

# iReceptor+:

## From small project to international research software engineering - Part II

Brian Corrie

Technical Director, iReceptor

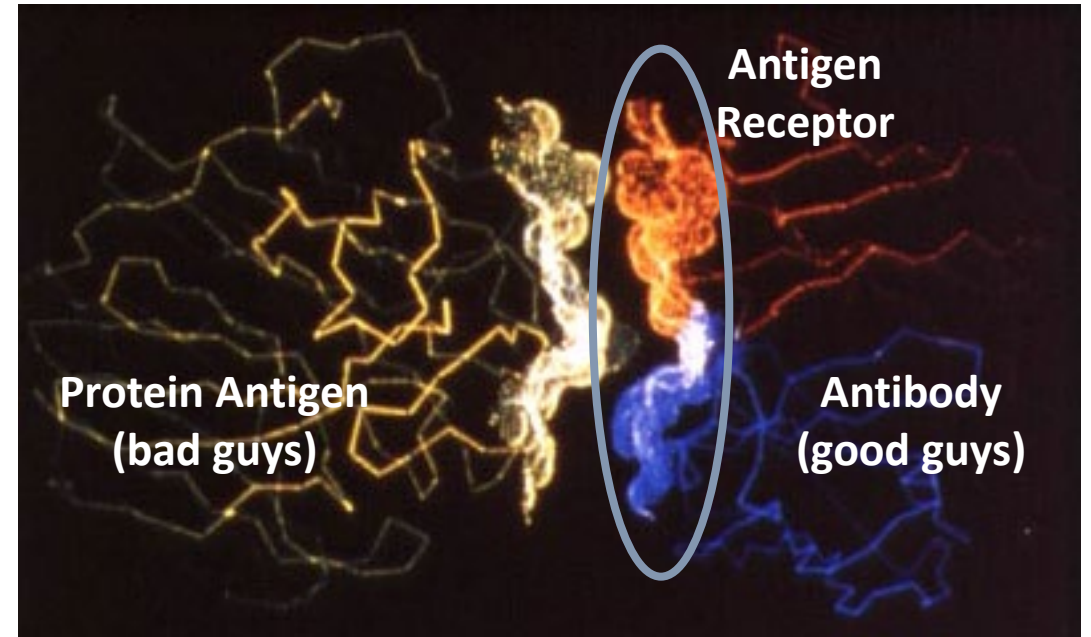
Technical Manager, iReceptor+



# Introduction

- Immunogenetics
  - Genetics of our immune system
- Goal: Personalized Immunotherapy
  - Cancer, autoimmune, and infectious disease
- How does the immune system work?
  - Antibody-producing cells bind to antigens
  - $\sim 10^{13}$  (?) possible antigen receptors
  - Adaptive  $\rightarrow$  Effective  $\rightarrow$  Hard to Understand
- What's changed? Adaptive Immune Receptor Repertoire Sequencing (AIRR-seq)
  - Deep sequencing of the immune response (2009)

## Adaptive Immune System



# Why iReceptor

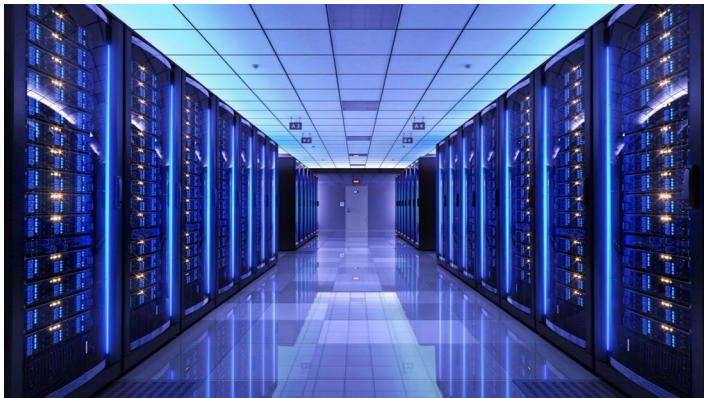
- Driver: Data Deluge from Next Gen Sequencing of Immune Response
  - $\sim 10^{13}$  possible antigen receptors
  - $\sim 10^8$  cells/person at a single time point
  - $> 10^6$  sequences per biological sample (statistical sampling)
  - $> 1$  time point/subject as immune system adapts
  - Many subjects/study, many studies/lab, 100s of labs
  - *New research area: emerging analysis tools, few standard processes*
- Researcher needs: AIRR-seq data that is FAIR
  - Find data of interest from multiple labs/institutions (securely)
  - Federate the data of interest (as authorized)
  - Perform complex analyses on that data (using advanced computational resources)



# The iReceptor Approach

## Advanced AIRR-seq Analysis

What do you do with all that data you just found?



Integration of analysis tools, management of data movement so user doesn't have to worry about it!

Job Management

Analysis Results

## iReceptor Scientific Gateway

Interactive data discovery, exploration, and analytics



Web based portal that hides the complexity of finding, searching and federating data across many repositories!

Data Federation

Data Query

## AIRR Data Commons

Distributed repositories, standards based



Based on standards developed by the international AIRR Community!



Jan 2014

Jan 2015

Jan 2016

Jan 2017

Jan 2018

Jan 2019

Jan 2020

Jan 2021

# CRSC 2018

iReceptor Timeline

CANARIE – NEP 131

EXT

CFI Cyberinfrastructure

CFI Cyberinfrastructure - Ops

Other proposals in the works...

Collaborations

AIRR Timeline

Pre-AIRR

AIRR Community

New AIRR Repositories

2014

2015

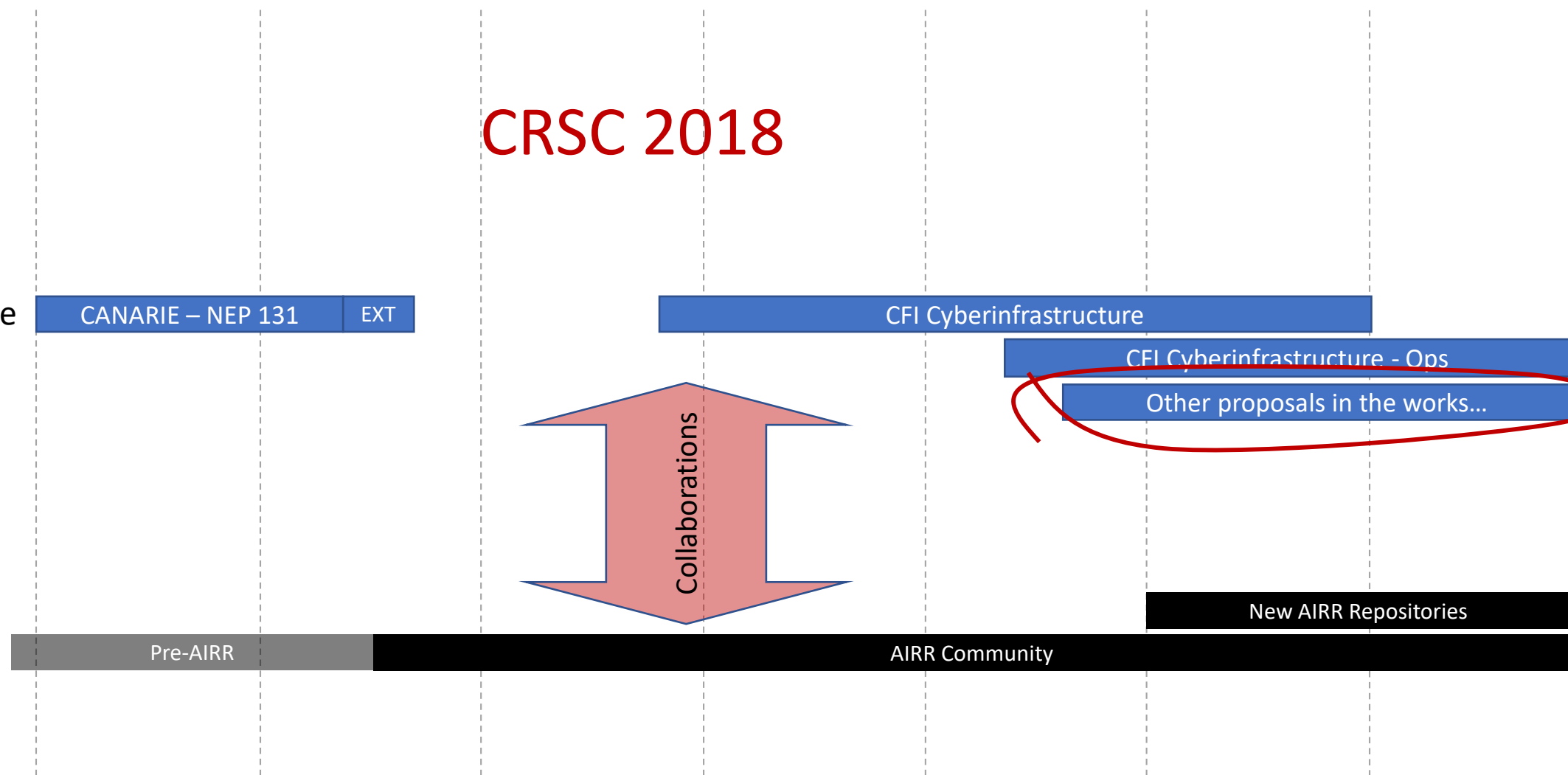
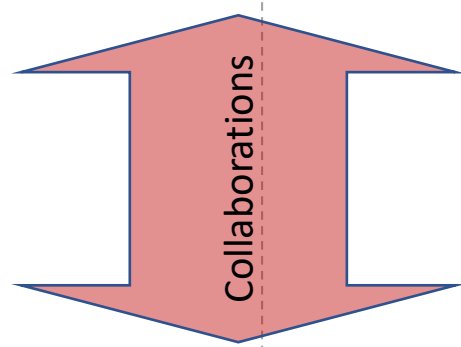
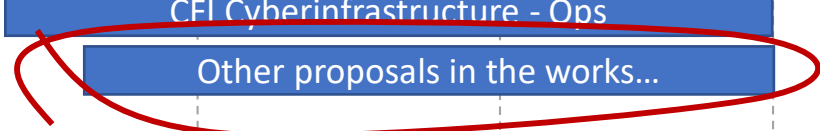
2016

2017

2018

2019

2020



Jan 2014

Jan 2015

Jan 2016

Jan 2017

Jan 2018

Jan 2019

Jan 2020

Jan 2021

CRSC 2019

iReceptor Timeline

CANARIE – NEP 131

EXT

CFI Cyberinfrastructure

CFI Cyberinfrastructure - Ops

CANARIE RS 238

CANARIE RS 238 - Ops

iReceptor+: EU/CIHR, 4 years

AIRR Timeline

Pre-AIRR

AIRR Community

New AIRR Repositories

2014

2015

2016

2017

2018

2019

2020

# What is iReceptor – CANARIE RS-238

- Expand to different immunology domains
- Expand AIRR Data Commons
  - iReceptor Public Archive (IPA)
  - VDJServer
  - Sorbonne
  - iReceptor Synthetic Archive (ISA)
  - **Your repository goes here!**
- iReceptor Turnkey Repository
  - Easy to install AIRR Repository
    - Docker containers
    - Mongo DB/Web Service
    - Web based query API



# What is iReceptor+

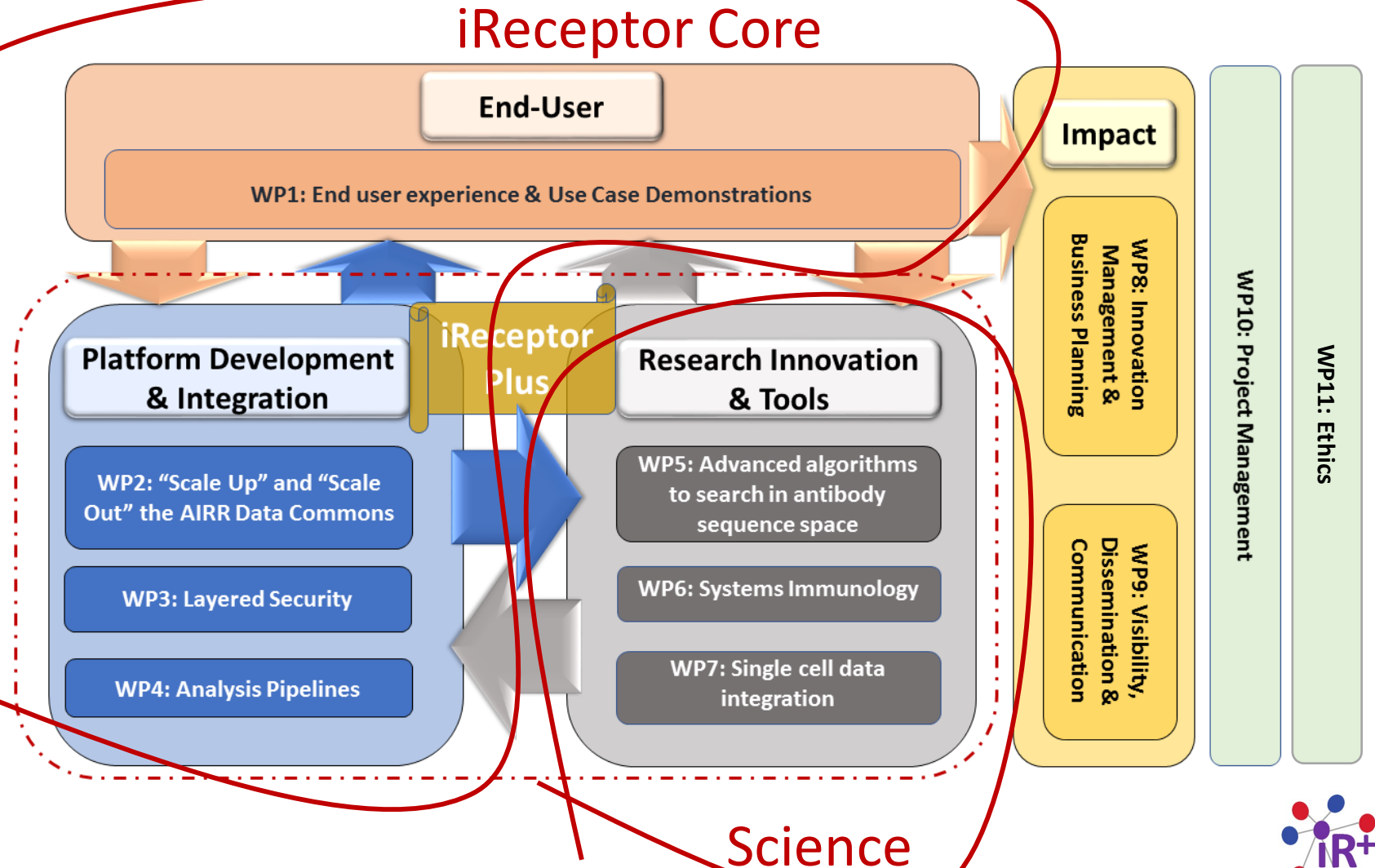
- Joint EU Horizon 2020/CIHR project (4 years, started Jan 2019)
  - Builds on CANARIE NEP-131, CFI Challenge 1, and CANARIE RS-238 funding
  - € 8.4M over four years, € 7.85M from EC, \$800 CAD from CIHR
  - 20 partners from 9 countries
    - Research (life sciences), research (computing science), clinic/hospital, industry
- It is big – and it is complex
  - Complex administratively!
  - Complex scientifically!
  - **Complex research software engineering project!**





# iReceptor+ Work Packages (WPs)

- WP structure
  - WPs have leaders
  - WPs have deliverables
- iReceptor core
  - WP1 – WP4
- Science extensions
  - WP5 – WP7
  - WP4 integration



# iReceptor+ RSE Challenges/Benefits

- Challenges

- 20 Partners = 20 different RSE processes
- 18 FTE over four years on research/technology/software
- 10 time zones (West coast North America to Israel)
- Language/culture (although English is the working language)
- Purpose of the project is to be integrative (need everyone to work together)

- Benefits

- A fair bit of money for personnel resources to do work!!!
- Many partners are from the AIRR Community – we already work with them well
- We have some strong (but new to the AIRR Community) RSE partners
- We have broad participation (clinical/hospital and industry)



# iReceptor+ RSE Approach

- iReceptor has always been user driven
  - Fundamental part of iR+ proposal
  - Working with partners (science/clinical/industry) to define use cases
- Use cases drive implementation in WPs
  - WPs map use cases to WP deliverables
  - WP deliverable mapping generates requirements for implementation
  - WP leaders manage WP implementation
- Oversight/Integration
  - Technical Manager provides iReceptor Core oversight/integration
  - Scientific Manager provides scientific oversight



# Can iR+ be agile(-ish)?

This cannot be enforced...

- Groups have their own processes, need to support how WPs want to work

... but it can be encouraged

- SFU iReceptor team is agile-ish
- AIRR Community and our AIRR iR+ partners are agile-ish in how we work
- Our software engineering partners are highly active in the agile community

Goal (personal\*):

- To have an agile-ish team and a continuously adaptable iR+ Core platform
- I will let you know how we are doing next year!

\*The positions expressed in this goal in no way reflects the goals of other partners in the iR+ project 8-)



Questions?

# Useful References

- Websites
  - iReceptor web site: [www.ireceptor.org](http://www.ireceptor.org)
  - iReceptor Scientific Gateway: [gateway.ireceptor.org](http://gateway.ireceptor.org)
  - iReceptor Plus: [www.ireceptorplus.org](http://www.ireceptorplus.org)
  - AIRR Community: [www.airr-community.org](http://www.airr-community.org)
- GitHub
  - <https://github.com/sfu-ireceptor>
  - <https://github.com/airr-community>
- Publications
  - iReceptor paper: Immunological Reviews – (DOI: 10.1111/imr.12666)
  - AIRR Community paper: Frontiers in Immunology – (DOI: 10.3389/fimmu.2017.01418)
  - AIRR Data Standards paper: Nature Immunology – (DOI: 10.1038/ni.3873)
  - AIRR Data Rep paper: Frontiers in Immunology – (DOI: 10.3389/fimmu.2018.02206)

