

A web server for assembling, annotating and visualizing bacterial genomes

Jason R. Grant¹, Eric Enns², Eric Marinier², Arnab Saha-Mandal³, Chih-Yu Chen^{2,3}, Morag Graham^{2,3}, Gary Van Domselaar^{2,3}, Paul Stothard¹

¹University of Alberta, Edmonton, Canada

²National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Canada

³University of Manitoba, Winnipeg, Canada

What is Proksee?

- Proksee is an analytical toolkit and high-performance genome browser for rapid and in-depth characterization of bacterial genomes and plasmids.
- Proksee makes it easy to annotate a bacterial genome and create high-quality, publication-ready genome maps
- Proksee is designed to be user friendly, responsive, easy to maintain and extensible.
- Proksee is available here: proksee.ca

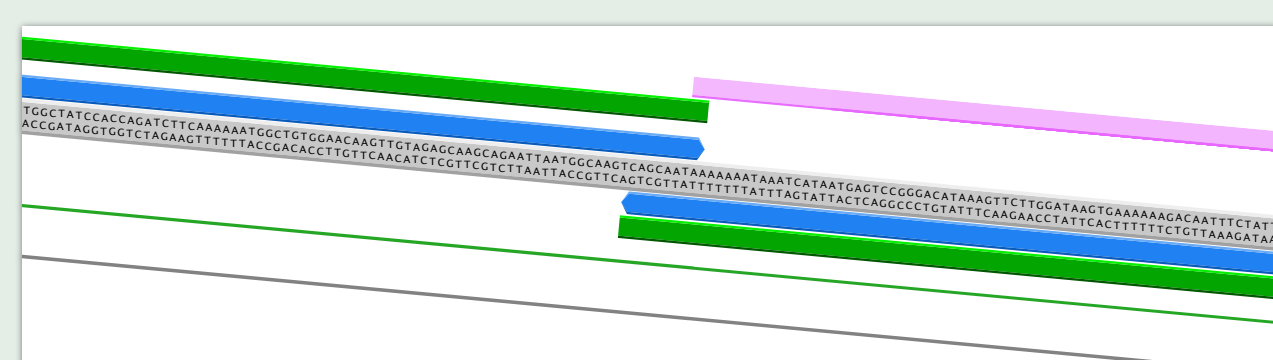
Start a Project

Proksee accepts 3 types of input for starting a project:

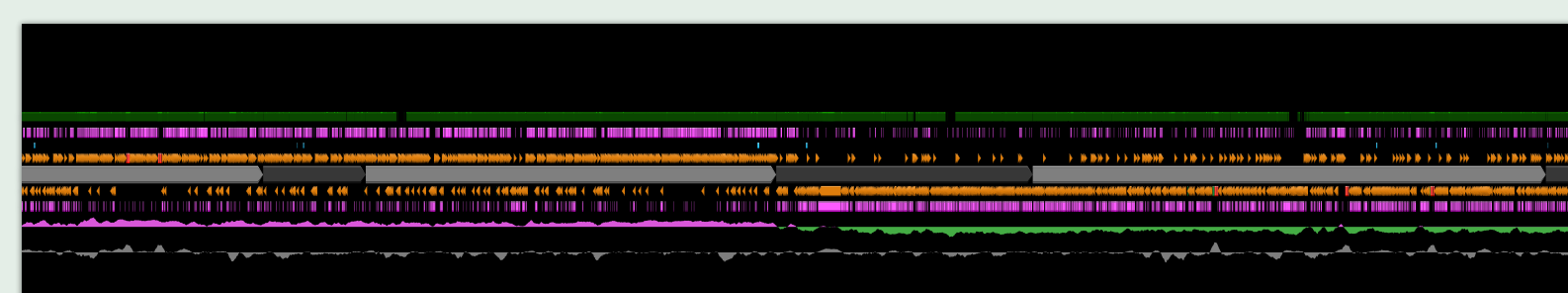
- Genome files**
 - Genbank, EMBL, or FASTA
 - NCBI accession can be provided instead of file
 - Sequences can be up to 10 Mbp
 - Annotations are added to map
- Sequence reads** COMING SOON
 - FASTQ files (Forward and Reverse)
 - Files up to 1 GB each
- JSON**
 - Map files created by Proksee or CGView.js

View

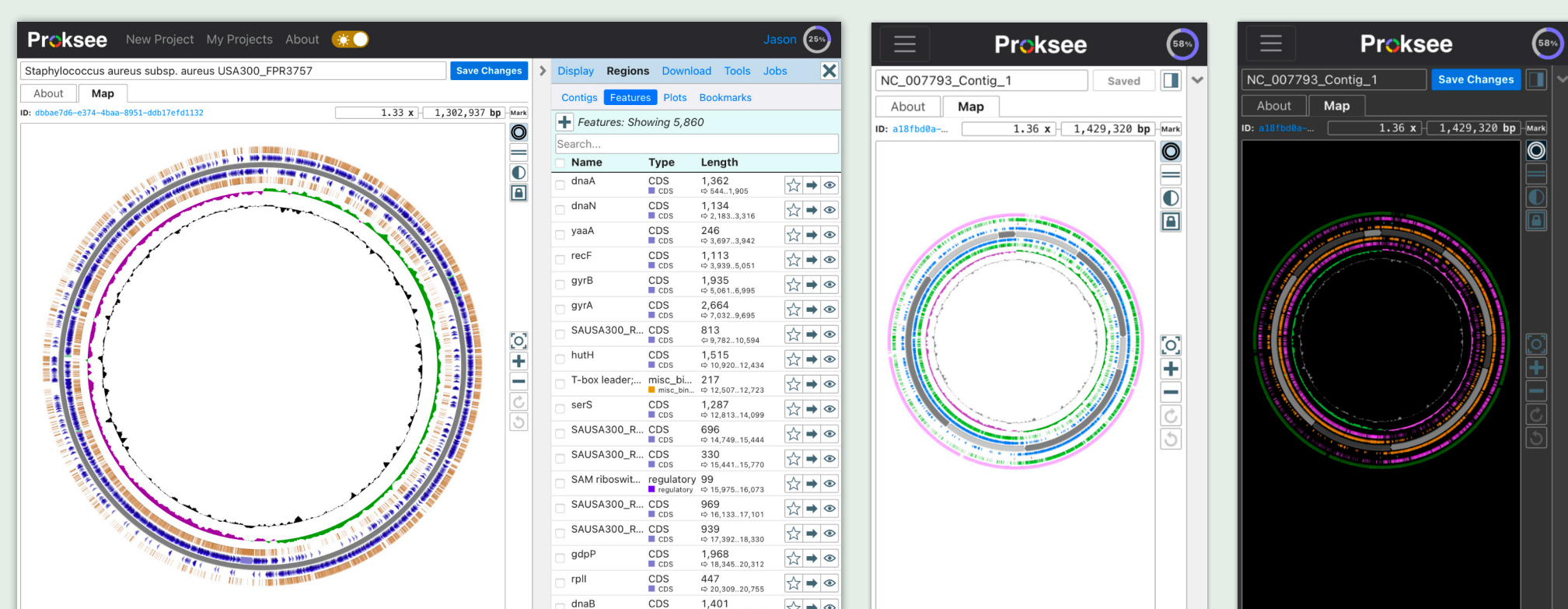
- Interactive maps are viewed with the CGView.js genome browser (<https://js.cgview.ca>) allowing navigation and zooming down to the sequence level



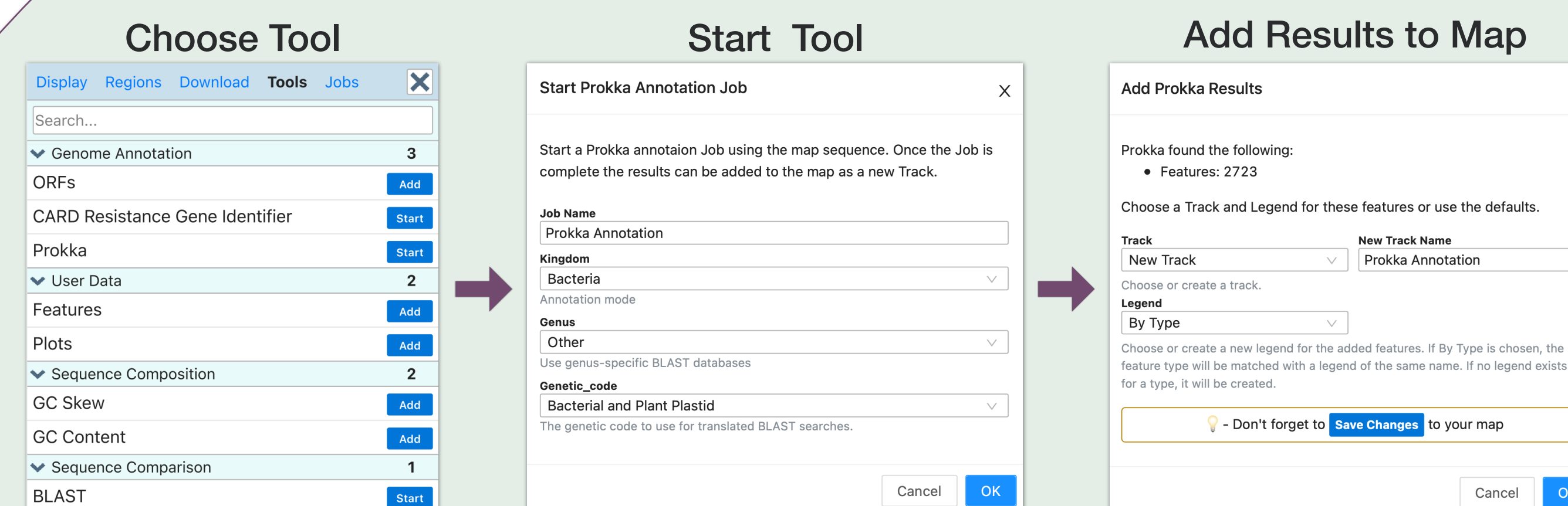
- Maps can viewed in a circular or linear format



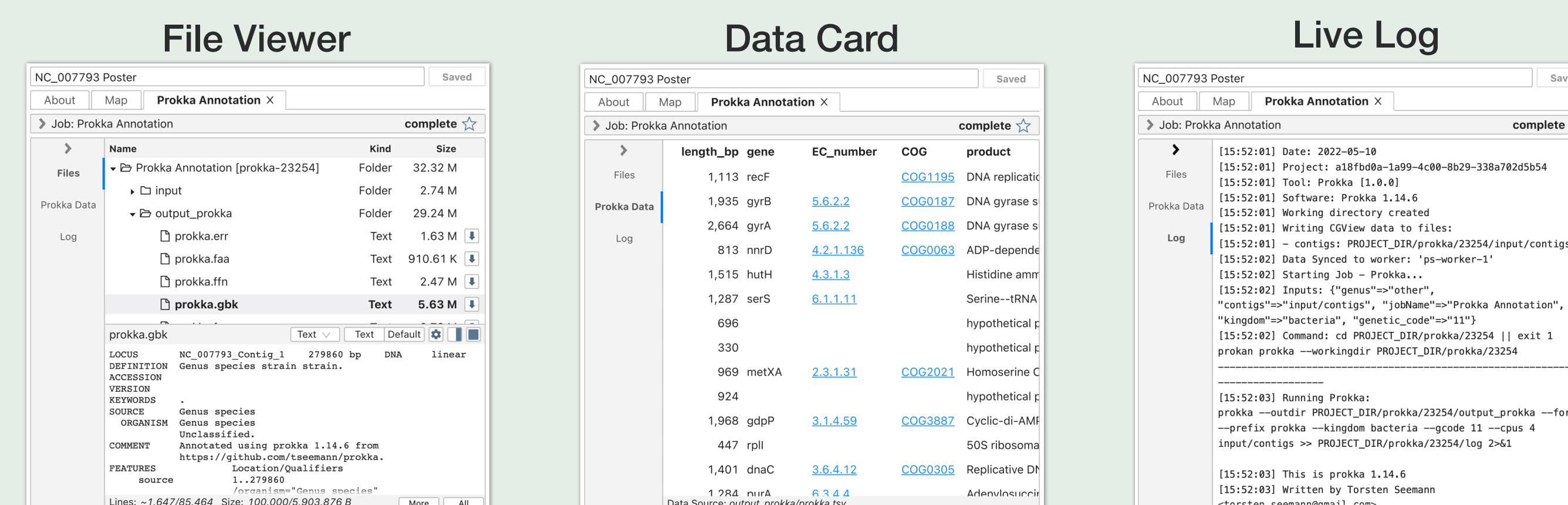
- Proksee works on computers, tablets and smart phones and has a light and dark mode



Annotate



Annotation Tools. A growing collection of integrated tools is provided to the user for performing various analyses and the results are easily added to the map for comparison to other features.



Annotation Results. The results can be examined or downloaded with the integrated file viewer. Data cards are tool specific and use specialized viewers for tabular data, HTML reports and images. A live log shows the real time progress of jobs running on the server.

Poster Map

The *Staphylococcus aureus* USA300 (MRSA) genome (NC_007793) was submitted to Proksee. DNA BLAST was performed against the *S. aureus* reference genome (NC_007795). The CARD Resistance Gene Identifier, Prokka, ORF, GC Skew, and GC Content tools were run with default settings.

BLAST
ORF
CARD
Prokka
GC Skew +
GC Skew -
GC Content

Edit

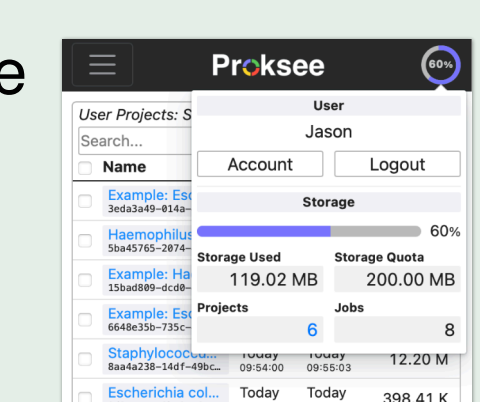
- Numerous options are available for customizing the map, tracks, legend, captions, labels, ruler, dividers and more
- Colours are easily changed with a custom colour picker
- Fonts for any map text can be altered
- Features can have different decorations
- Features can also be drawn with shading

Download

- PNG Images**
 - Bitmap-based images up to 8,000x8,000 px
- SVG Images**
 - Vector-based images
- JSON**
 - Map files that can be saved and reloaded later
- Sequence Data**
 - Whole genome or any selection of features
 - Features can be downloaded as DNA or AA
- Annotation Results**

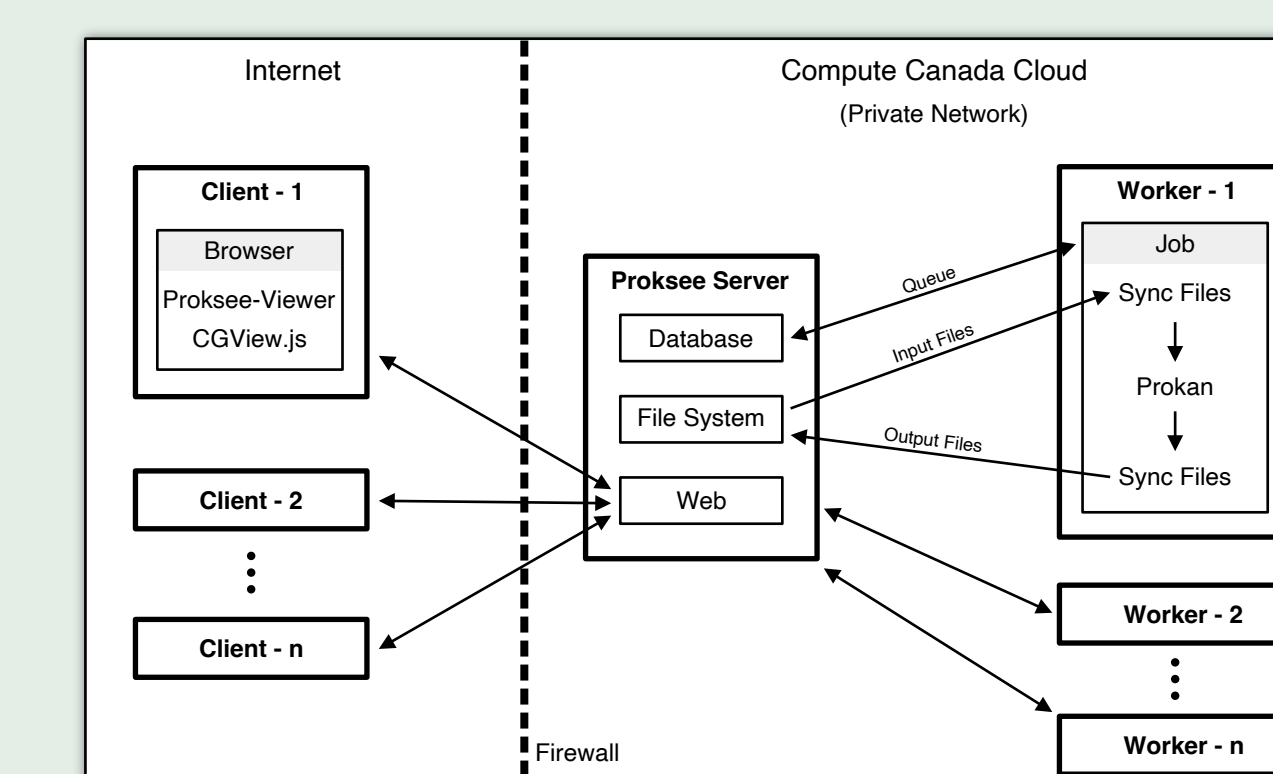
User Accounts

- Optional user accounts provide 200 MB for project storage
- Non-user projects are deleted 7 days after the last save date



Implementation

- Written in Ruby on Rails and JavaScript
- Annotation tools use Bioconda for external dependencies
- Multi-server cloud-based architecture



Contact: jason.grant@ualberta.ca