODE	Мерчантагийн төрөл	DIM_MERCHANT table-н харгалзах MERCHANT_CODE оруулна
	DIM_ISSUED_CARD_ID	Банкны картын эзэмшлийн төрөл	DIM_ISSUED_CARD table-н харгалзах ISSUED_CARD_ID оруулна
	DIM_CARD_SERVICE_ORG_ID	Картын төрөл	DIM_CARD_SERVICE_ORG table-н харгалзах CARD_SERVICE_ORG_ID оруулна
	DIM_TRANSACTION_TYPE_ID	Картын лэд төрөл	DIM_CARD_TYPE table-н харгалзах CARD_TYPE_ID оруулна
	DESCRIPTION	Гүйлгээний утга	Текст утгаар оруулна
	IS_CASH	Бэлэн мөнгөний гүйлгээ эсэх	1 = Тийм, 0 = Угуй
	IS_ORDERED_TRX	Захиалгат гүйлгээ эсэх	1 = Тийм, 0 = Угуй
	AMOUNT	Гүйлгээний дүн	Тоон утгаар оруулна
	IS_CONFIRM	Хүчинтэй эсэх	Банкууд ITRS кодыг зассан тохиолдолд 1 гэж дугаарлана. Өмнөх буруу утгатай гүйлгээг 0 болгож өөрчилнө.
	MODIFIED_DATE	Өөрчилсөн өдөр, цаг, минут, секунд	YYYY-MM-DD HHMMSS форматаар байна. Гүйлгээний мэдээллийг зассан өдөр цагийг оруулна.

FACT_PAYMENT_QUANTITY table

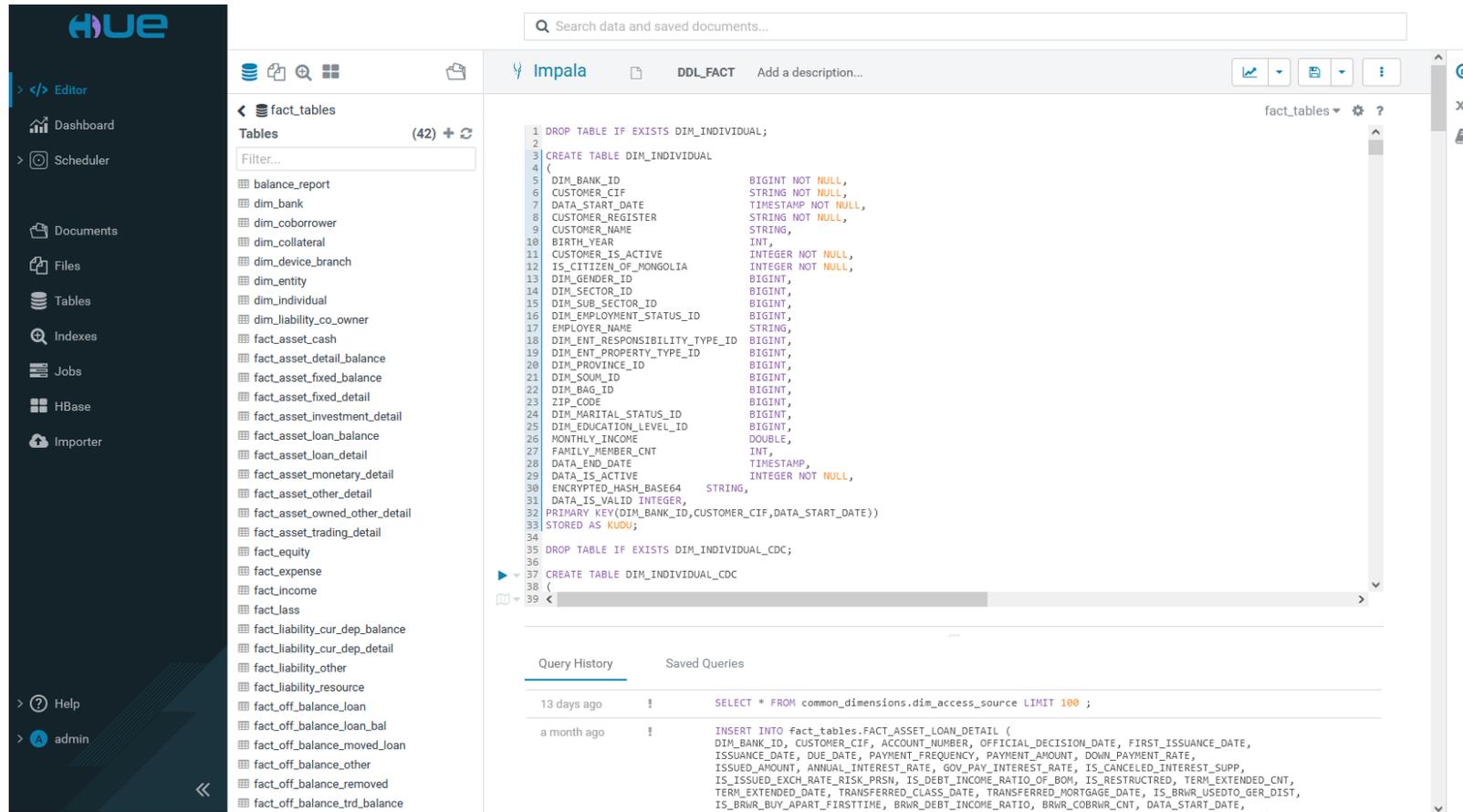
Хөдөлгөөнт банкны хэрэглэгч, картын тоо болон Гэрээт хамтрагчдийн тоон мэдээллийг FACT_PAYMENT_QUANTITY table-руу оруулна.

COLUMN_NAME	NAME	DESCRIPTION
ASSOCIATED_YEAR	Тайлант жил	Сар бүр нэгдсэн тоогоор татах тул тухайн жилийн тоог оруулна
ASSOCIATED_MONTH	Тайлант сар	Сар бүр нэгдсэн тоогоор татах тул тухайн сарын тоог оруулна
DIM_BANK_ID	Банк, салбар	Хэрэглэгчийн бүртгүүлсэн салбар нэгж
DIM_PARTICIPANT_TYPE_ID	Оролцогчдийн хандах төрөл	Карт, хөдөлгөөнт банк, гэрээт хамтрагч гэсэн ангилал бий
DIM_CARD_SERVICE_ORG_ID	Картын үйлчилгээ үзүүлэгчийн карт	DIM_CARD_SERVICE_ORG table-ийн CARD_SERVICE_ORG_ID оруулна
DIM_TRANSACTION_TYPE_ID	Картын төрөл	Зөвхөн картын хэрэглэгчийн тоон дээр нөхөгдөнө
DIM_CARD_ACC_TYPE_ID	Картын дэд төрөл	Зөвхөн картын хэрэглэгчийн тоон дээр нөхөгдөнө
DIM_ACCESS_SOURCE_ID	Хандалтын эх үүсвэр	Зөвхөн хөдөлгөөнт банкны ангилалын ID авч ашиглана
DIM_CUSTOMER_STATUS_ID	Нийт хэрэглэгчийн төлөв	Идэвхитэй, болон нийт, тайлант сард г.м ангилалыг оруулна
VOLUME	Тоо хэмжээ	Тоон утгаар оруулна

Cloudera

Bank of Mongolia's Big Data cluster includes more than 20 services and 8+ servers.

- ▼ SERVICE Clear
-  Atlas 8
 -  Core Configuration 8
 -  Cruise Control 1
 -  Data Analytics Studio 2
 -  HBase 7
 -  HDFS 8
 -  Hive 8
 -  Hive on Tez 8
 -  Hue 2
 -  Impala 6
 -  Kafka 5
 -  Kudu 8
 -  Livy 8
 -  Cloudera Management Ser... 1
 -  Oozie 2
 -  YARN Queue Manager 1
 -  Ranger 2
 -  Ranger KMS 1
 -  Schema Registry 8
 -  Solr 8
 -  Spark 8
 -  SQOOP_CLIENT 8
 -  Streams Messaging Manag... 1
 -  Tez 8
 -  YARN 8
 -  Zeppelin 1
 -  ZooKeeper 3



The screenshot displays the Hue interface. On the left, a sidebar lists various services and their counts. The main area is divided into two panes. The top pane shows a list of tables under the 'fact_tables' schema, with 42 tables listed. The bottom pane is an SQL editor for Impala, showing a DDL script for creating a table named 'DIM_INDIVIDUAL'. The script includes a 'DROP TABLE IF EXISTS' statement, followed by a 'CREATE TABLE' statement with various columns and data types. Below the editor, there are sections for 'Query History' and 'Saved Queries'.

Tables List:

- balance_report
- dim_bank
- dim_coborrower
- dim_collateral
- dim_device_branch
- dim_entity
- dim_individual
- dim_liability_co_owner
- fact_asset_cash
- fact_asset_detail_balance
- fact_asset_fixed_balance
- fact_asset_fixed_detail
- fact_asset_investment_detail
- fact_asset_loan_balance
- fact_asset_loan_detail
- fact_asset_monetary_detail
- fact_asset_other_detail
- fact_asset_owed_other_detail
- fact_asset_trading_detail
- fact_equity
- fact_expense
- fact_income
- fact_lass
- fact_liability_cur_dep_balance
- fact_liability_cur_dep_detail
- fact_liability_other
- fact_liability_resource
- fact_off_balance_loan
- fact_off_balance_loan_bal
- fact_off_balance_moved_loan
- fact_off_balance_other
- fact_off_balance_removed
- fact_off_balance_trd_balance

SQL Script:

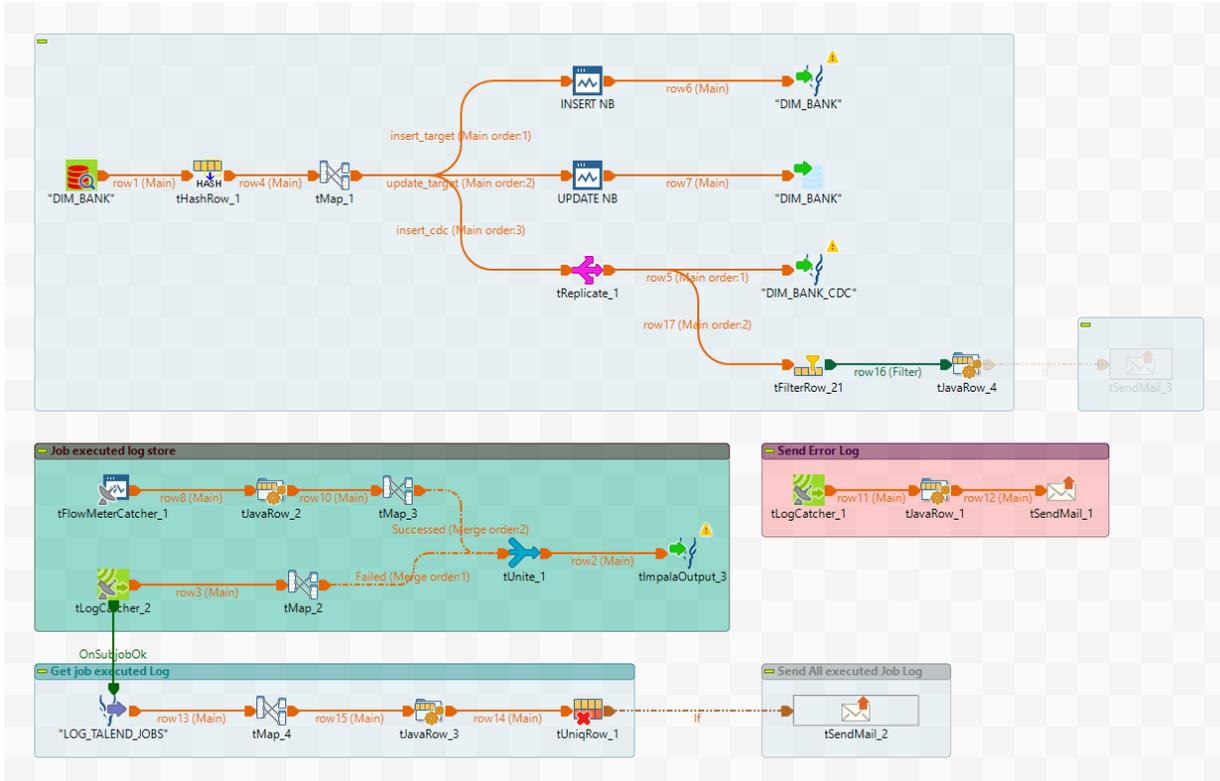
```

1 DROP TABLE IF EXISTS DIM_INDIVIDUAL;
2
3 CREATE TABLE DIM_INDIVIDUAL
4 (
5   DIM_BANK_ID          BIGINT NOT NULL,
6   CUSTOMER_CIF         STRING NOT NULL,
7   DATA_START_DATE    TIMESTAMP NOT NULL,
8   CUSTOMER_REGISTER   STRING NOT NULL,
9   CUSTOMER_NAME       STRING,
10  BIRTH_YEAR           INT,
11  CUSTOMER_IS_ACTIVE  INTEGER NOT NULL,
12  IS_CITIZEN_OF_MONGOLIA INTEGER NOT NULL,
13  DIM_GENDER_ID       BIGINT,
14  DIM_SECTOR_ID       BIGINT,
15  DIM_SUB_SECTOR_ID   BIGINT,
16  DIM_EMPLOYMENT_STATUS_ID BIGINT,
17  EMPLOYER_NAME       STRING,
18  DIM_ENT_RESPONSIBILITY_TYPE_ID BIGINT,
19  DIM_ENT_PROPERTY_TYPE_ID BIGINT,
20  DIM_PROVINCE_ID     BIGINT,
21  DIM_SOUM_ID         BIGINT,
22  DIM_BAG_ID          BIGINT,
23  ZIP_CODE            BIGINT,
24  DIM_MARITAL_STATUS_ID BIGINT,
25  DIM_EDUCATION_LEVEL_ID BIGINT,
26  MONTHLY_INCOME     DOUBLE,
27  FAMILY_MEMBER_CNT  INT,
28  DATA_END_DATE      TIMESTAMP,
29  DATA_IS_ACTIVE     INTEGER NOT NULL,
30  ENCRYPTED_HASH_BASE64 STRING,
31  DATA_IS_VALID      INTEGER,
32  PRIMARY KEY (DIM_BANK_ID, CUSTOMER_CIF, DATA_START_DATE)
33  STORED AS KUUU;
34
35 DROP TABLE IF EXISTS DIM_INDIVIDUAL_CDC;
36
37 CREATE TABLE DIM_INDIVIDUAL_CDC
38 (
39

```

Talend Big Data

Big data system of the Bank of Mongolia has installed the TAC with 4 servers and 10 cores.



JOB CONDUCTOR

Refresh Add Duplicate Deploy Run Kill
 Enable auto refresh Delete Pause task(s) Resume task(s) Recover last execution Show

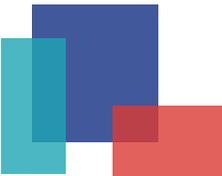
State	Status	Error status	Label	Trigg...	Actions	Time left ...	Last run	Project
Project: BOM (6 items)								
	Ready to run		SERV1-KERBEROS-AU...			4h 49min ~	2022-11-23 08:5...	BOM
	Ready to run		SERV2-KERBEROS-AU...			4h 42min ~	2022-11-23 08:5...	BOM
	Ready to run	Connection t...	SERV3-KERBEROS-AU...				2022-11-20 16:3...	BOM
	Ready to run	Connection t...	SERV4-KERBEROS-AU...				2022-11-19 16:5...	BOM
	Ready to depl...		AB_DIM_ACCESS_SOU...				2022-11-18 16:5...	BOM
	Ready to run		DIM_AB_ALL					BOM

Page 1 of 1 | 10

Triggers Context parameters JVM parameters

Refresh Add trigger... Duplicate Delete Pause trigger Resume trigger

Label	Type	Status	Previous trig...	Time left befo...	Next triggerin...	Next Triggeri...	Number of t
	Add simple trigger						
	Add CRON trigger						
	Add file trigger						



Dynamic Form

A

Оюунтолгойн сарын мэдээ (ОТ)
Нэр: — Тайлант огноо: |

Хайх

Гадаад дах дансны мэдээлэл

I. Хаягийн хэсэг

II. Үндсэн хэсэг

- Экспортын үзүүлэлт
- Гадаад дах дансны мэдээлэл
- Худалдааны зээл /Экспорт/

	A	B	C	D
1	2. Гадаад дахь данс /бэлэн мөнгө, харилцах/			Мөнгөн дүн (ман. ам доллар)
2	2.1 Тайлант хугацааны эхний үлдэгдэл			
3	2.2 Дансны орлого			
4	Эдийн засгийн харьяат бусаас			
5	Хувьцаат капиталд оруулсан хөрөнгө оруулалт			
6	Хөрөнгө оруулагчаас авсан зээл			
7	Монголын банкин дэх өөрийн данснаас зузаатгал			
8	Экспортын орлого (зардлын нөхөн төлбөр)			
9	Экспортын урьдчилгаа			
10	Гадаадаас авсан зээл			
11	Бусад			
12	Эдийн засгийн харьяатаас			
13	Үйлчилгээний төлбөр			
14	Бараа бүтээгдүүний төлбөр			
15	Бусад			
16	2.3 Дансны зарлага			
17	Эдийн засгийн харьяат бус руу			
18	Импортын төлбөр (барааны)			
19	Импортын төлбөр (цахилгааны)			
20	Үйлчилгээний төлбөр			
21	Хөрөнгө оруулагчаас авсан зээлийн төлбөр			
22	Гадаадаас авсан зээл, түүний хүүгийн төлбөр			
23	Бусад			
24	Эдийн засгийн харьяат руу			
25	Монголын банкин дэх өөрийн данс руу зузаатгал			
26	Үйлчилгээний төлбөр			
27	Бараа бүтээгдүүний төлбөр			
28	Бусад			
29	2.4 Тайлант хугацааны эцсийн үлдэгдэл			

B

СУДАЛГАА

Аж ахуй нэгжийн үзүүлэлт /Төлбөрийн тэнцэлийн мэдээ/ (BOP)
2022-11-03 2022-11-08

1 Эхлээгүй
2 Үргэлжилж байгаа
0 Дууссан

3 Үйл ажиллагаа явуулж байгаа

Хайлтыг цэвэрлэх

Дугаар	Нэр	ISIC	Үйл ажиллагааны төрөл	Аймаг/Хот	Сум/Дүүрэг	Мэргэжилтэн	Эхлэх	Дуусах	Үйл ажиллагаа явуулж байг	Мэдээний төлөв	Үйлдэл
2075377	Голомт банк	6419	Мөнгөний зуучлалын бусад үйл ажиллагаа	Улаанбаатар	Чингэлтэй	S.Sudlaach1, Судлаач	2022-11-03	2022-11-08	Үйл ажиллагаа явуулж байгаа	Үргэлжилж байгаа	👁
2693321	ХААН банк	6419	Мөнгөний зуучлалын бусад үйл ажиллагаа	Улаанбаатар	Сүхбаатар	S.Sudlaach1, Судлаач	2022-11-03	2022-11-08	Үйл ажиллагаа явуулж байгаа	Үргэлжилж байгаа	👁
2578697	ХАС банк төв	6419	Мөнгөний зуучлалын бусад үйл ажиллагаа	Улаанбаатар	Сүхбаатар	S.Sudlaach1, Судлаач	2022-11-03	2022-11-08	Үйл ажиллагаа явуулж байгаа	Эхлээгүй	👁

C

"Dynamic Form" List of Reports

No	Report Name
1	Monthly Sales and Transportation Data of Coal Exporting Mongolian Enterprises (MST)
2	Foreign Trade Statistics (FTS)
3	Oyun Tolgoi Monthly Export Forecast (OME)
4	Mortgage Bonds Held by the Bank of Mongolia (MBH)
5	Foreign and Domestic Ownership Concentration Survey (FDO)
6	Oyun Tolgoi Monthly News (OT)
7	Balance of Payments (BOP)
8	E-Money Report (E-MR)
9	Settlement Representative Report (SRR)
10	Mongolian Embassy's Abroad Expenditure Report (MEAER)
11	Diplomatic Mission Expenditure Report (DMEP)
12	State Budget News (SBN)
13	Cross-Border Passenger Survey (CPS)
14	Payment Card Clearing Report (PCCR)
15	Correspondent Bank News (CBN)

