

Kaka: Towards a Distributed Genotype Information System

Helge Dzierzon Samantha Baldwin

October 22, 2015

Kaka: Towards a Distributed Genotype Information System

Helge Dzierzon, Samantha Baldwin, John McCallum

TOC

► Why?

TOC

- ▶ Why?
- ▶ How?

TOC

- ▶ Why?
- ▶ How?
- ▶ Status

TOC

- ▶ Why?
- ▶ How?
- ▶ Status
- ▶ Conclusions

Why?

- ▶ PFR distributed

Why?

- ▶ PFR distributed
- ▶ Genotype information created

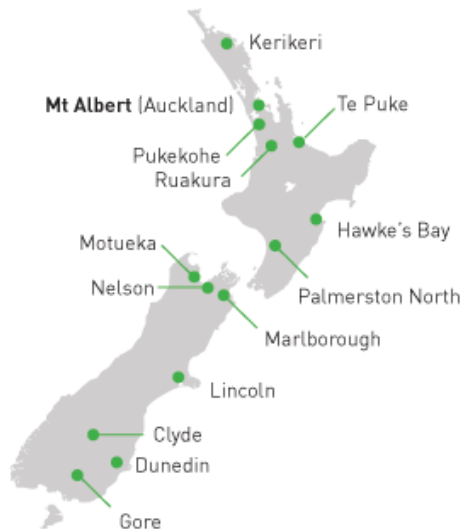
Why?

- ▶ PFR distributed
- ▶ Genotype information created
- ▶ No central place to o to

Why?

- ▶ PFR distributed
- ▶ Genotype information created
- ▶ No central place to o to
- ▶ Nobody knows what others do

PFR Sites Across New Zealand



- ▶ Genotype information is produced on a lot of these sites

- ▶ Genotype information is produced on a lot of these sites
- ▶ Analyst usually somewhere else and needs access to the data

- ▶ Genotype information is produced on a lot of these sites
- ▶ Analyst usually somewhere else and needs access to the data
- ▶ Data currently spread on K-Drives

Existing Databases

- ▶ EBrida (provenance)

Existing Databases

- ▶ EBrida (provenance)
- ▶ KEA (sample tracking)

Existing Databases

- ▶ EBrida (provenance)
- ▶ KEA (sample tracking)
- ▶ Ensembl

DBs in development

- ▶ Kaka

DBs in development

- ▶ Kaka
- ▶ Kakapo

How?

- ▶ Python Django

How?

- ▶ Python Django
- ▶ PGSQL and JSON

How?

- ▶ Python Django
- ▶ PGSQL and JSON
- ▶ Docker and continuous integration

Status

- ▶ DB has been developed as a proof of concept

Status

- ▶ DB has been developed as a proof of concept
- ▶ Some example data have been loaded

Status

- ▶ DB has been developed as a proof of concept
- ▶ Some example data have been loaded
- ▶ RStudio, notebooks databases etc. have been attached and can access api

Status

- ▶ DB has been developed as a proof of concept
- ▶ Some example data have been loaded
- ▶ RStudio, notebooks databases etc. have been attached and can access api
- ▶ API has been developed

Conclusions

- ▶ Concept is simple but a bit nerdy

Conclusions

- ▶ Concept is simple but a bit nerdy
- ▶ It is a great tool already but needs a nerd

Conclusions

- ▶ Concept is simple but a bit nerdy
- ▶ It is a great tool already but needs a nerd
- ▶ Blubb

Thank you! Questions?

Image