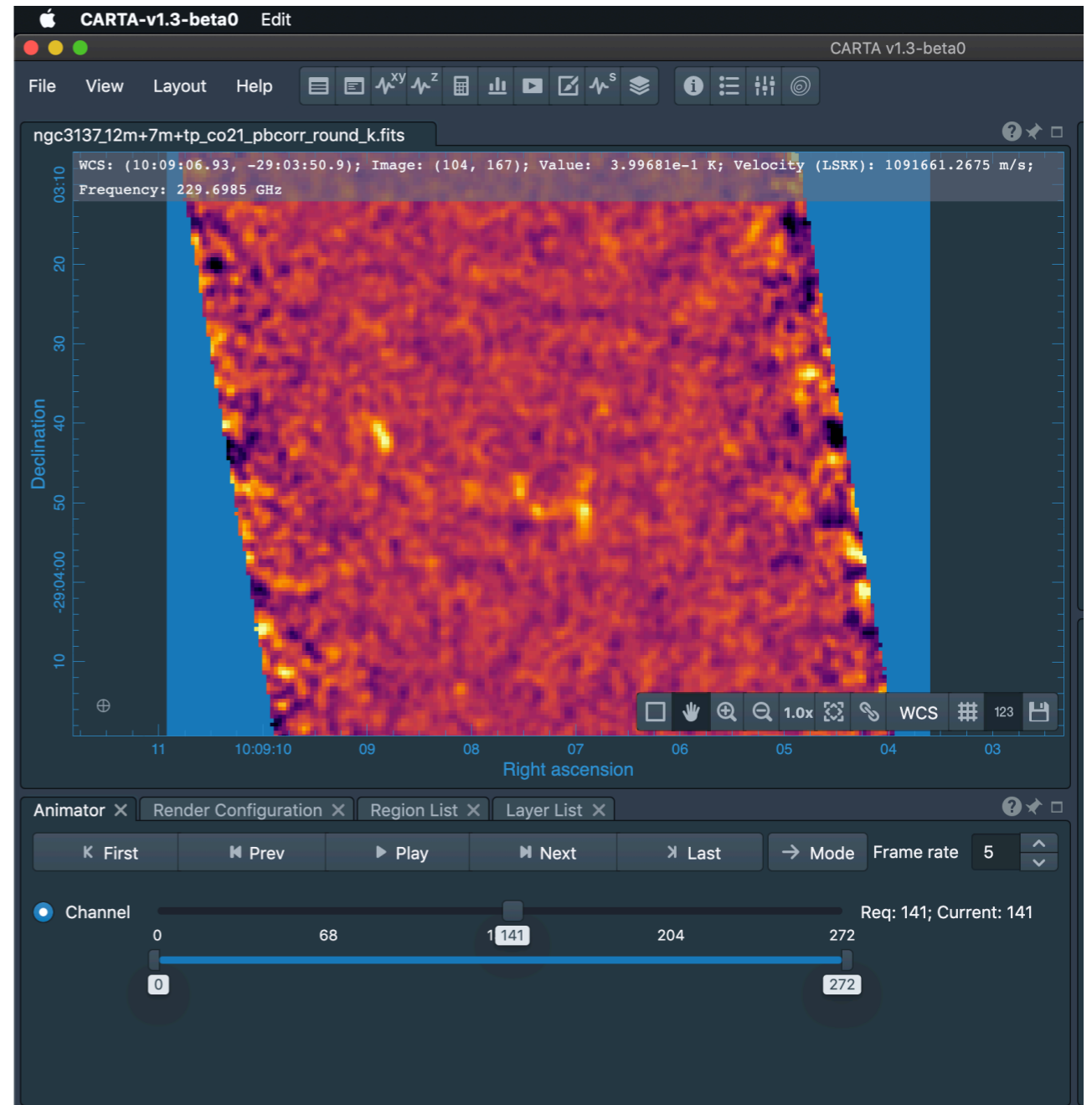


# Prime-Cam Data Reduction

Erik Rosolowsky

# Introduction

- Assoc. Professor in Canada
- Pipeline development and data delivery for surveys (Bolocam Galactic Plane Survey; 3 GBT Large Programs, PHANGS survey on ALMA/HST/VLT-MUSE)
- Software delivery for Canadian cyber-I projects (cyberSKA) and NRAO (CARTA)



# Pending Project Role

- Pending CFI funding
- Develop and maintain the Prime-Cam imaging pipeline with UBC team led by Douglas Scott (SCUBA2 SMURF software)
- This meeting will be “requirements capture” for me.

# Requirements Capture

Data rates

How much data will flow and at what stages? What will be the downsampled “raw” data rate (2 EB/yr → ?)

Data Transport

How does data move from telescope to processing to archive?

Data retention

What data needs to be retained at what stages?

Calibration strategies

How? How much and how rapidly will these evolve?

Computational requirements

How much CPU / GPU time is needed for imaging? How does this vary by products?

Imaging Products

What products are needed? What are their quality specifications?

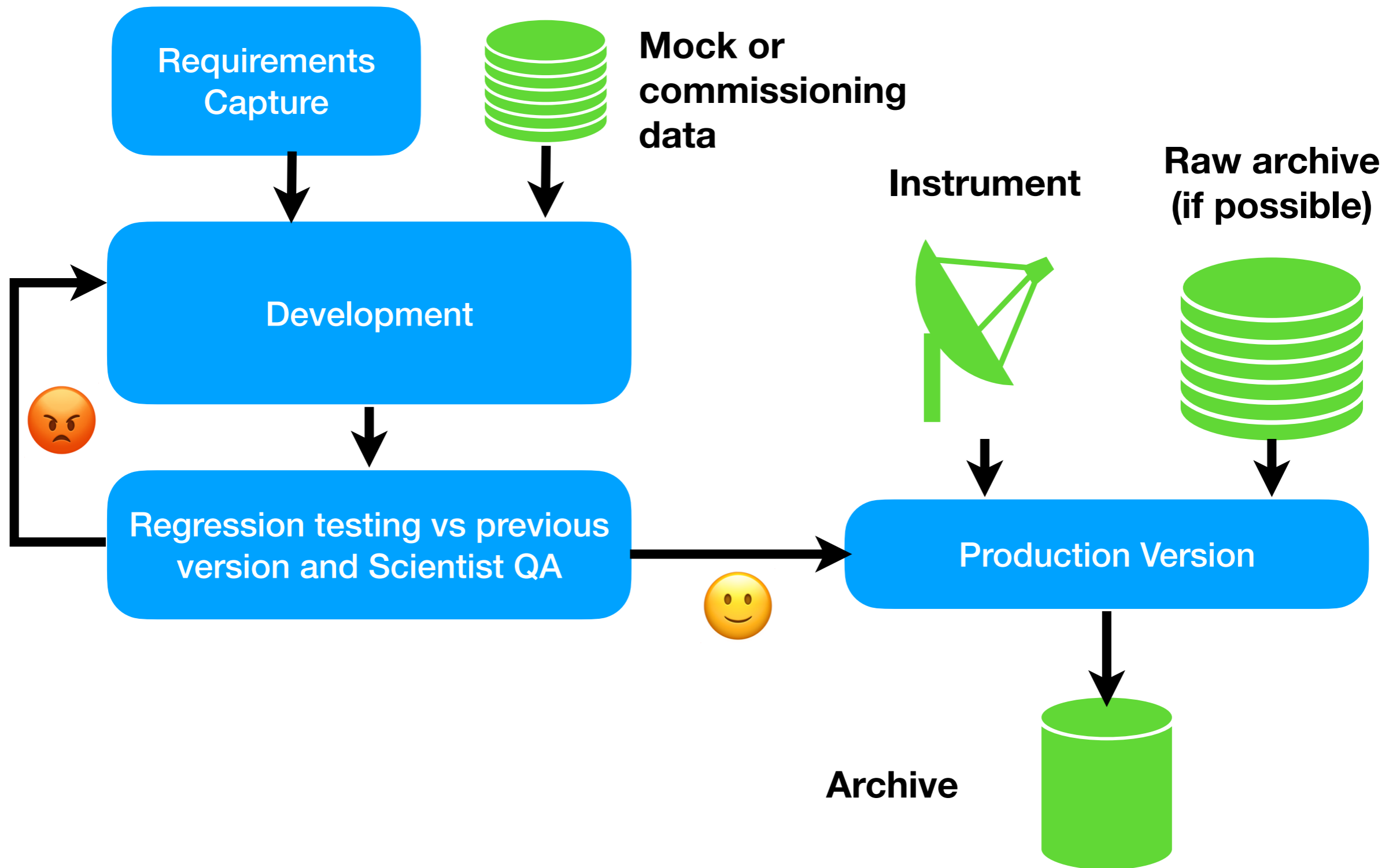
Responsiveness

How fast will some products be needed (transient science)?

Archiving

What data will be stored in an archive? Who will have access on what timescale? Capacity for reprocessing?

# Pipelining



# Development Strategies

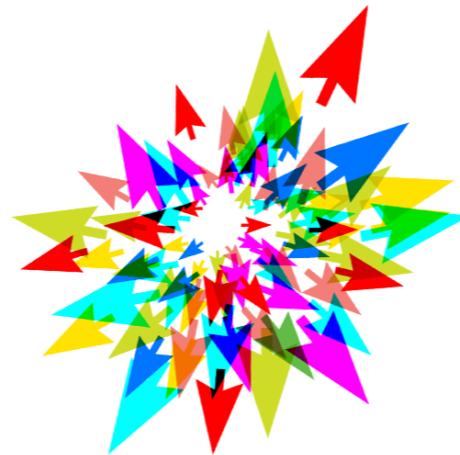
## Straw person

- Adopt-and-optimize wherever possible (e.g., use ACT software and strategies)
- Glue together pieces in a rapid prototyping language (probably python)
- Focus paid development time on optimization and/or porting, shift to GPU processing where needed
- Applicability of imaging software is less clear to me; do we need a separate reduction package?

# Archiving Strategy

Straw person

- Use Canadian national services to host archive and processing



**compute** | **calcul**  
canada | canada

- Create versioned public-release-lookalikes for internal use and access.

# Timelines and Contingencies

- No significant bandwidth until Q3 2020
- If no funding, I'll plan to contribute documentation back to the project
- Discussion on approach, timelines, strategy?



# Requirements Capture

Data rates      How much data will flow and at what stages?

Data Transport      How does data move from telescope to processing to archive?

Data retention      What data needs to be retained at what stages?

Calibration strategies      How? How much and how rapidly will these evolve?

Computational requirements      How much CPU / GPU time is needed for imaging? How does this vary by products?

Imaging Products      What products are needed? What are their quality specifications?

Responsiveness      How fast will some products be needed (transient science)?

Archiving      What data will be stored in an archive? Who will have access on what timescale? Capacity for reprocessing?