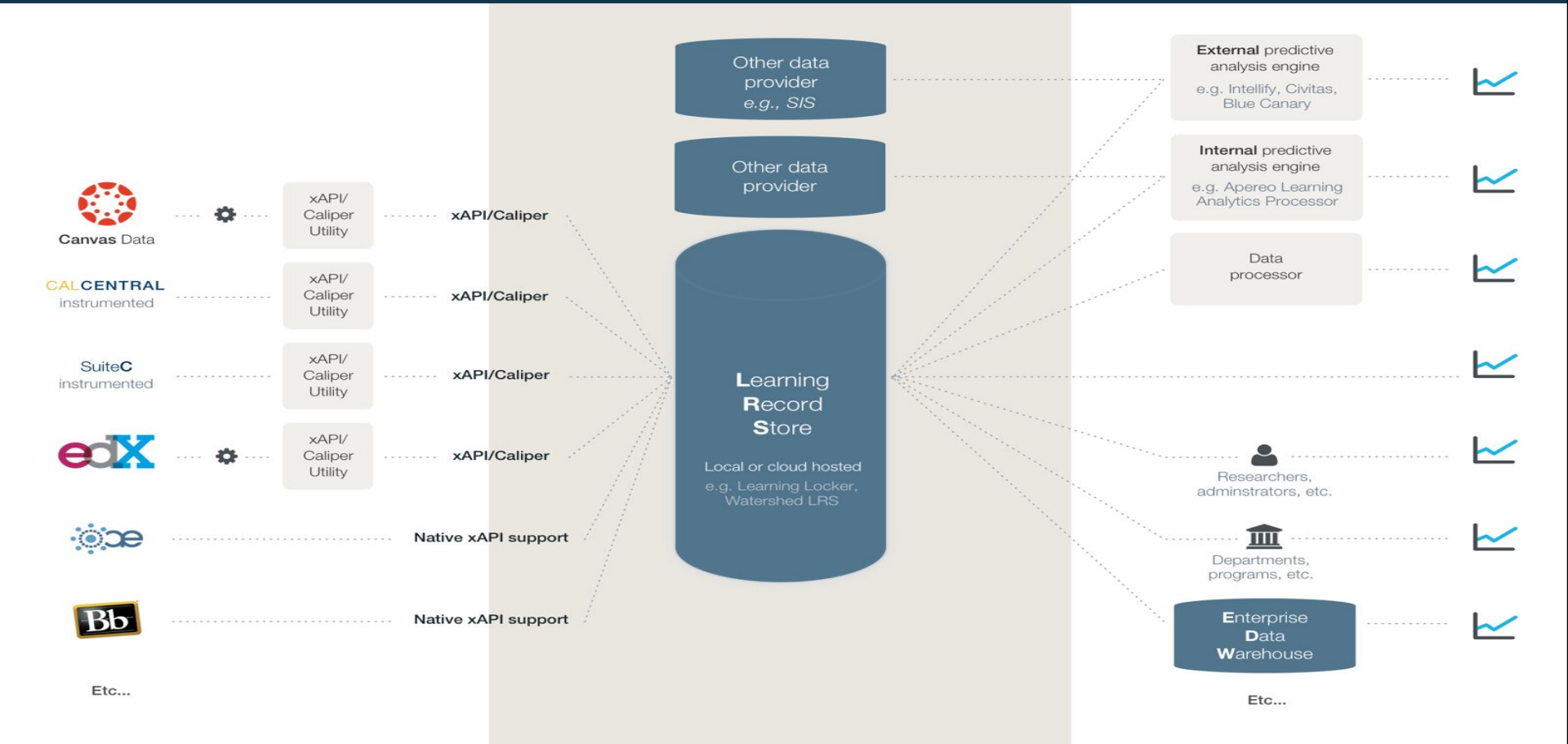
Building with Legos: UC Berkeley

Jenn Stringer, Oliver Heyer, Sandeep Jayaprakash

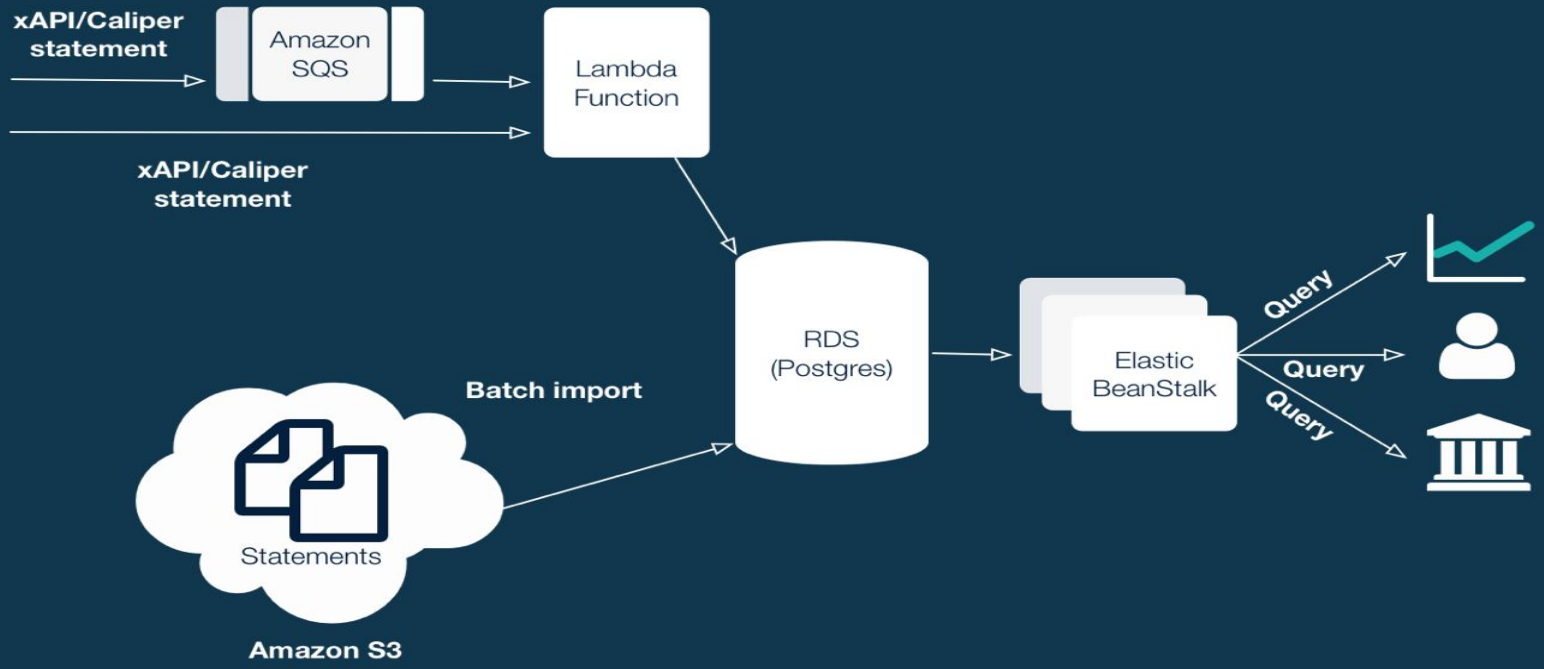
Ecosystem


Canvas Data
CALCENTRAL instrumented
SuiteC instrumented
edX
Etc...
xAPI/ Caliper Utility
xAPI/ Caliper Utility
xAPI/ Caliper Utility
xAPI/ Caliper Utility
xAPI/Caliper
xAPI/Caliper
xAPI/Caliper
xAPI/Caliper
Native xAPI support
Native xAPI support
Other data provider e.g., SIS
Other data provider
Learning Record Store
Local or cloud hosted
e.g. Learning Locker, Watershed LRS
External predictive analysis engine
e.g. Intellify, Civitas, Blue Canary
Internal predictive analysis engine
e.g. Apereo Learning Analytics Processor
Data processor
Researchers, adminstrators, etc.
Departments, programs, etc.
Enterprise Data Warehouse
Etc...

Learning Record Store

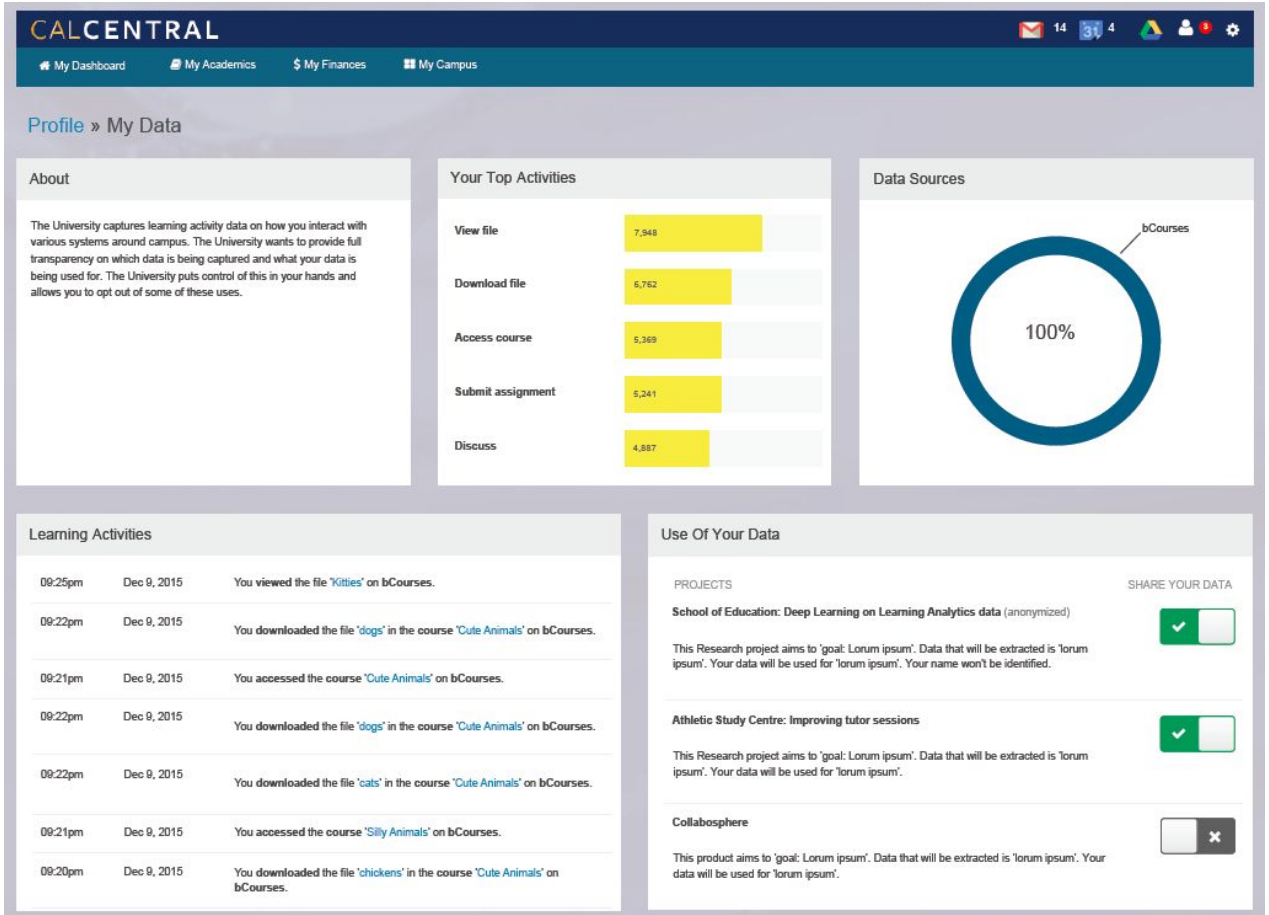
1. **Amazon Web Services**-based Learning Record Store
2. **Multi-tenant** LRS that can support multiple institutions at once
3. **Scalability and cost effective**
4. **Faster deployments** - Lower Dev/Ops overhead
5. **Lambda Architecture** which emcompasses both Batch and Real time processing capabilities

Learning Record Store

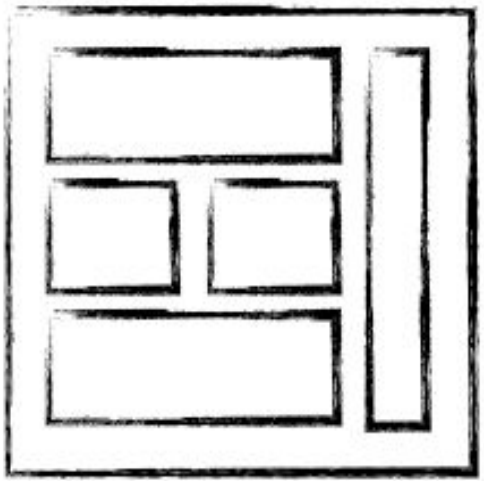


Quick Demo

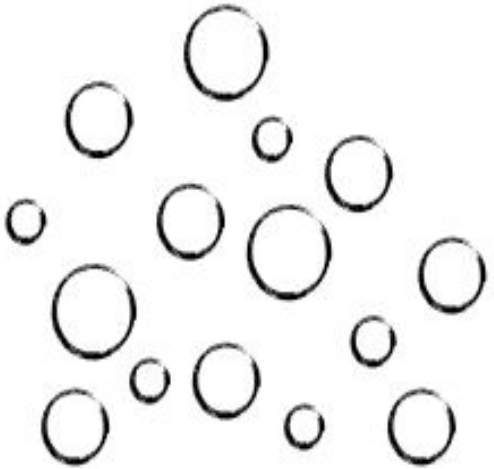
Student Agency and Privacy



Monolith Architectures vs Microservices



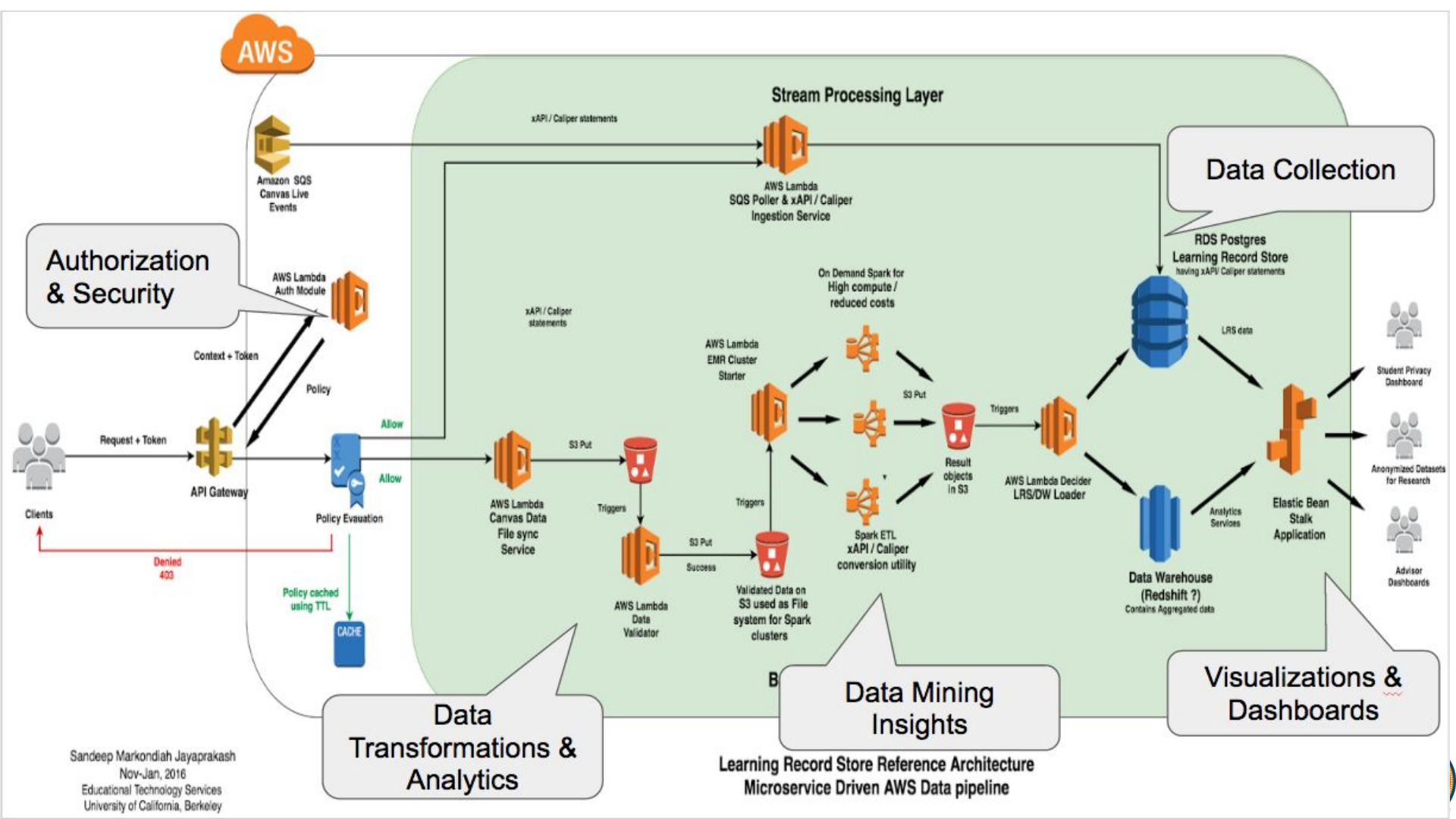
MONOLITHIC/LAYERED



MICRO SERVICES

Micro-services driven Architectures

1. Easy to understand code
2. Faster deployments & easy to scale
3. System resilience - Failure of a service doesn’t affect rest of the application
4. Ability to use different technology stacks (Best tool for the job)
5. Integrate new services easily since they are independent.
6. Embraced by big names like Amazon, Netflix, eBay etc


AWS
Stream Processing Layer
Data Collection
Authorization & Security
Data Transformations & Analytics
Data Mining Insights
Visualizations & Dashboards
xAPI / Caliper statements
Amazon SQS Canvas Live Events
AWS Lambda SQS Poller & xAPI / Caliper Ingestion Service
RDS Postgres Learning Record Store
having xAPI/ Caliper statements
AWS Lambda Auth Module
Context + Token
Policy
Request + Token
Clients
API Gateway
Policy Evaluation
Allow
Allow
Denied 403
Policy cached using TTL
CACHE
xAPI / Caliper statements
AWS Lambda Canvas Data File sync Service
S3 Put
Triggers
AWS Lambda Data Validator
S3 Put
Success
Validated Data on S3 used as File system for Spark clusters
Triggers
AWS Lambda EMR Cluster Starter
On Demand Spark for High compute / reduced costs
Spark ETL xAPI / Caliper conversion utility
S3 Put
Result objects in S3
Triggers
AWS Lambda Decider LRS/DW Loader
LRS data
Data Warehouse (Redshift ?)
Contains Aggregated data
Analytics Services
Elastic Bean Stalk Application
Student Privacy Dashboard
Anonymized Datasets for Research
Advisor Dashboards
Learning Record Store Reference Architecture
Microservice Driven AWS Data pipeline
Sandeep Markondiah Jayaprakash
Nov-Jan, 2016
Educational Technology Services
University of California, Berkeley

Next Steps

Enhancing Cloud Security for FERPA data in collaboration with AWS
https://d0.awsstatic.com/whitepapers/AWS_FERPA_Whitepaper.pdf

Ingest Canvas live Caliper feeds for an end-end pilot

Identify Learning Analytics use cases (Design phases)

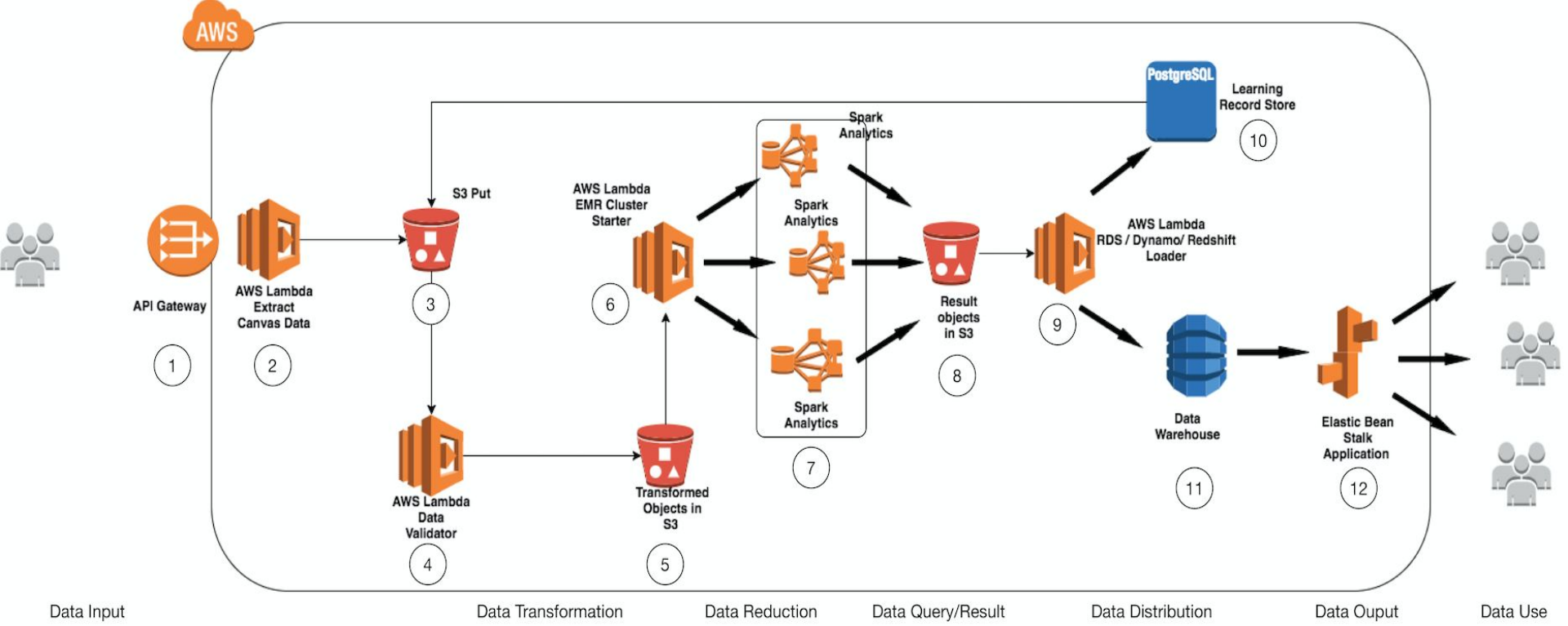
- NSF Impact Studio for SuiteC
- Athletics Study Center Advisor Dashboards

Collaboration with campus experts on policy, infrastructure and architecture.

Historical Replays of Canvas data

1. Currently working on converting Canvas activity data into open standards compliant formats xAPI / Caliper statements
2. Includes about 1 billion event transactions
3. Leveraging 2 Node JS applications
    a. Canvas Data Processor (migrating to Spark/Python)
    b. xAPI/Caliper utility

Appendix



Learning Analytics Cloud Architecture
Microservice Driven AWS Data pipeline workflow

1. Data input
2. ETL Canvas Data
3. Store transformed Canvas data in AWS S3
4. Validate data
5. Store transformed, validated data
6. Map/reduce data
7. Parallel data queries
8. Store results
9. Split data between LRS and DW
10. LRS data back to S3
11. DW data sent to apps/viz
12. Deliver apps to users