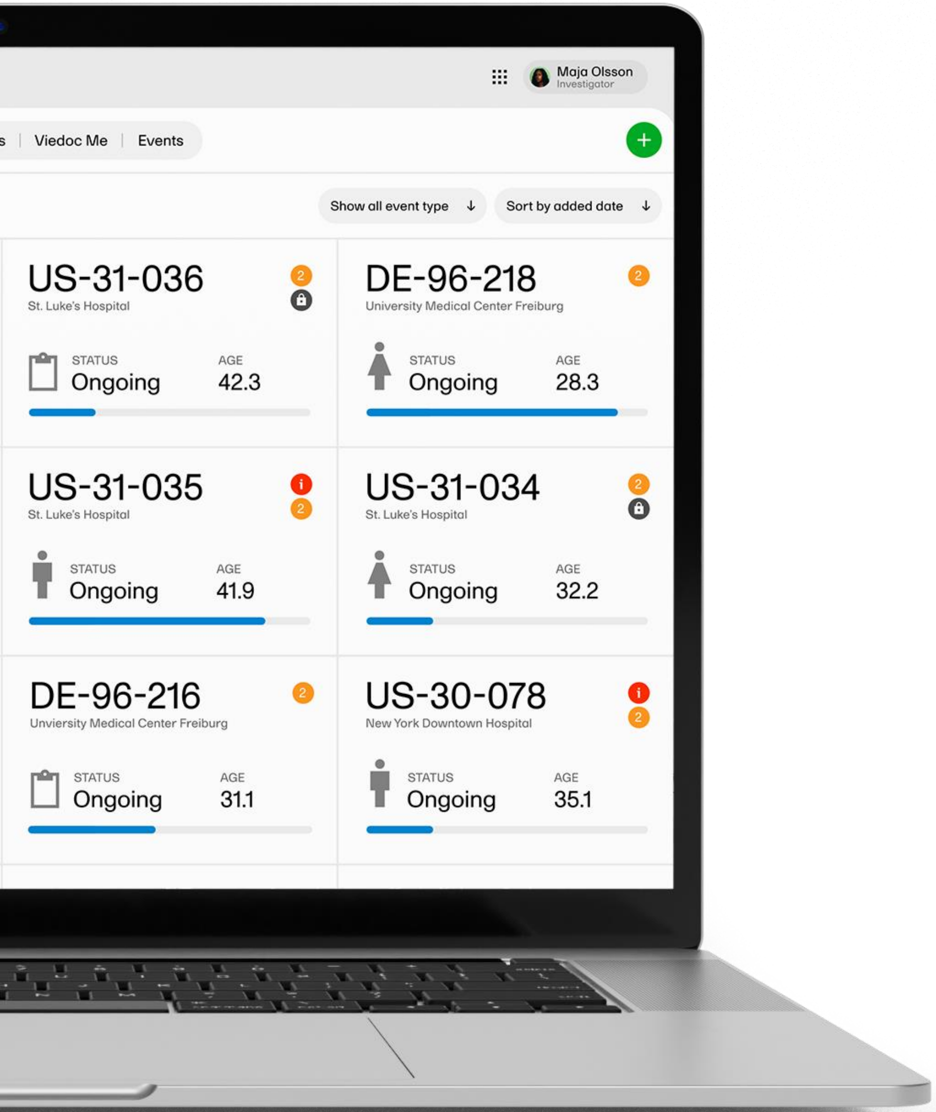
Background & challenges

Paper-based data entry into CRFs has evolved into sophisticated electronic data capture systems (EDCs)

- The transition from paper-based CRFs to sophisticated EDC systems has revolutionized data management in clinical trials
- EDC systems became the gold standard in clinical data management
- Adopted industry-wide, improving efficiency and accuracy

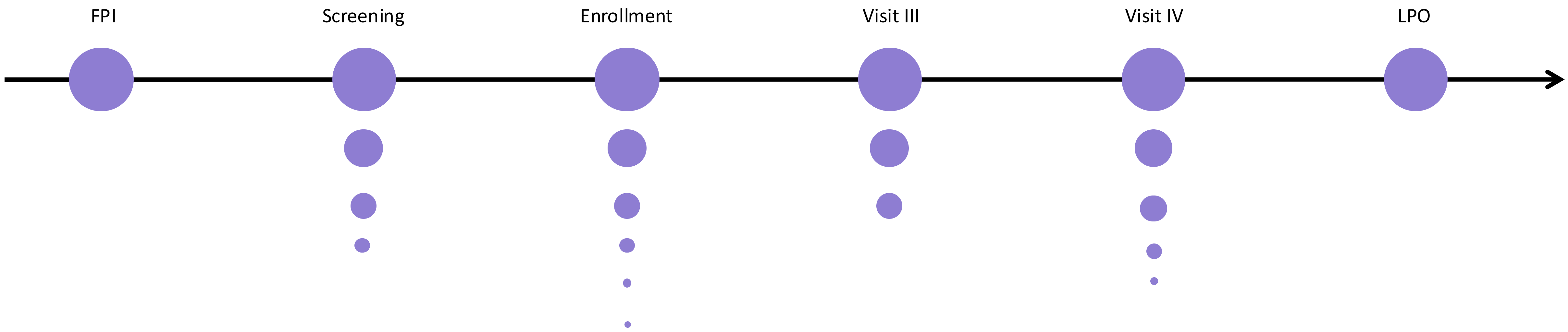


Today's EDC systems emphasize data quality, integrity, and regulatory compliance

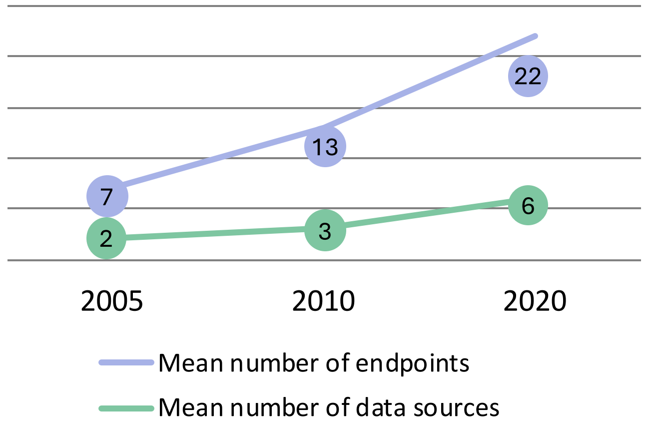
- Designed around efficient data entry
- Automated validations for improved data quality and integrity
- Support collaboration through secure data access and control
- Include comprehensive audit trails for full accountability
- Align with key regulatory standards, including FDA 21 CFR Part 11

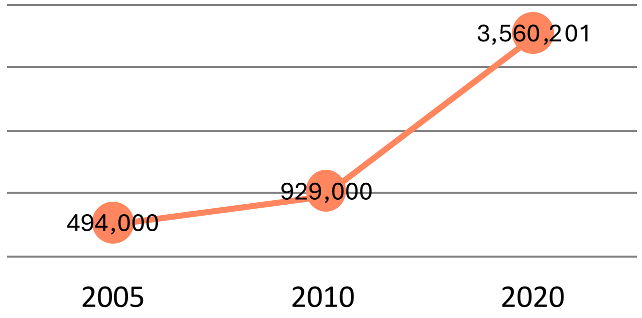
Despite advances, manual data entry remains a critical component of clinical research

- EDC systems focus on specific visits and forms
- Data management often siloed within a study's lifecycle (FPI to LPO), with little to no integration with other systems
- Source data verification (SDV) remains essential for ensuring data accuracy



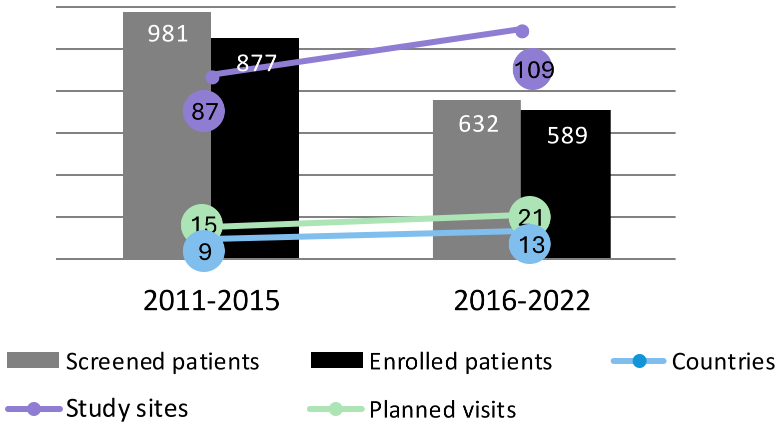
Clinical trials face rising complexity and data volumes alongside challenges in attracting and retaining patients





Average number of data points in clinical trials

Phase III clinical trials



Phase III clinical trial
3.6 million data points (3x compared to last decade)
~26 endpoints (>1 primary endpoint)
13 countries, 87 study sites
Endpoints ↑, Procedures ↑, Countries ↑, Sites ↑, Duration ↑, Cost ↑
Number of recruited patients ↓

- Marked increase in trial complexity, especially in Phase III
- Rapid growth in data volume
- Rising costs, longer timelines, and challenges in patient enrollment

Tufts Center for the Study of Drug Development, Tufts University - Impact report vol 23, number 1 (Jan/Feb 2021) and vol 25, number 3 (May/June 2023)

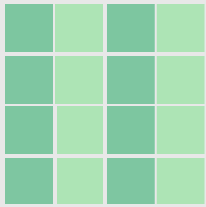
The surge in clinical trial data volume demands innovative collection, analysis, and data management strategies

EDC
ePRO / eCOA
Lab
Imaging

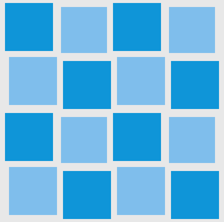
Clinical trial data

EHR
Wearables / sensors
Genomics / biomarkers

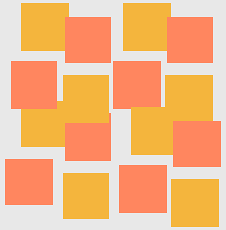
Data standardization



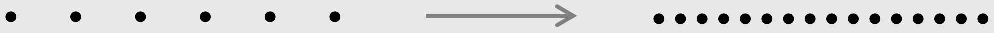
Structured Data



Semi-structured Data



Unstructured Data



Data points

Data streams

Current EDC systems face limitations in managing expanding data sources and volumes

Over

70%

of clinical data is now collected outside traditional EDC systems

77%

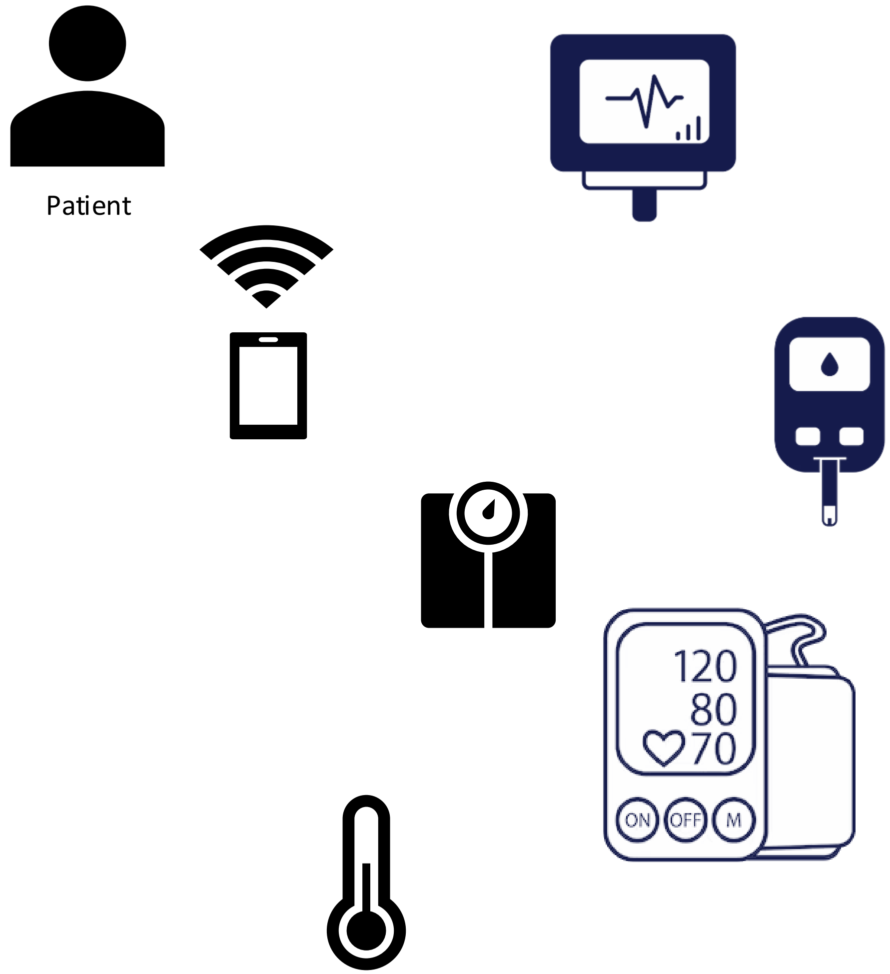
of sponsors and CROs cite difficulty loading data into their primary EDC system

- Data volume and diversity strain EDC capabilities, affecting timelines and efficiency
- Sponsors and CROs use an average of six applications to support clinical trial activities
- Challenges in integrating diverse data sources across clinical applications - 77% of sponsors and CROs cite difficulty loading data into their primary EDC system
- Over 70% of clinical data is now collected outside EDC systems, complicating data management
- Traditional manual data checks and queries lag behind
- Need for holistic data management strategies beyond traditional EDC functions

Tufts Center for the Study of Drug Development, Tufts University - Impact report vol 20, number 1 (Jan/Feb 2018)

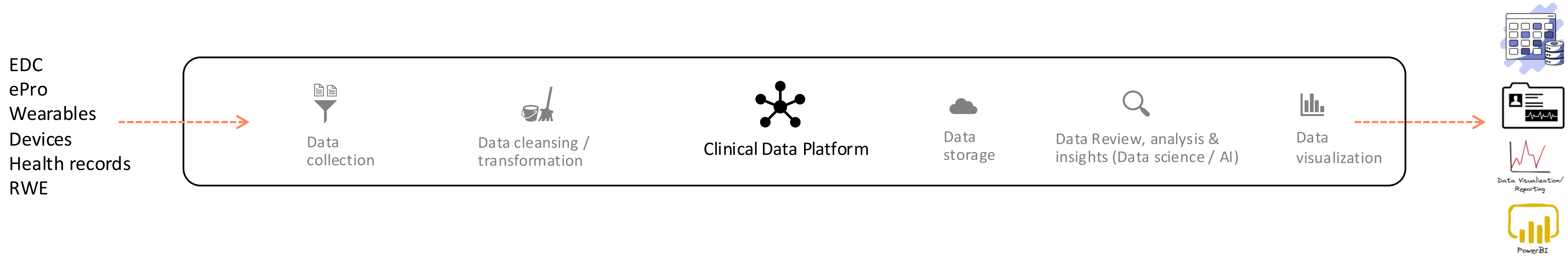
Automated data capture (eSource) emerges as a key solution to reduce manual data entry, enhance efficiency, and limit time and cost overruns

- Direct transfer of data from electronic sources to trial database
- Reduces manual data entry and associated errors
- Eliminates redundancy of double data entry
- Eliminates the need for source data verification
- Decreases workload on site staff
- Simplifies patient participation in trials



Towards Innovative Solutions

The transition from electronic data capture to comprehensive clinical data platforms is essential for future trials

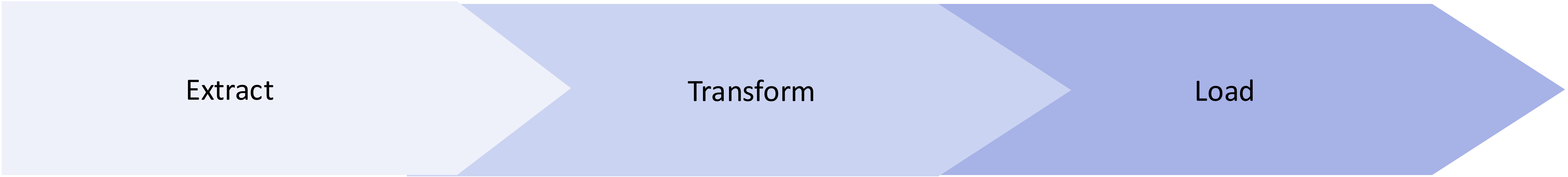


- Data integration from diverse sources for seamless data management
- Real-time support for diverse and continuous data types
- Advanced data management using AI for efficient data handling
- Enhanced decision-making and patient safety with sophisticated analytics
- Supports clinical data managers evolving into data scientists
- Maintain rigorous standards of security, access control, and data integrity
- Investigators retains control and oversight over patient data and have the necessary tools to monitor and act on insights efficiently.

Data pipelines streamline data collection, processing, and integration

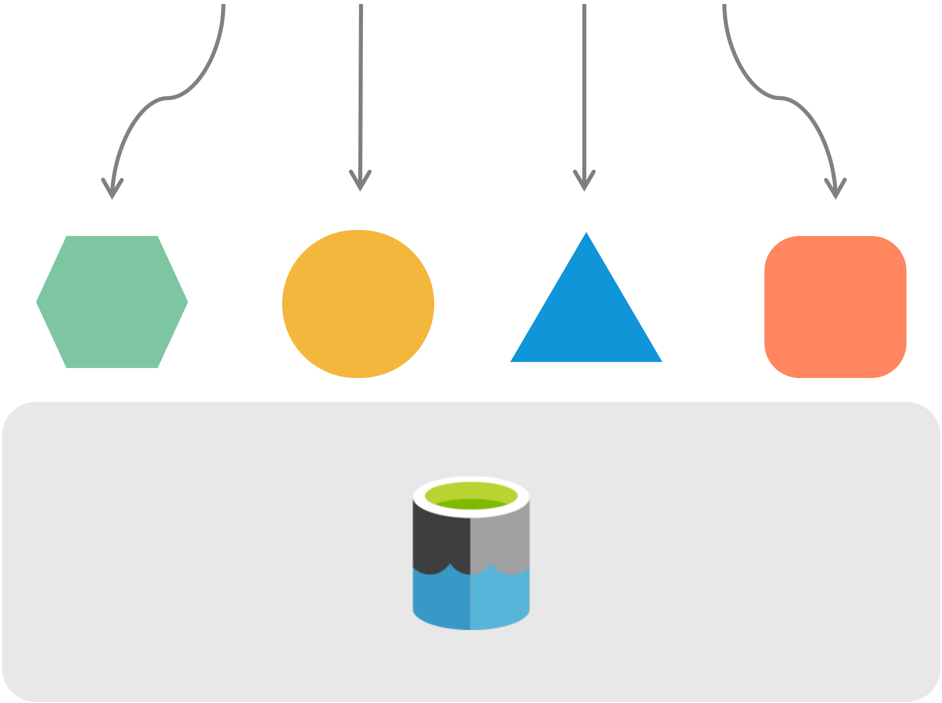
Data pipelines (ETL pipelines) automate data handling from various sources…

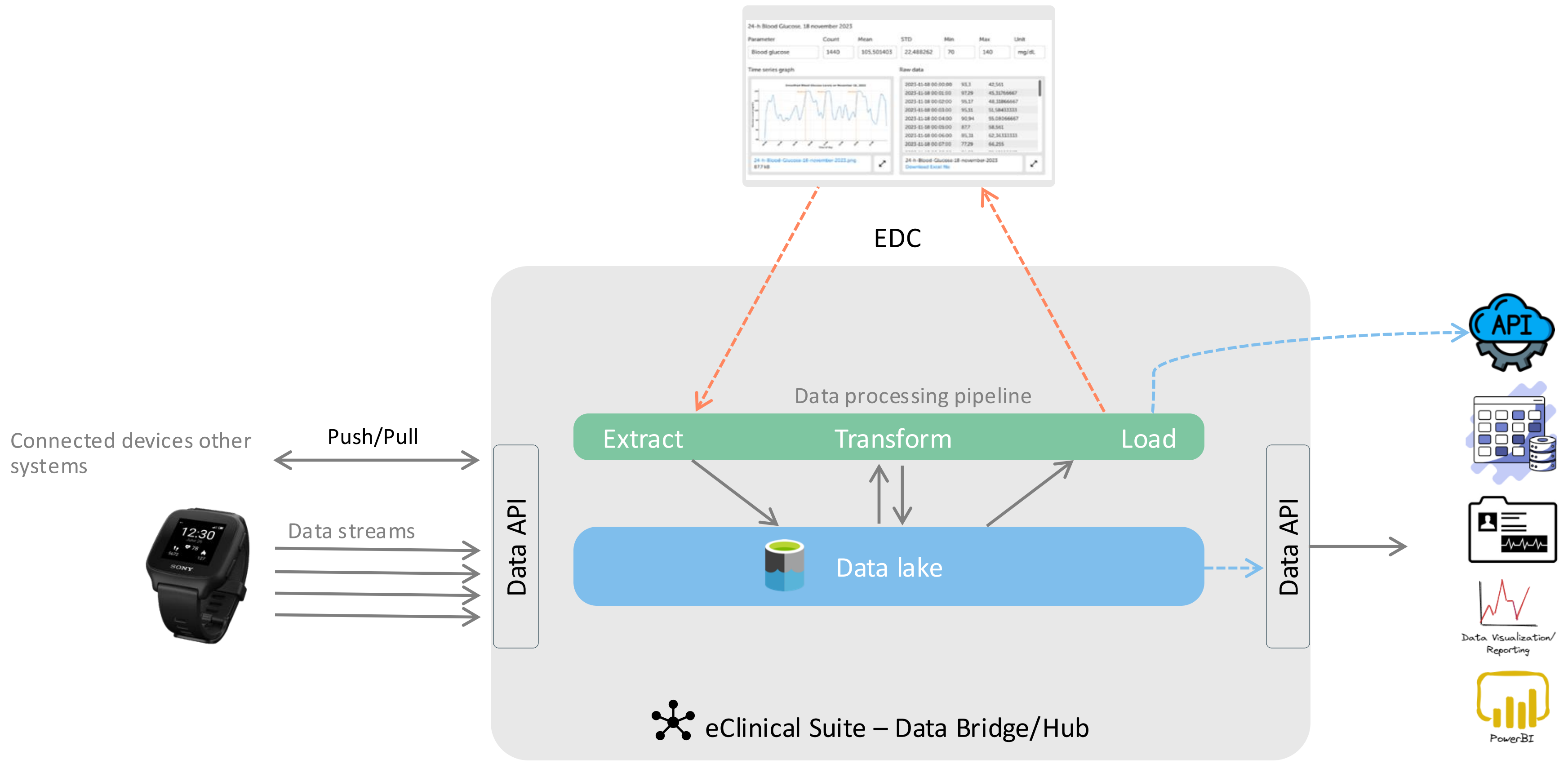
- Streamline data ingestion
- Facilitate data transformation and standardization
- Supporting Real-Time Data Access and Analysis
- Enabling Efficient Data Movement



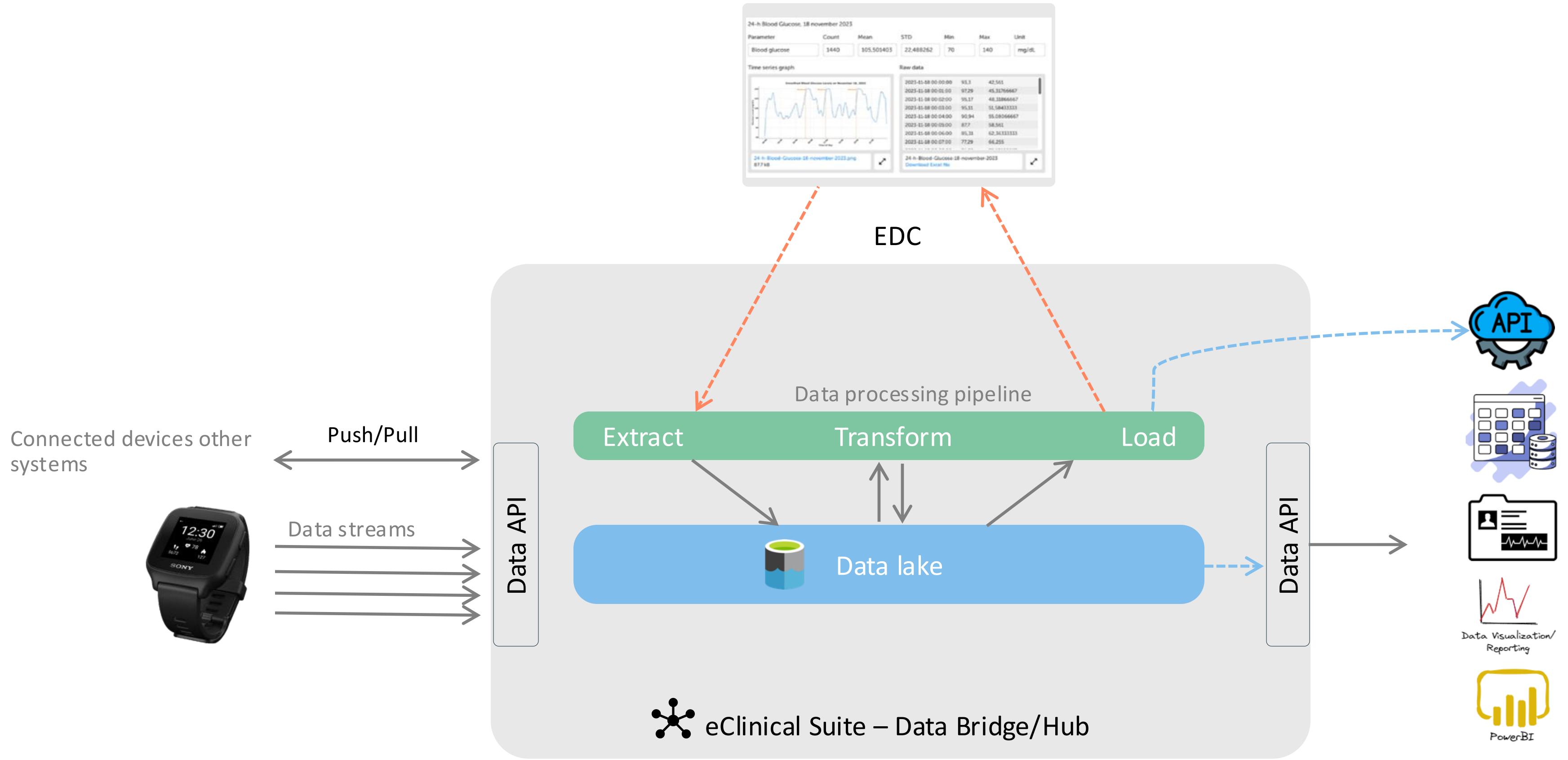
Data lakes address the volume, diversity, and aggregation of clinical trial data through scalable storage

- Accommodating diverse data formats
- Facilitating scalable data storage
- Enhancing data accessibility and discovery
- Enabling Comprehensive Data Analysis
- Ensuring Data Security and Compliance

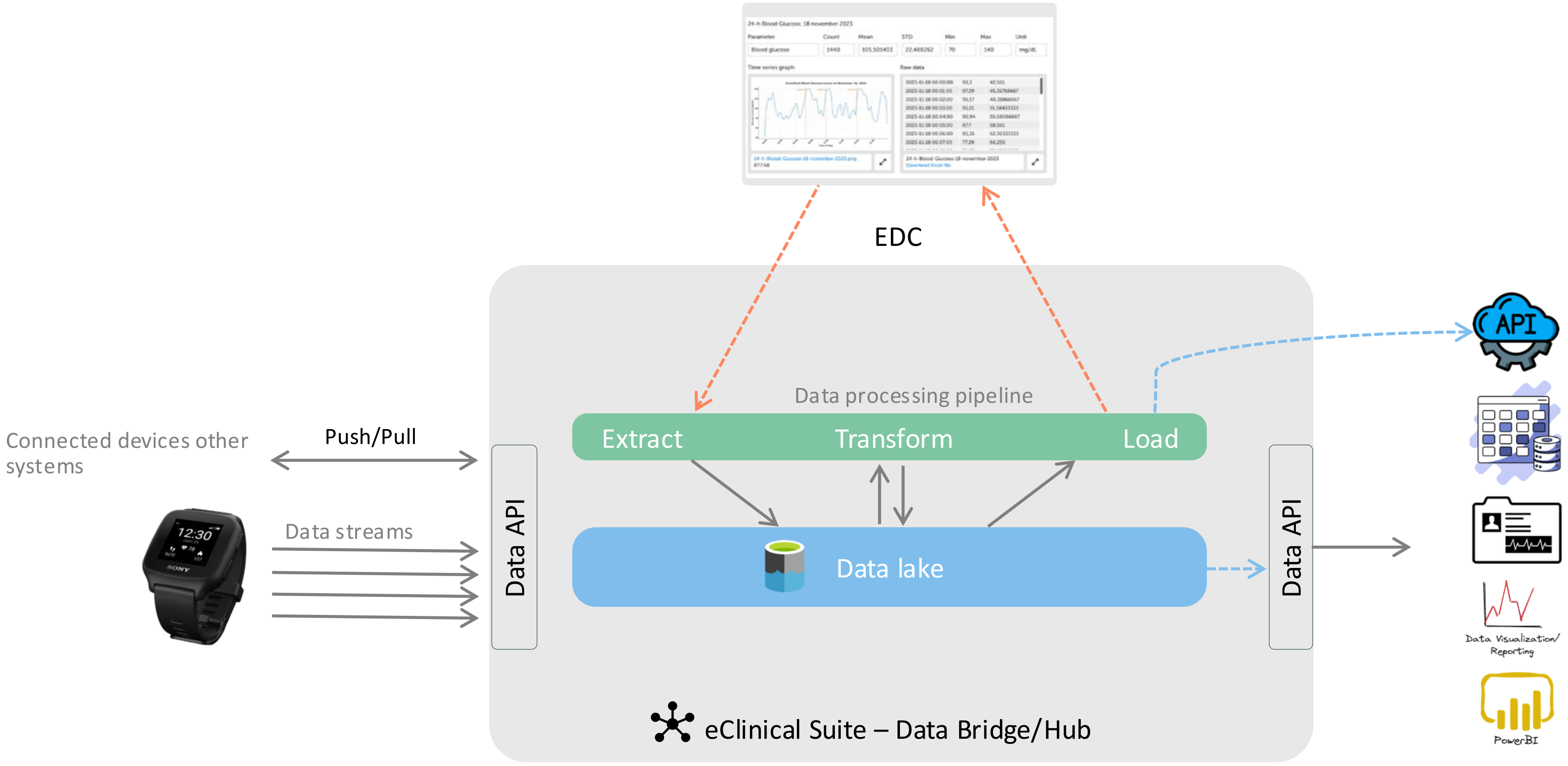




Data Pipelines	Data lake	Custom connectors	Empowerment of data scientists	Enhanced EDC functionality	Integration of wear-able device data	Comprehensive data utilization



	EDC	Clinical Data Platform
SCOPE	Study centric	Data collection centric
DATA INGESTION	Manual data entry	Automatic data collection
DATA STRUCTURE	Focus on data entry (form mode)	Focus on data aggregation and listings (tabular mode)
DATA SCHEMA	Relational database: schema on save	Data lake: schema on read



Platform for integrations	Repeatability	Custom code (Python and R)	Reusability

Basic
(Python or R)

: custom code

Tasks

: get data from system X
: transform X to Y
: push data to system Y

Workflows

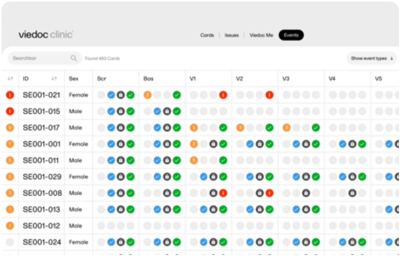
: end-to-end integrations
(various tasks)

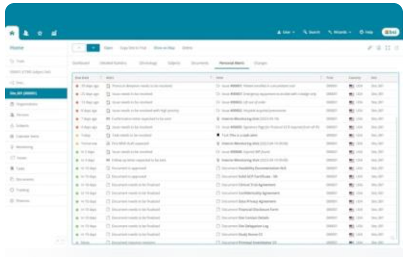
No programming required
Customizable
Repurposable

Examples

CTMS integration case study:
Streamlining clinical trial management through automated data pipelines

Scheduled trigger





Extract → **Transform** → **Load**

Extract

Trigger a Viedoc clinical data export:
POST/api/data export/start

Extract the exported CSV files in the collection data lake
BSI1000101/Raw

Transform

CSV-to-CSV mapping:
Data transformation to align with the destination APIs

Transformed data is also stored in the collection data lake
BSI1000101/Processed

Load

PUT:
/api/dataloader/trials/BS1000101/subjects/{{subjectId}}

- Loop through CSV data rows
- Transform to JSON { }
- Authenticate with the destination API
- Push the data

CTMS integration case study:
Streamlining clinical trial management through automated data pipelines

Extract → **Transform** → **Load**

Extract

Trigger a Viedoc clinical data export:
POST/api/data export/start

Extract the exported CSV files in the collection data lake
BSI1000101/Raw

Transform

CSV-to-CSV mapping:
Data transformation to align with the destination APIs

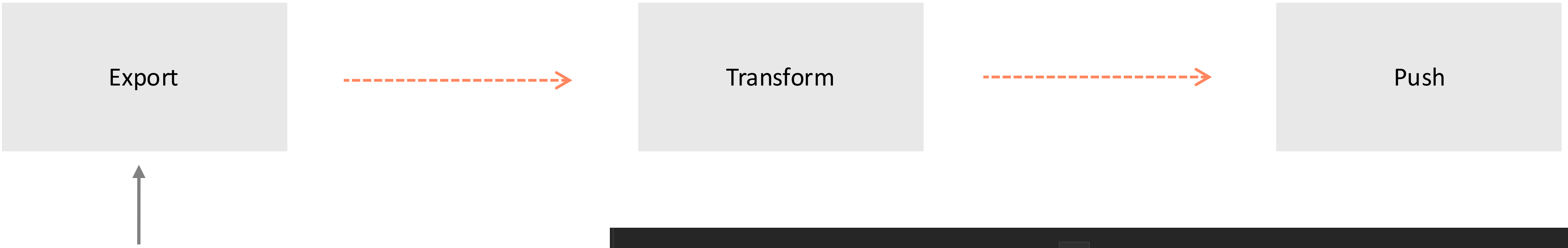
Transformed data is also stored in the collection data lake
BSI1000101/Processed

Load

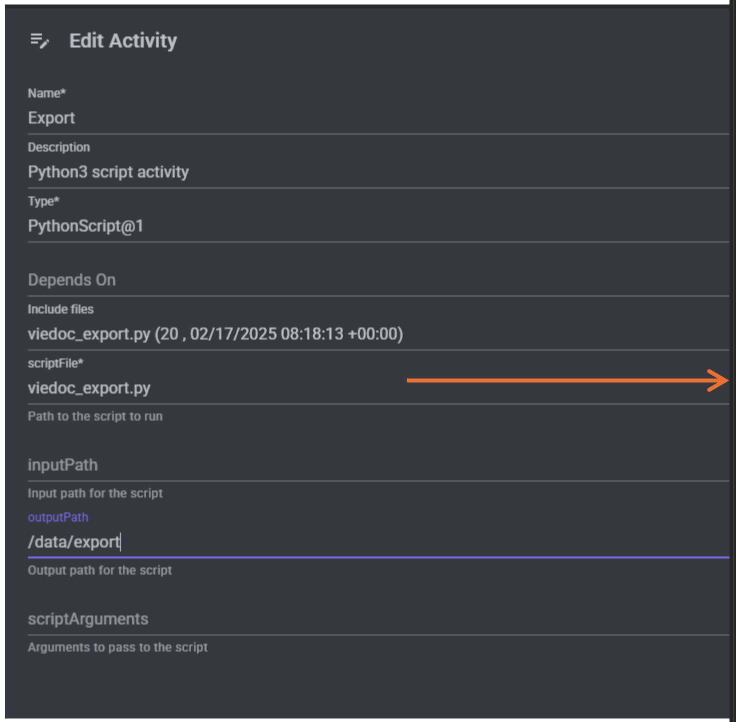
PUT:
/api/dataloader/trials/BS1000101/subjects/{{subjectId}}

- Loop through CSV data rows
- Transform to JSON { }
- Authenticate with the destination API
- Push the data

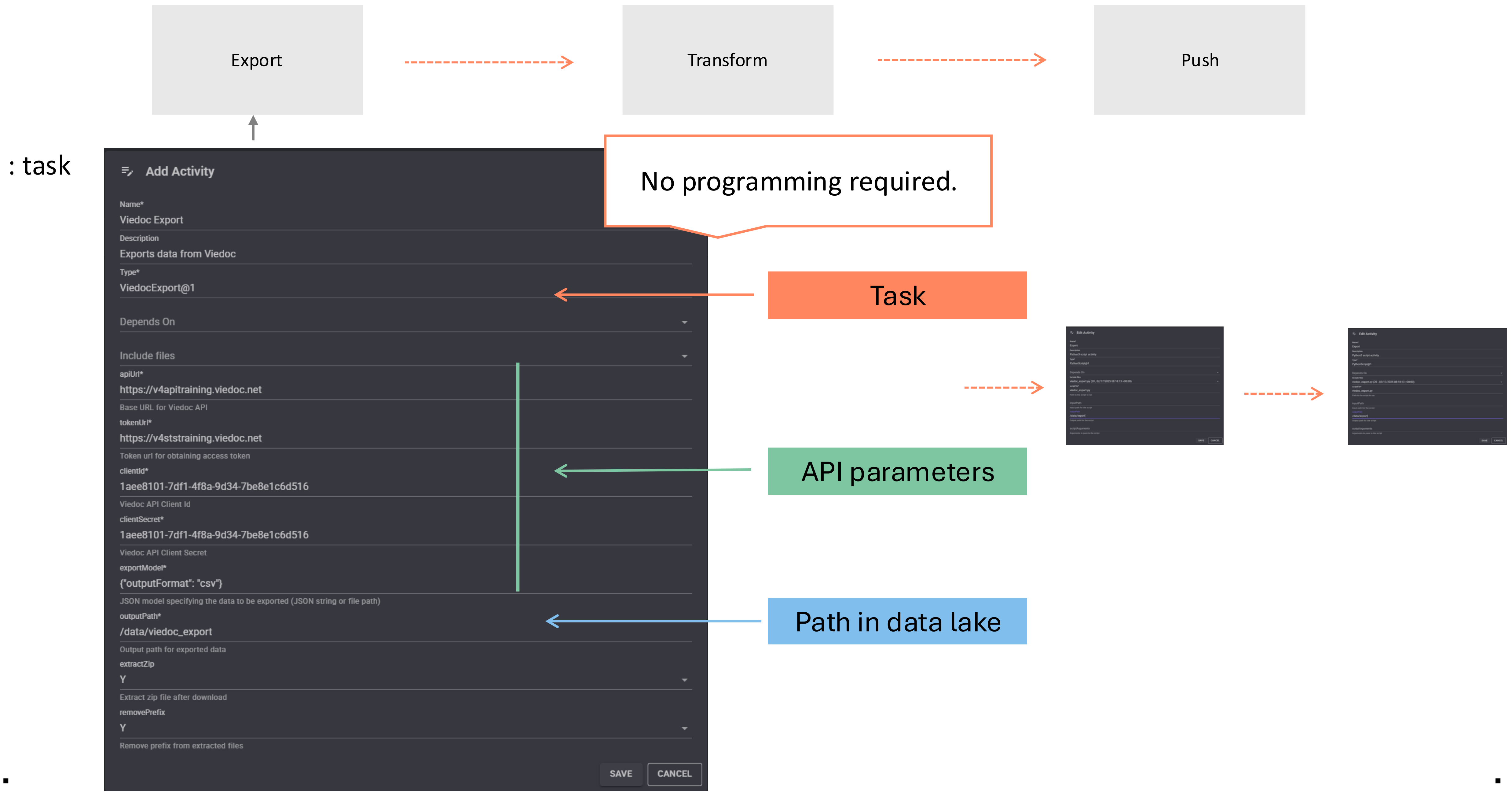
Export → Transform → Push



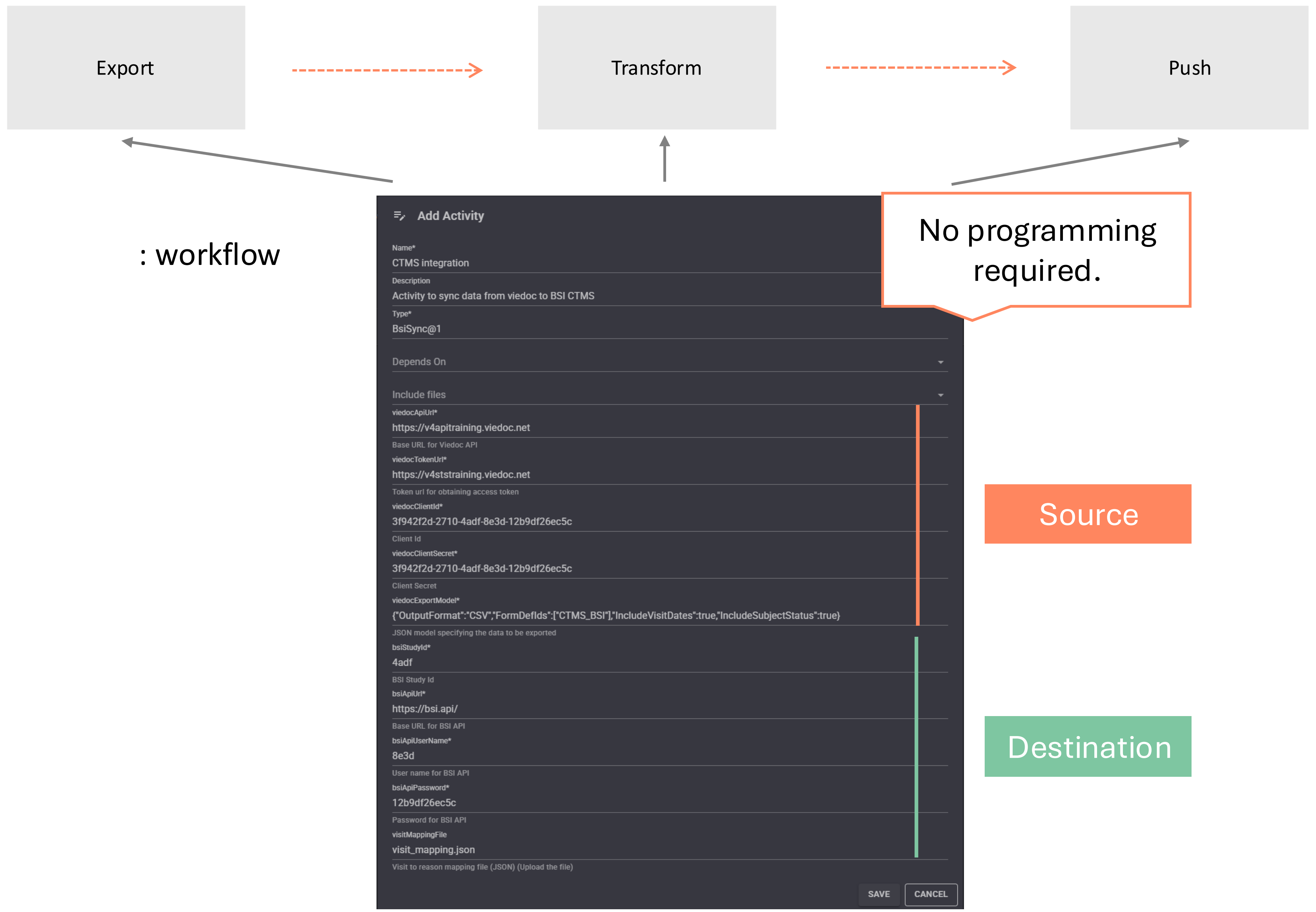
: custom code

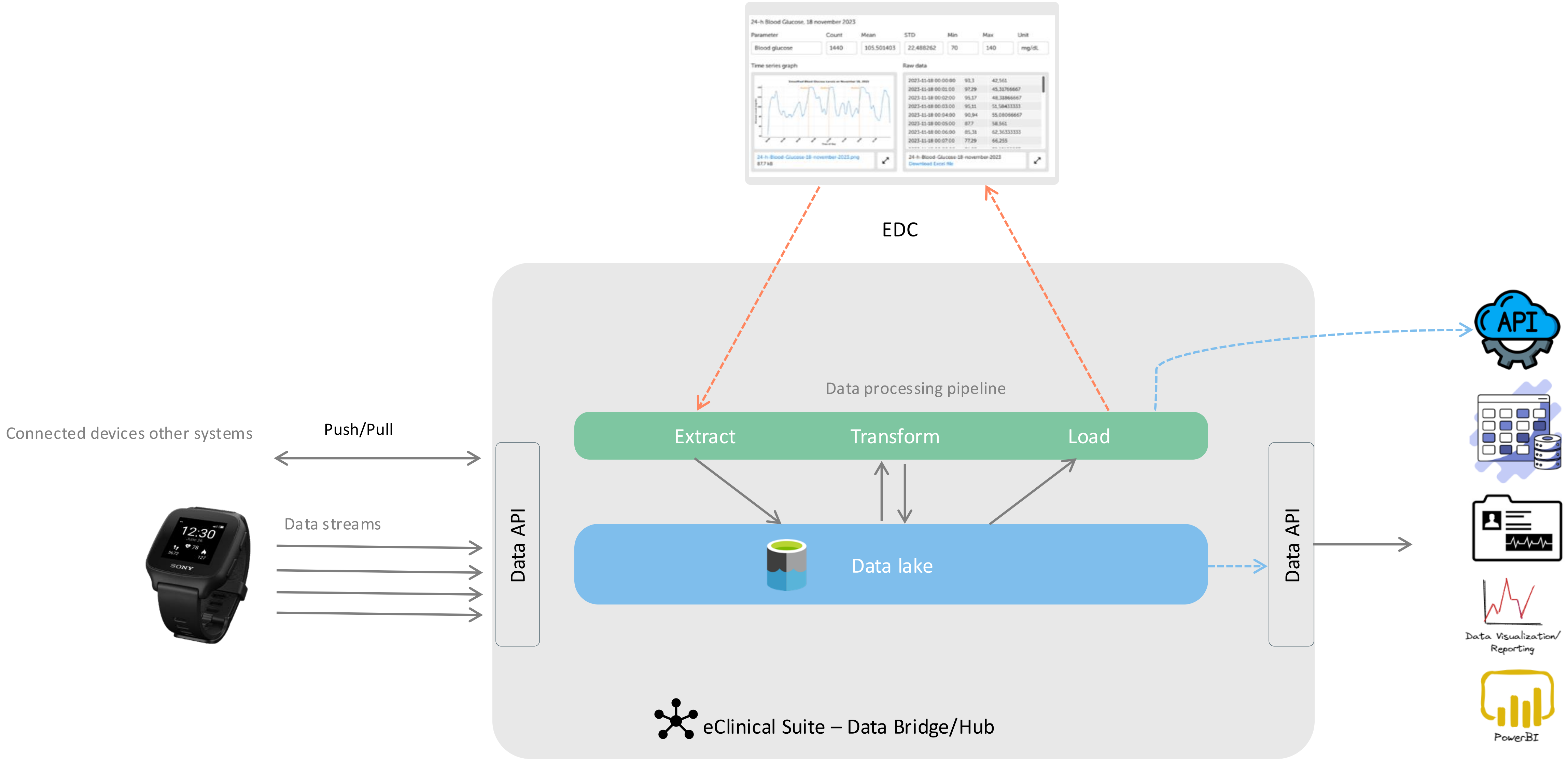



```python
def start_export(url, token, export_model):...

def get_export_status(url, token, export_id):...

def download_export(url, token, export_id, output_path, extract_zip, remove_prefix):...

def main(token_url, api_url, client_id, client_secret, export_model, output_path, extract_zip, remove_prefix):
    start_export_url = api_url + "/clinic/dataexport/start"
    check_status_url = api_url + "/clinic/dataexport/status"
    download_url = api_url + "/clinic/dataexport/download"

    # Get the access token
    token_manager = TokenManager()
    token = token_manager.get_token(token_url, client_id, client_secret)

    # Start the export process
    export_id = start_export(start_export_url, token, export_model)
    logging.info("Export ID: %s", export_id)

    # Check the export status until it's ready
    while get_export_status(check_status_url, token, export_id) != "Ready":
        token = token_manager.get_token(token_url, client_id, client_secret)

    # Download the export file
    download_export(download_url, token, export_id, output_path, extract_zip, remove_prefix)
```

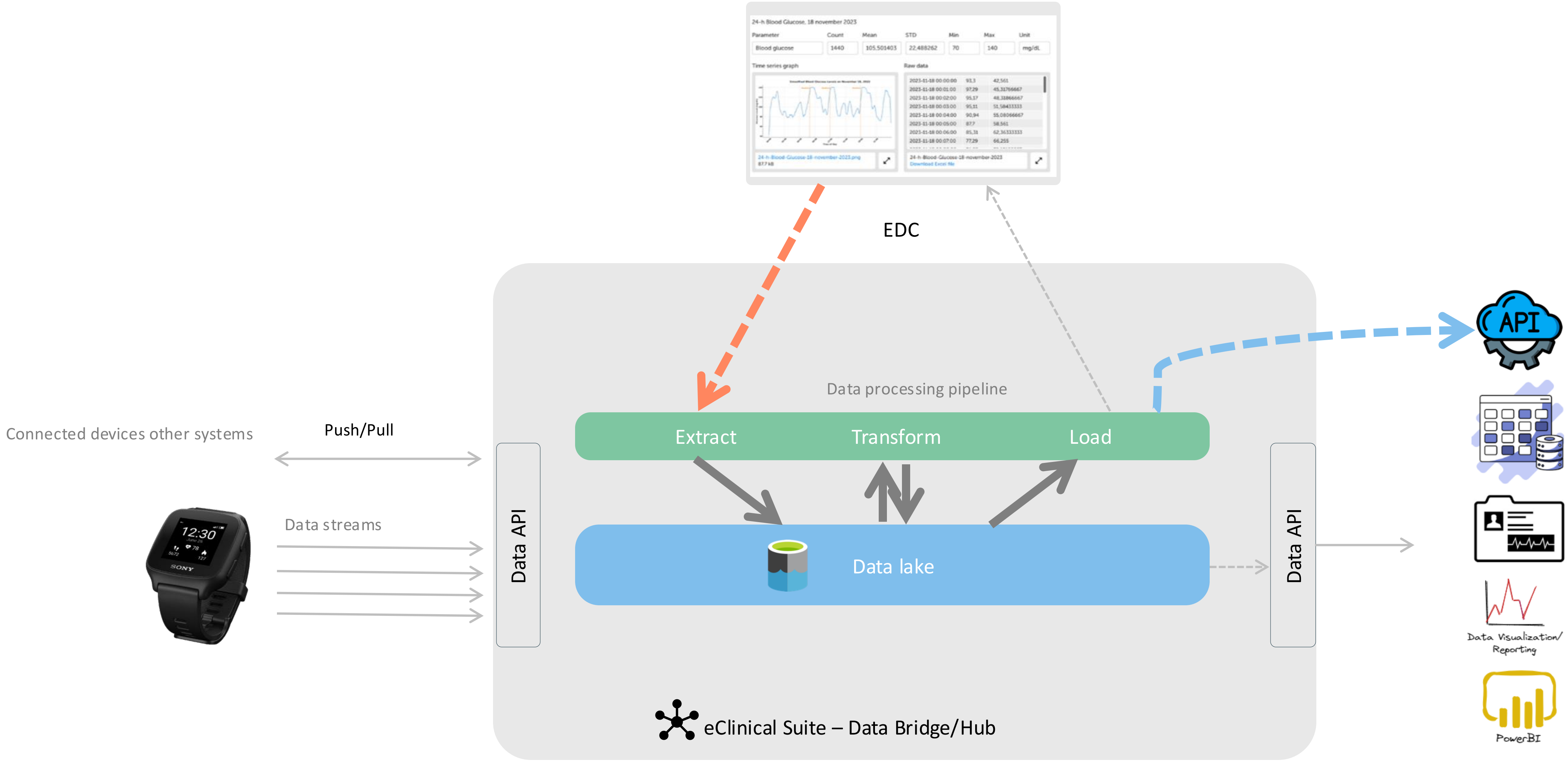


Export
Transform
Push
: task
Add Activity
Name*
Viedoc Export
Description
Exports data from Viedoc
Type*
ViedocExport@1
Depends On
Include files
apiUrl*
https://v4apitraining.viedoc.net
Base URL for Viedoc API
tokenUrl*
https://v4ststraining.viedoc.net
Token url for obtaining access token
clientId*
1aee8101-7df1-4f8a-9d34-7be8e1c6d516
Viedoc API Client Id
clientSecret*
1aee8101-7df1-4f8a-9d34-7be8e1c6d516
Viedoc API Client Secret
exportModel*
{"outputFormat": "csv"}
JSON model specifying the data to be exported (JSON string or file path)
outputPath*
/data/viedoc_export
Output path for exported data
extractZip
Y
Extract zip file after download
removePrefix
Y
Remove prefix from extracted files
SAVE
CANCEL
No programming required.
Task
API parameters
Path in data lake

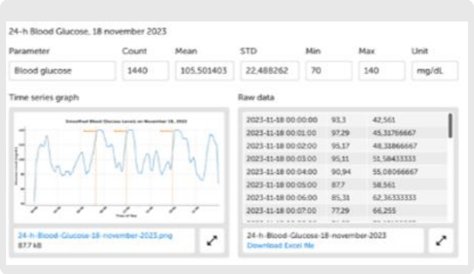

Export
Transform
Push
: workflow
Add Activity
Name*
CTMS integration
Description
Activity to sync data from viedoc to BSI CTMS
Type*
BsiSync@1
Depends On
Include files
viedocApiUrl*
https://v4apitraining.viedoc.net
Base URL for Viedoc API
viedocTokenUrl*
https://v4ststraining.viedoc.net
Token url for obtaining access token
viedocClientId*
3f942f2d-2710-4adf-8e3d-12b9df26ec5c
Client Id
viedocClientSecret*
3f942f2d-2710-4adf-8e3d-12b9df26ec5c
Client Secret
viedocExportModel*
{"OutputFormat":"CSV","FormDefIds":["CTMS_BSI"],"IncludeVisitDates":true,"IncludeSubjectStatus":true}
JSON model specifying the data to be exported
bsiStudyId*
4adf
BSI Study Id
bsiApiUrl*
https://bsi.api/
Base URL for BSI API
bsiApiUserName*
8e3d
User name for BSI API
bsiApiPassword*
12b9df26ec5c
Password for BSI API
visitMappingFile
visit_mapping.json
Visit to reason mapping file (JSON) (Upload the file)
SAVE
CANCEL
No programming required.
Source
Destination



EDC
Data processing pipeline
Extract
Transform
Load
Data lake
Data API
Data API
Connected devices other systems
Push/Pull
Data streams
eClinical Suite – Data Bridge/Hub
API
Data Visualization/ Reporting
PowerBI


EDC
Data processing pipeline
Connected devices other systems
Push/Pull
Data streams
Data API
Extract
Transform
Load
Data lake
Data API
API
Data Visualization/
Reporting
PowerBI
eClinical Suite – Data Bridge/Hub

Smart watch integration case study:
Enhancing patient monitoring with real-time data collection



Post: /api/data

1 file every 2min!

Ingest | Aggregate | Transform | Insights | Land

Raw data stored in:
Sony/2024-04-08/401/20240408181818123z.json Sony/2024-04-08/401/20240408181818123z_metadata.json

Create aggregate files for each device:

- Hourly data
- Daily data

Stored in:
Sony/2024-04-08/401/20240408.json Sony/2024-04-08/401/20240408_00.json Sony/2024-04-08/401/20240408_01.json

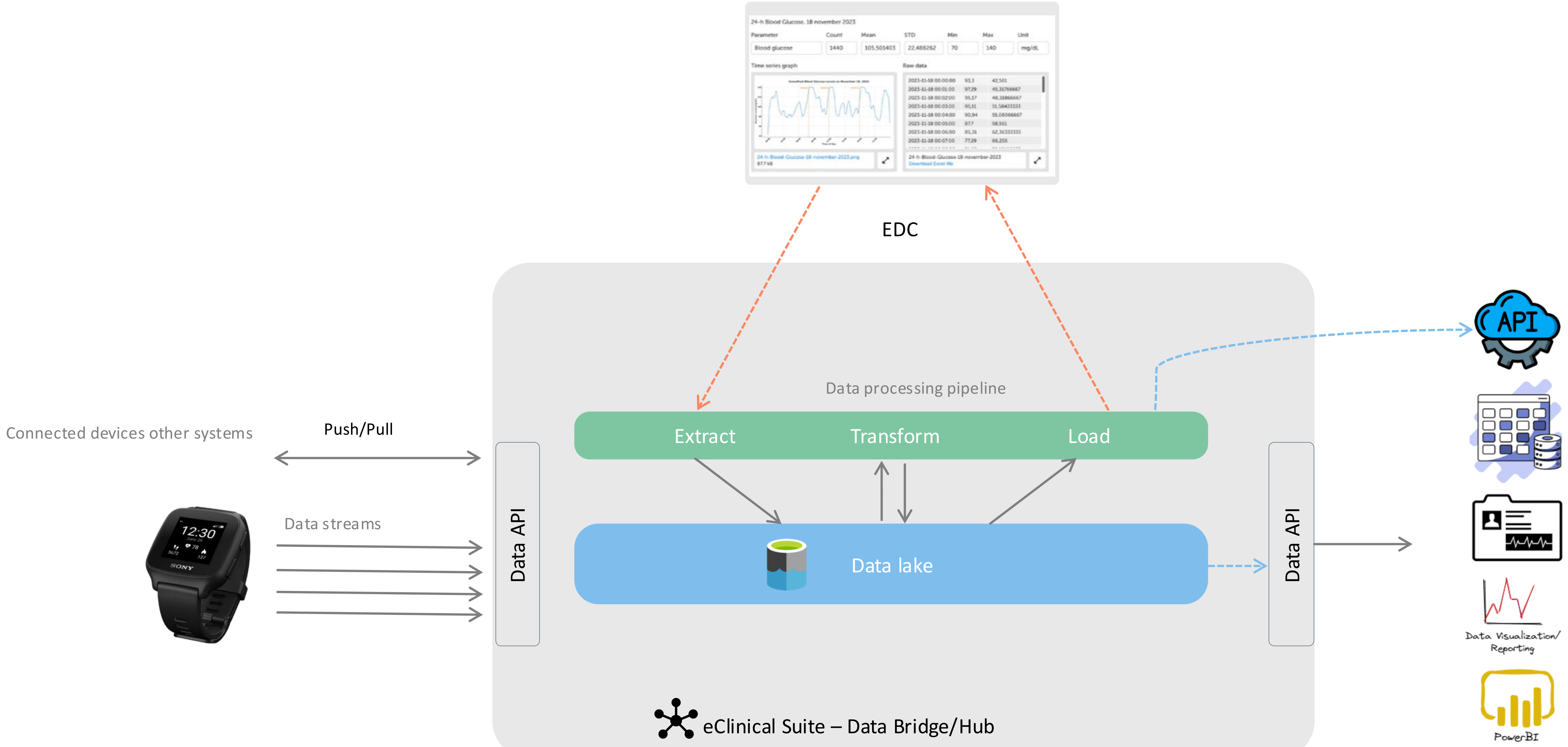
Hourly and daily files are transformed to a standardized format for easier analysis

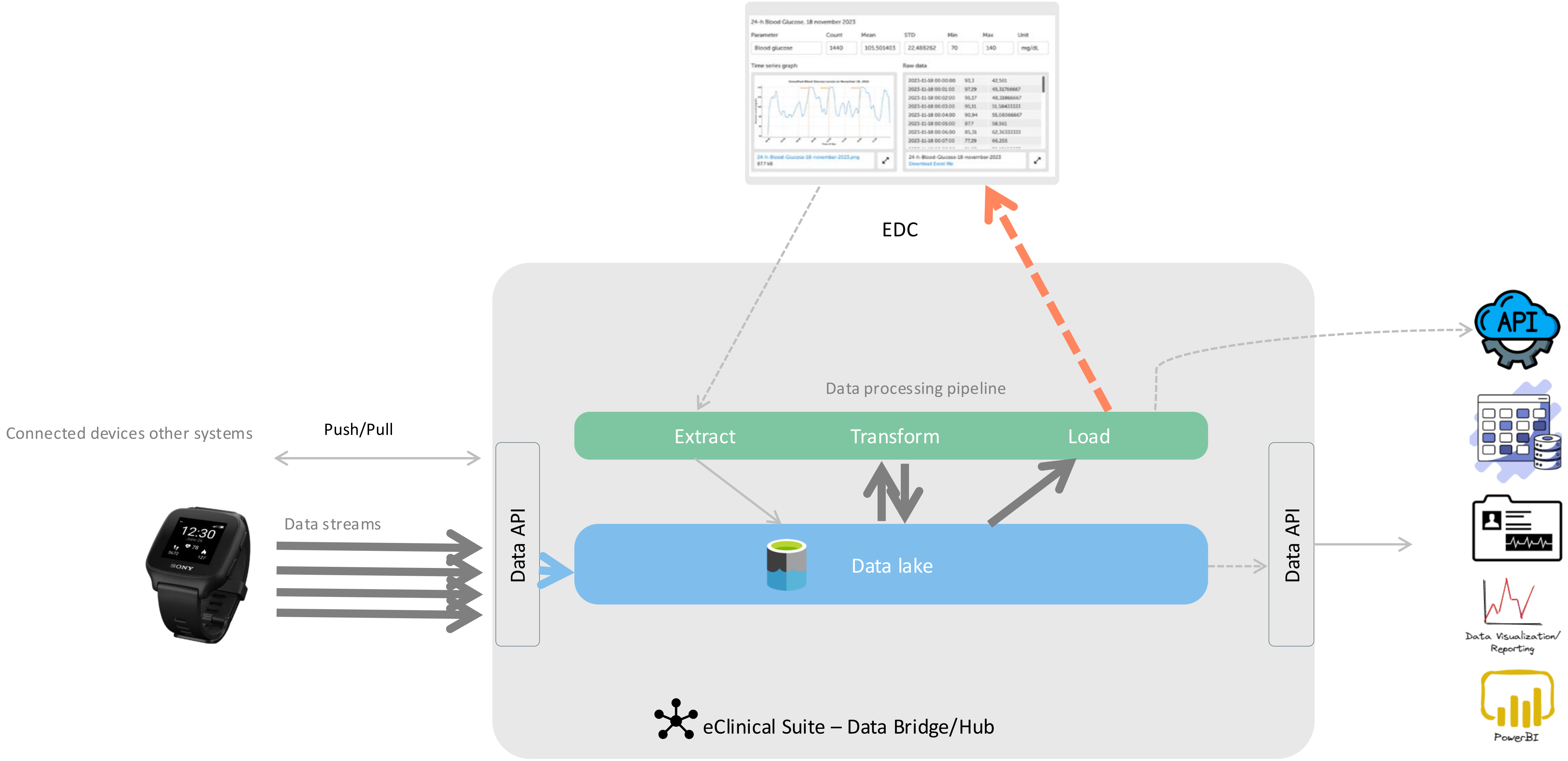
Stored in:
Sony/2024-04-08/401/20240408.csv Sony/2024-04-08/401/20240408_00.csv Sony/2024-04-08/401/20240408_01.csv

Summary statistics:
Sony/2024-04-08/401/20240408_summary.csv

Custom graphs:
Sony/2024-04-08/401/20240408_graph.png

Raw daily data summary, graphs are submitted to EDC


24-h Blood Glucose, 18 november 2023
Parameter
Count
Mean
STD
Min
Max
Unit
Blood glucose
1440
105,501403
22,488262
70
140
mg/dL
Time series graph
Raw data
EDC
Data processing pipeline
Connected devices other systems
Push/Pull
Data streams
Data API
Extract
Transform
Load
Data lake
Data API
API
Data Visualization/
Reporting
PowerBI
eClinical Suite – Data Bridge/Hub


EDC
Data processing pipeline
Connected devices other systems
Push/Pull
Extract
Transform
Load
Data streams
Data API
Data lake
Data API
API
Data Visualization/
Reporting
PowerBI
eClinical Suite – Data Bridge/Hub

R custom reports and LSAF integration

Export

CSV to RDS

Custom Reports

Load

Trigger a Viedoc clinical data export:
POST/api/data export/start

R script:

Convert CSV files to RDS (R data format)

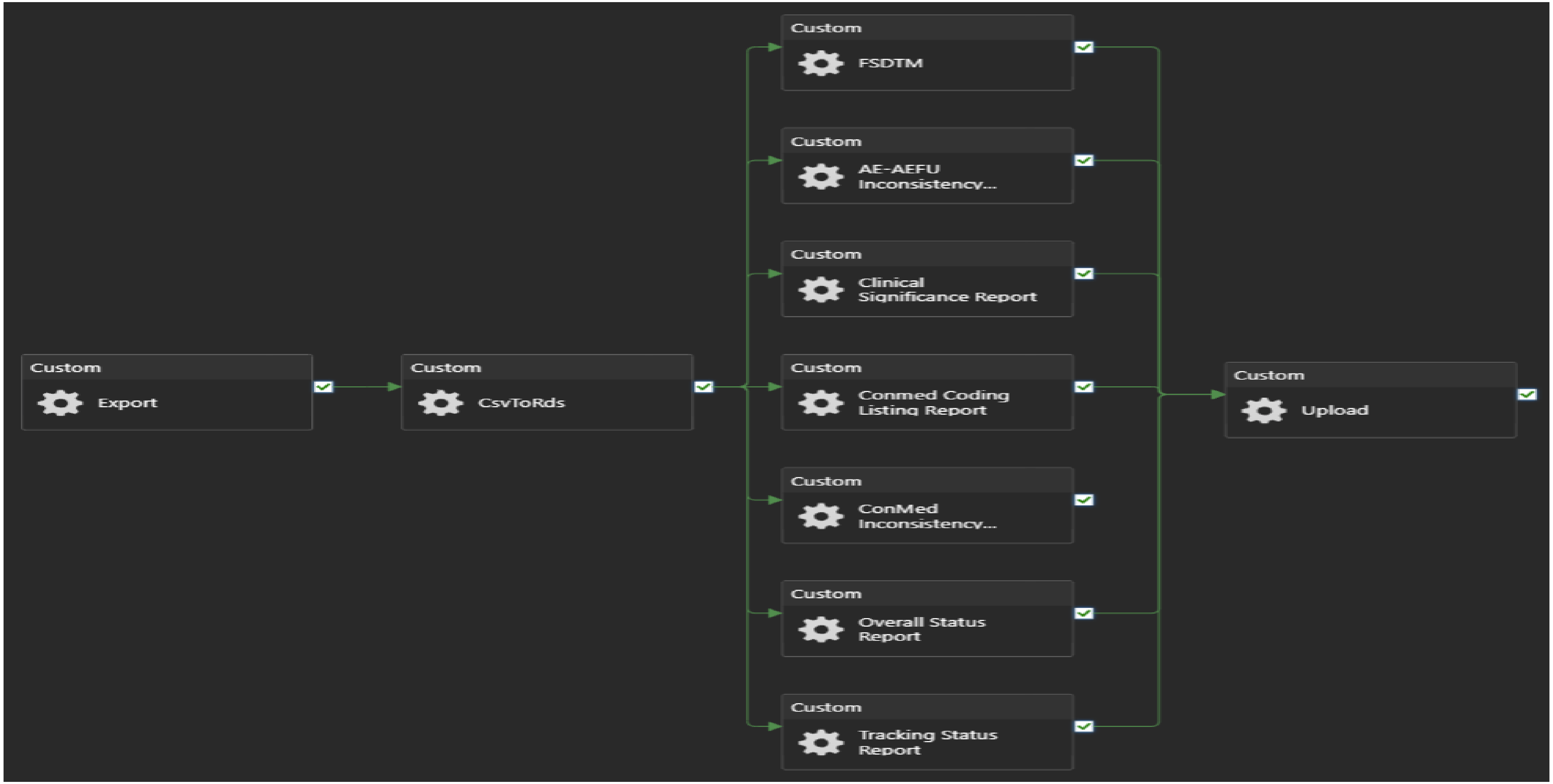
7 custom R reports

- Any custom report used in Viedoc works out of the box

A .zip file containing _all_ the reports is pushed to LSAF

R custom reports and LSAF integration

Export | CSV to RDS | Custom Reports | Load



Thank you

viedoc™