Digging through the data with MOOCczar

John Zornig

Andrew Dekker

- Everything about UQx, who we are, etc. :-)

- Very hard to engage with UQx learners
- Staff and students are disconnected, a bad signal to noise ratio
- Hard to tell how students are progressing
- Which questions are difficult?
- How (and whether) students are actually learning?

Understanding our users?

- An enormous amount of data
- Fantastic for long-term research, not fantastic for current courses
- Very rich data (**every** click), yet very shallow data (every **click**)
- All this data, what does it mean?
- How do we disseminate the data?
  - Internally?
  - To other researchers?
  - To the public?

---

edX research package

GitHub, Inc. [US] https://github.com/UQ-UQx/MOOCcz

eload this page

README.md

MOOCczar

A modular system for managing MOOC research data. These systems are in early testing stages, and contributions and feedback is welcome.

Data Analytics Platform

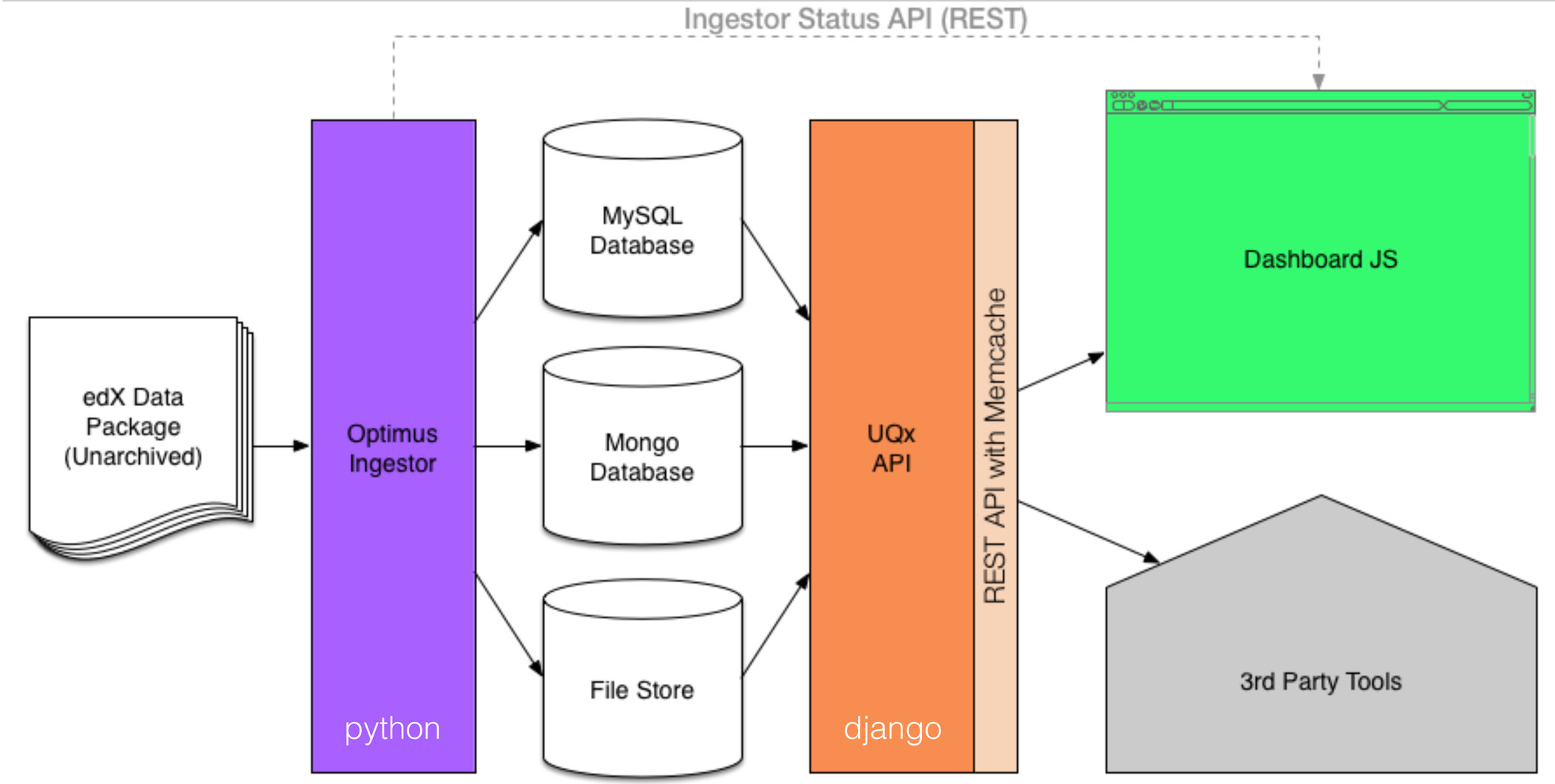
The Data Analytics Platform is a series of services which are used together to provide various access to the edX Research Data Package, and can be used as an API gateway to other data sources. The platform is segmented into three services, each of which work with eachother. Each service is extensible and can be modified independantly of the other services. Contributions are welcome to all of the services. They are best setup in order:

1. Optimus Ingestor - https://github.com/UQ-UQx/optimus_ingestor (ingests raw data into databases)
2. UQx API - https://github.com/UQ-UQx/uqx_api (provides REST endpoints to ingested data)

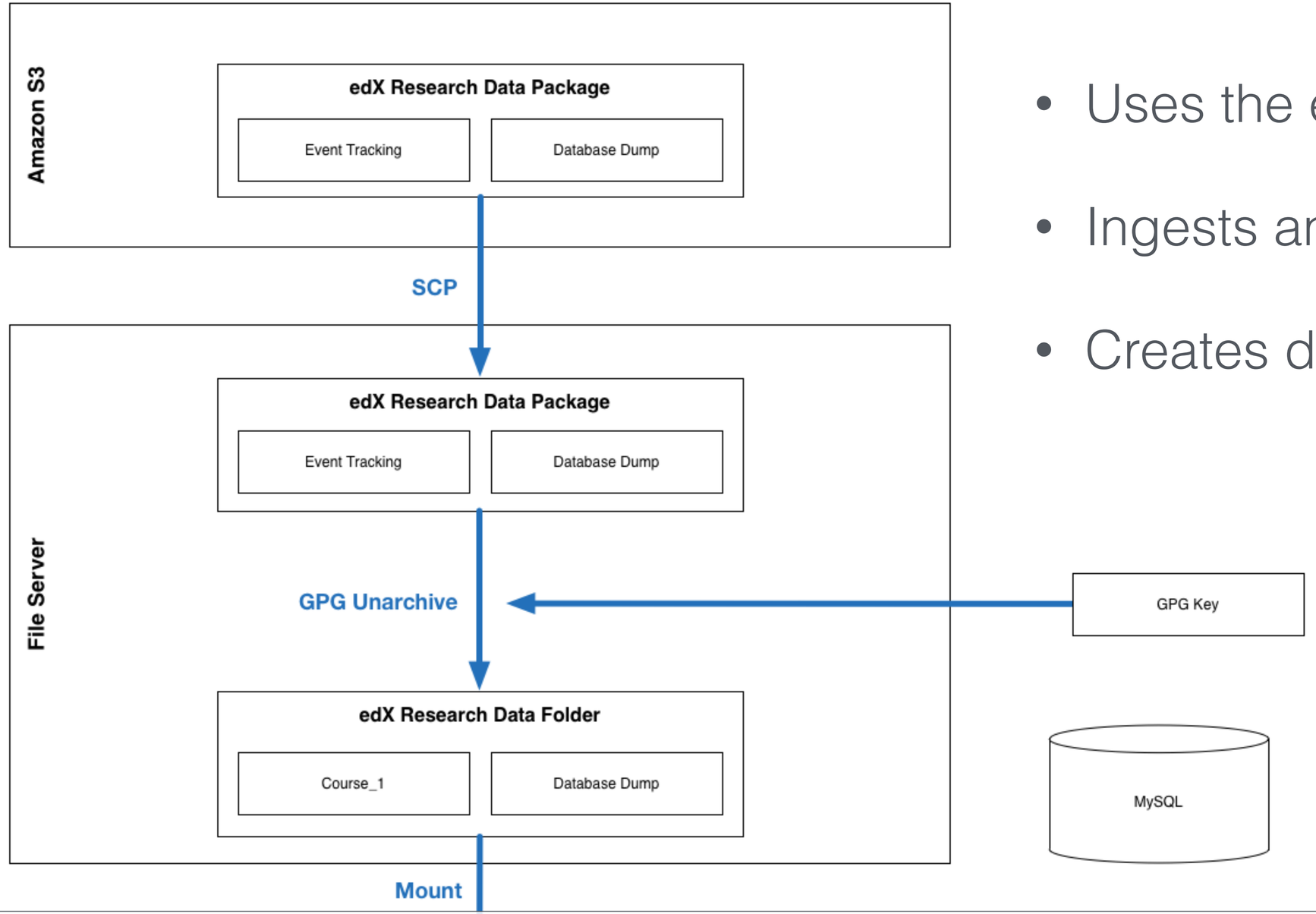
MOOCczar on Github

- Our first step into understanding and structuring our research data
- Architected to be extensible, usable and sharable
- Separation of concerns
  - Ingestor - Structuring the data
  - API - Sharing the data
  - Dashboard - Understanding the data

MOOCczar Platform

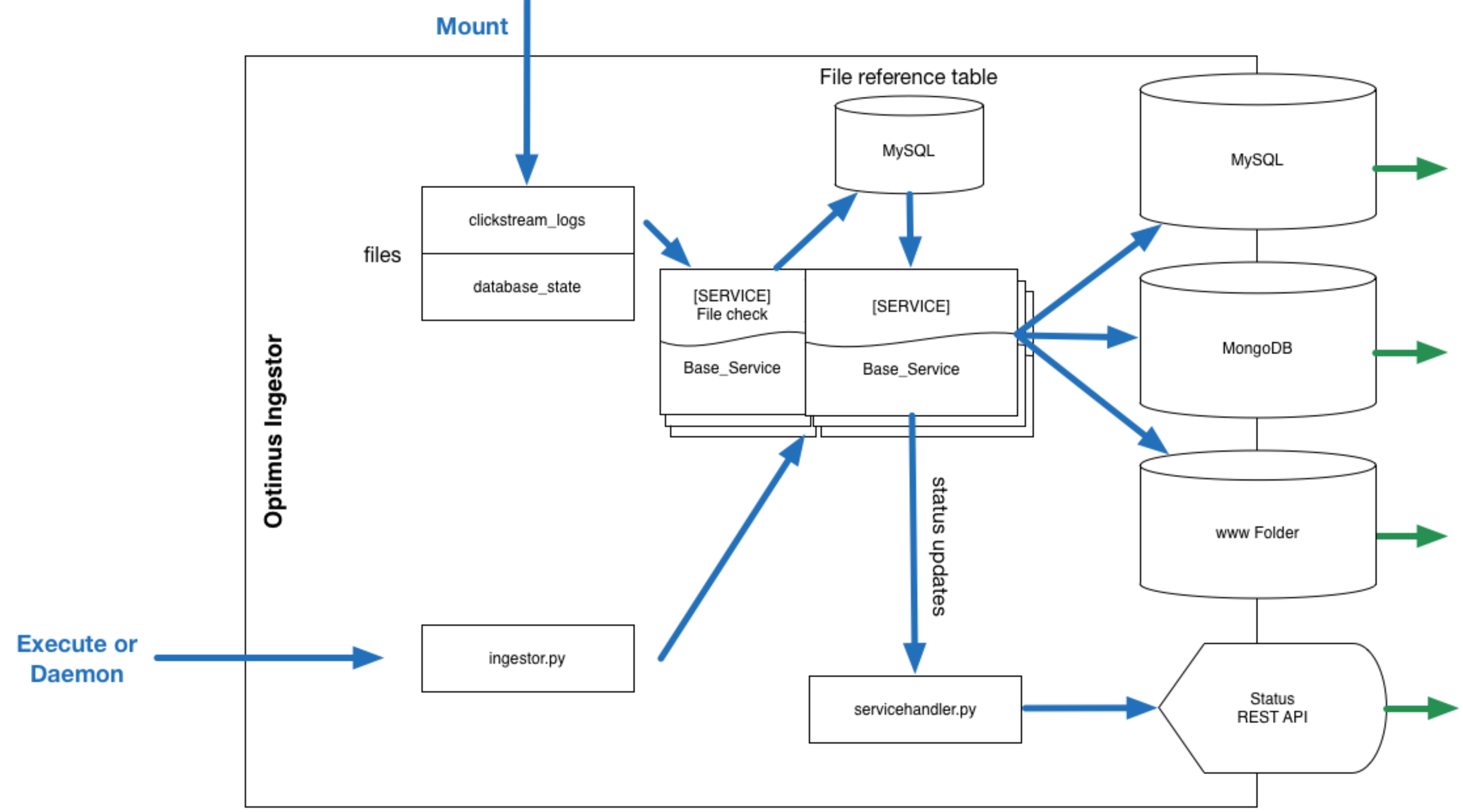


MOOCczar Platform



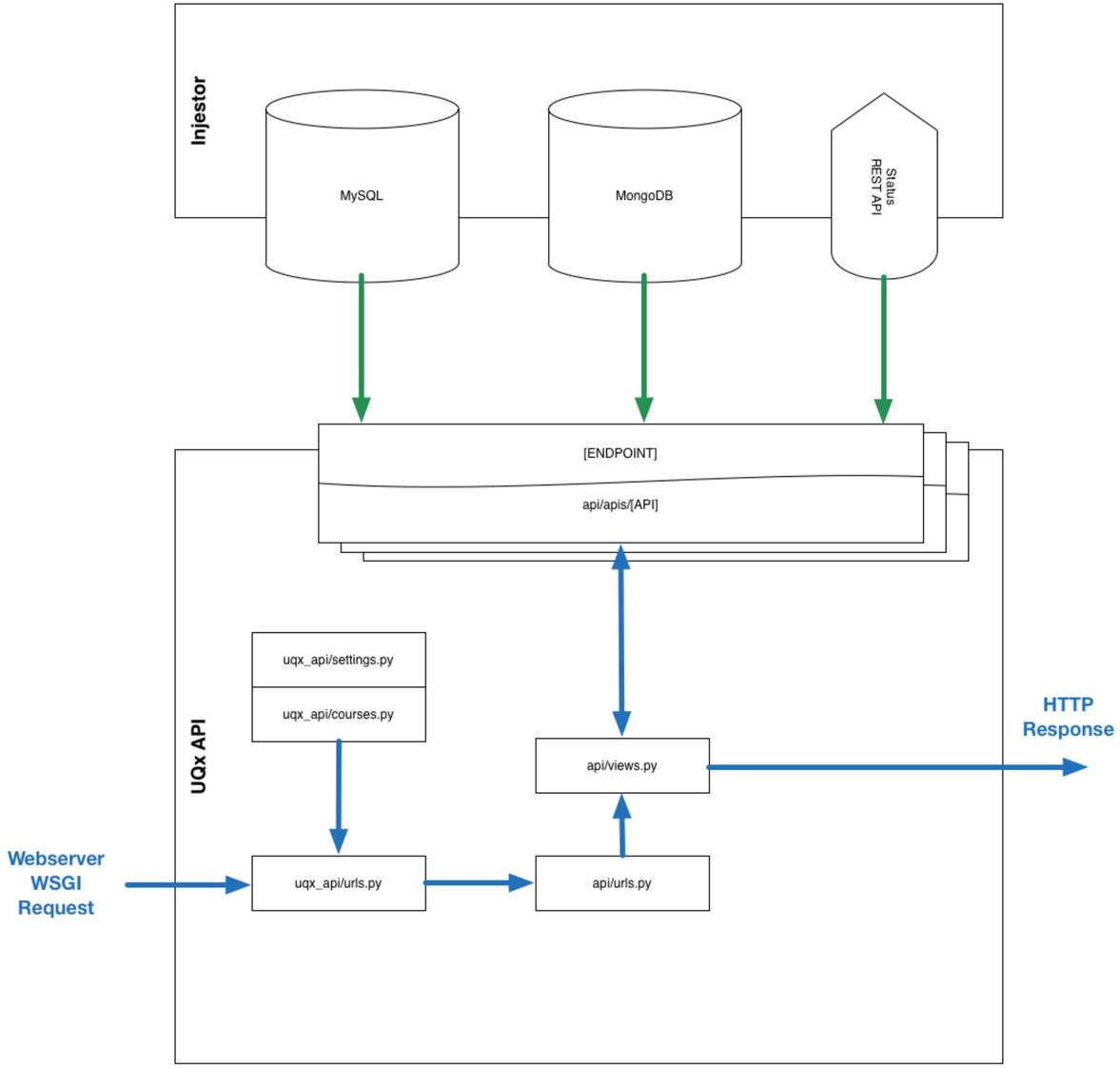
- Uses the edX data package
- Ingests and updates live DBs
- Creates derived data-sets

Ingestor


Mount
File reference table
MySQL
MySQL
clickstream_logs
files
database_state
[SERVICE]
File check
Base_Service
[SERVICE]
Base_Service
MongoDB
www Folder
Optimus Ingestor
status updates
Execute or
Daemon
ingestor.py
servicehandler.py
Status
REST API

Ingestor

- Standard REST architecture
- Uses django-rest-framework
- Caching layer included (memcache)
- Built around course JSON file
- Modular authentication (inc. LDAP)



API

- Angular and d3 based dashboard
- Fully client side stack, driven by API
- Modular d3 visualisations
- Course + Visualisation navigation
- Support for embeddable reports
- Provides interface to API caching mechanism

Dashboard

THE UNIVERSITY OF QUEENSLAND AUSTRALIA UQX DASHBOARD

Welcome edx (Logout)

Dashboard | Reports | Data Status | UQx

All Courses | Enrolments | Enrolment Map | Demographics | Structure | Active Students | Certificates | Discussions | Discussion Map

All Courses

Tropic 101x *2014*

Think 101x *2014*

World 101x *2014*

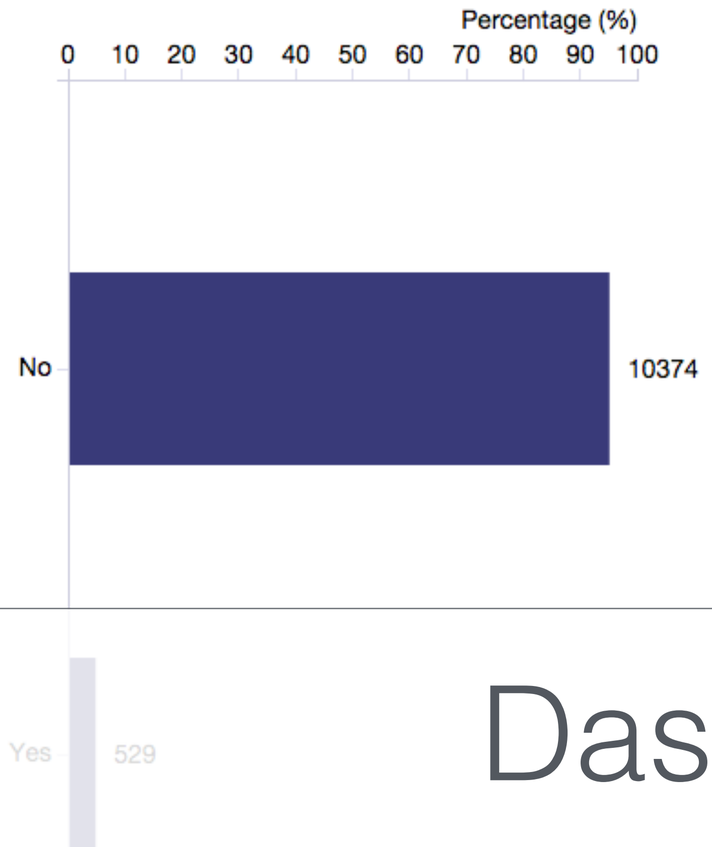
Write 101x *2014*

Hypers 301x *2014*

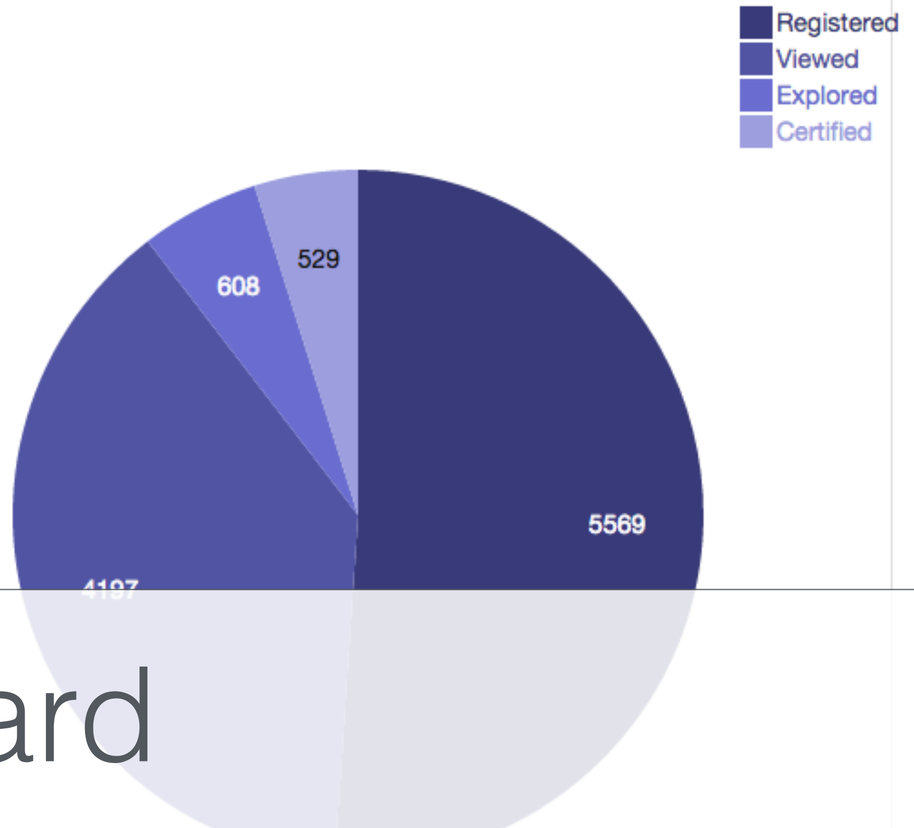
Certificates for Tropic 101x

CSV

Certificates Awarded



Progress



Dashboard

All Courses | Enrolments | Enrolment Map | Demographics | Structure | Active Students | Course Events | Certificates | Discussions | Discussion Map

All Courses

Tropic 101x *2014*

Think 101x *2014*

World 101x *2014*

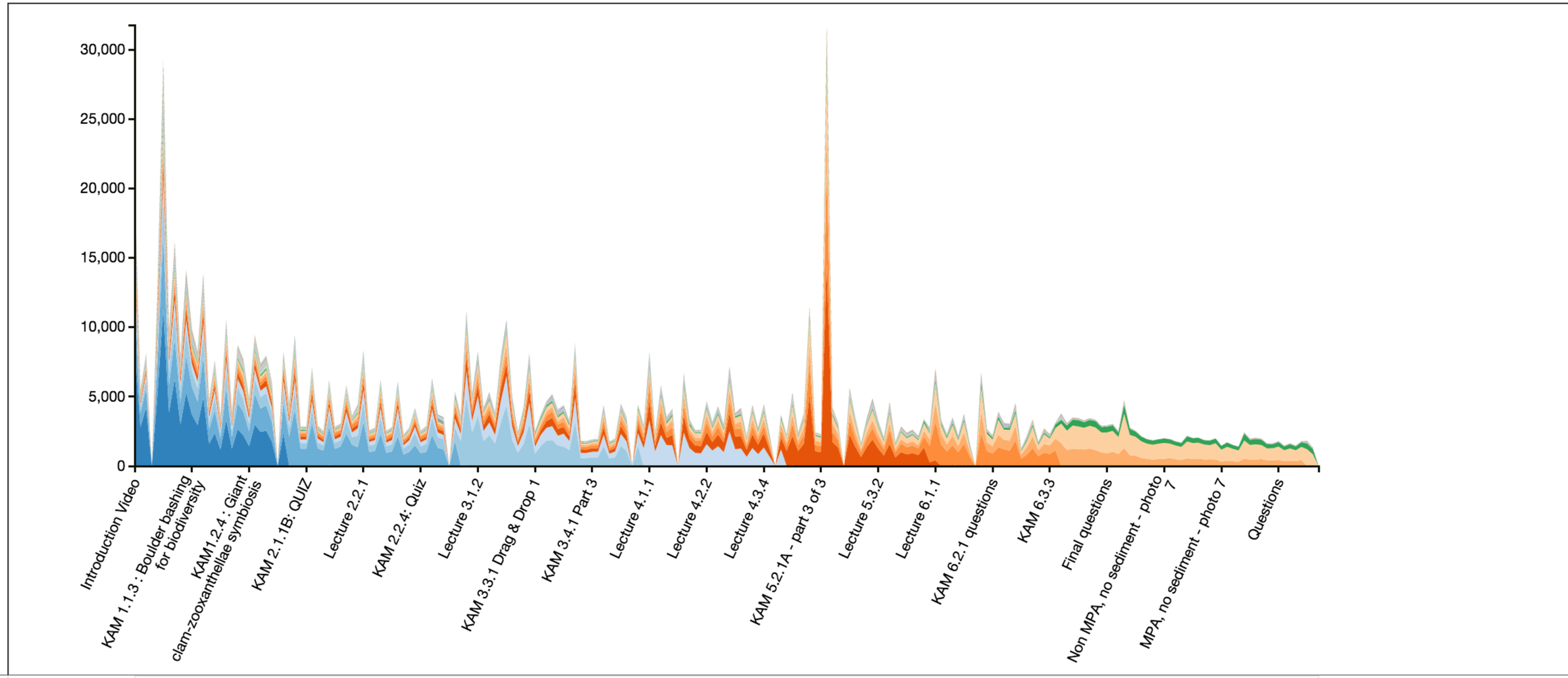
Write 101x *2014*

Hypers 301x *2014*

Bioimg 101x *2014*

Crime 101x *2014*

Course Events for Tropic 101x



Dashboard

363772

total enrolments

2004

average enrolments per day over last 30 days

306917

unique students (ignoring registrations for those who enrol in multiple UQx Courses)

Enrolments For Courses

think 101x 1T2014

128372

total enrolments

551

enrolments per day

5034

hypers 301x 1T2014

8705

total enrolments

52

enrolments per day

242

bioimg 101x 1T2014

19256

total enrolments

111

enrolments per day

1226

tropic 101x 1T2014

10903

total enrolments

65

enrolments per day

529

Dashboard Reports

Take-aways

- We hope this platform can be helpful to not only us
- Separation of concerns, you can focus on the structure of the data, the sharing of the data, or the understanding of the data
- Constantly evolving with the edX platform
- Digging through the data together

Thank you for your time

John Zornig

Andrew Dekker

http://uqx.uq.edu.au || http://edx.org || http://github.com/UQ-UQx || http://github.com/edx