Improving Research Through Advanced REDCap Interfaces

scott.s.burns@vanderbilt.edu

Education and Brain Science Research Lab

EBRL

We study reading disabilities in children using behavioral and MR imaging measures

- Very wide databases
- Very expensive datasets
- Novel tasks (in & out of magnet)
- Many projects

Before REDCap

- Members touched every piece of data
- Issues joining across paradigms
- Saved and shared data in spreadsheets
- Always behind in analyses
- No traceable analyses

Input ≫ Output

After REDCap

- Analyze some data within milliseconds
- Automate everything possible
- Automate the automation
- Start analyses from a single source

Goals

- Advocate for advanced data management workflows
- Discuss solving problems using REDCap's Application Programming Interface
- Explain how Data Entry Triggers can connect infrastructure

Scaling Science

- More subjects & more captured data
- Humans don’t scale efficiently
- How to do better work in less time with less money?

Ideally...

Machines perform all **definable** analyses:

- Perform reproducible work
- Operate deterministically
- Orders of magnitude faster and cheaper

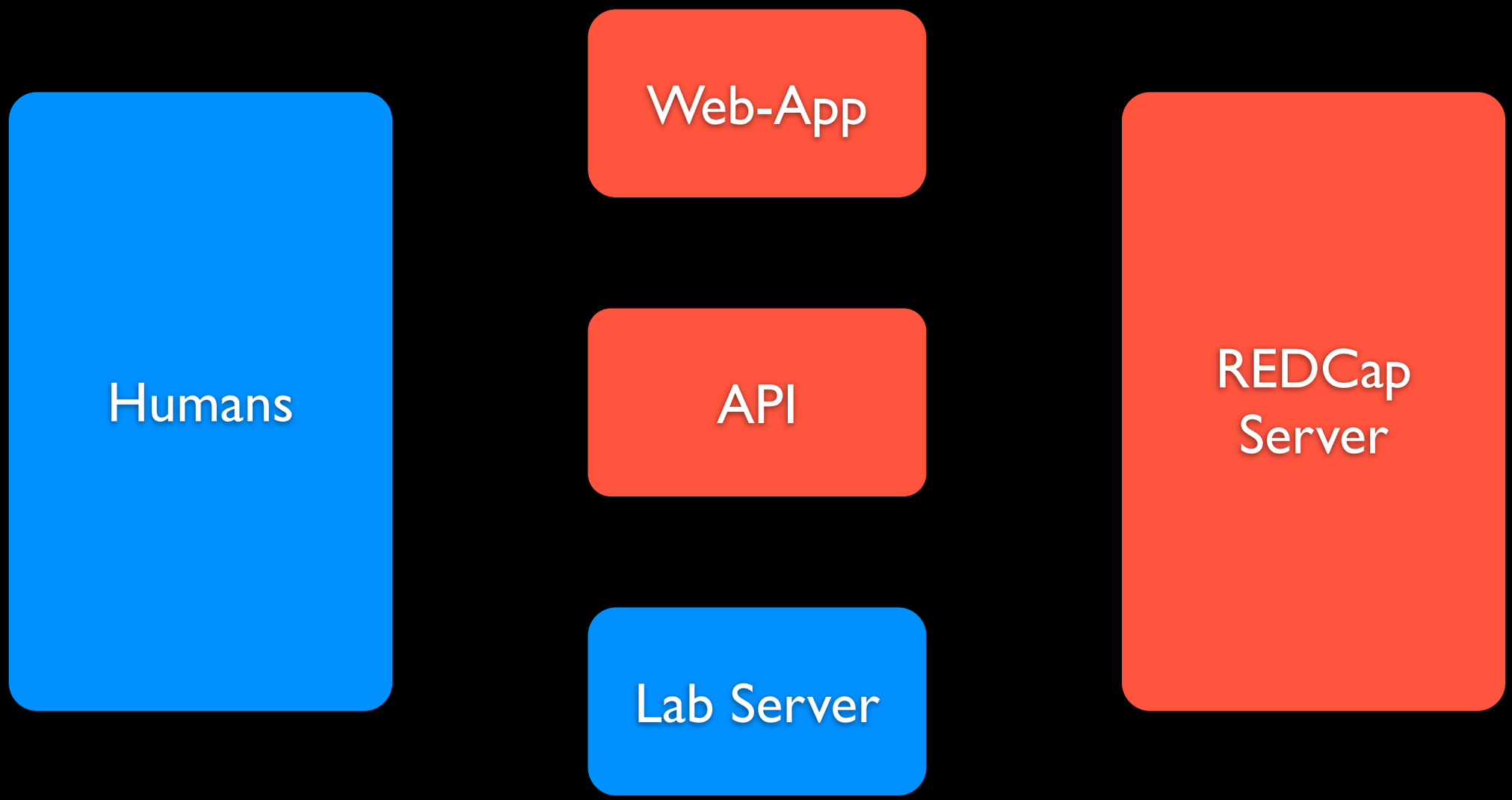
REDCap

- Is:
  - A service for collecting and storing data
  - Secure for the storage of PHI
  - An online spreadsheet
- Is not:
  - A relational database

Better than...

- A real database:
  - No administration
  - Easy schema definition
  - No security worries
- A spreadsheet:
  - GUI is browser-based
  - Client-Server architecture
  - Advanced web features

General Architecture



Advanced Features

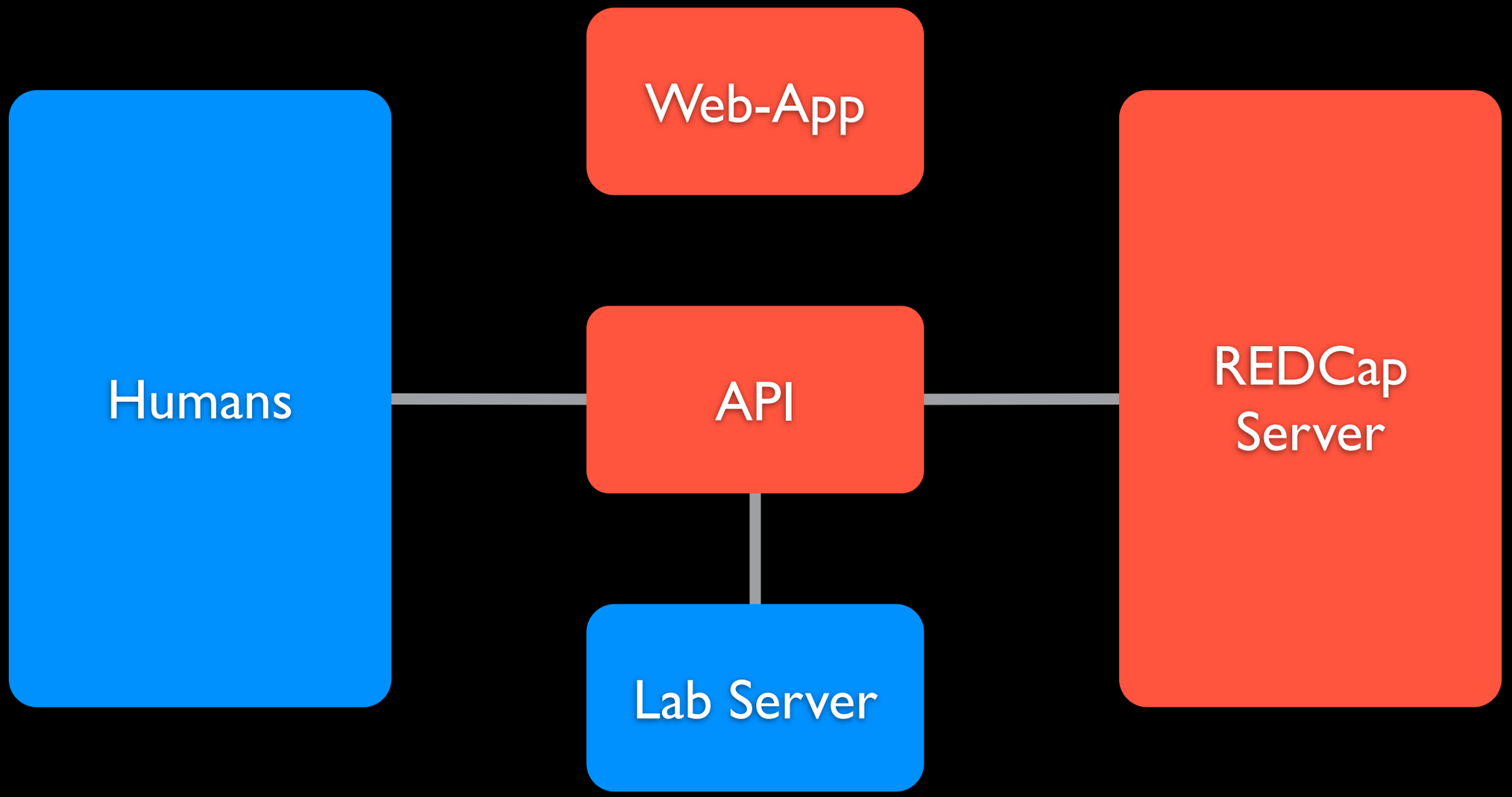
- Application Programming Interface (API)
  - Programmatic access to REDCap
- Data Entry Triggers
  - Automated notifications

All the building blocks we need

API

A method for software programs to ask for and push data to REDCap projects

API



Using the API

HTTP POST to API URL

Any programming environment with an HTTP library can use the API

(http://sburns.github.io/PyCap)

Major API Methods

- Metadata Export
- Data Export
- Data Import
- File Import, Export & Deletion

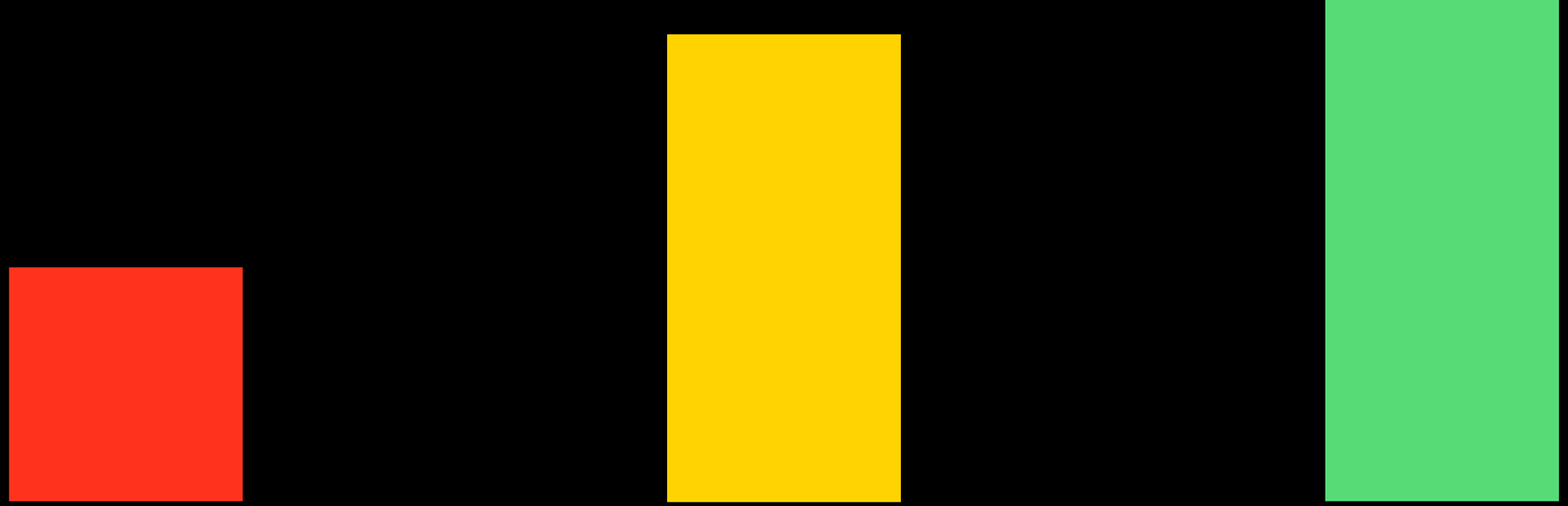
(https://redcap.vanderbilt.edu/api/help)

API: Possible Uses

- Advanced & automated field calculation
- Otherwise-impossible data upload
- REDCap as the input for external systems
- Shared Filesystem
- Across-project data movement

API: Field Calculation

Problem: How to update (many) fields across (many) records?



Download, Implement, Upload

REDCap Calculated Fields

API

Impossible Data Uploads

- Analyses can produce >1000 fields per record
- Collect 1000s of records per day

API: External Systems

- Hooks to external databases
- Reproducible cohort/group determination
- Automated database cleanup & backup

API: Shared Filesystem

How to insert or generate *intermediate* data to/from our analysis infrastructure?

- Secure
- Easy
- Automated

API: Shared Filesystem

File ⟶ fields:

- Software will:
  - Download file locally
  - Analyze file
  - Upload results to REDCap

API: Shared Filesystem

Fields ⟶ file:

- Software will:
  - Download data for that record
  - Substitute into a predefined template
  - Upload new report to REDCap
  - Alert lab members through email

API: Shuttle Data

- Capture data in one project
- Export and analyze through API
- Import results into same or other
- No need to duplicate data entry fields

API: Shuttle Data

- Capture data in public survey
- Manually verify
- Easily copy to new record in private project

API

API: Pitfall

API only serves external requests

- One-off scripts
- Scheduled programs

API: Pitfall

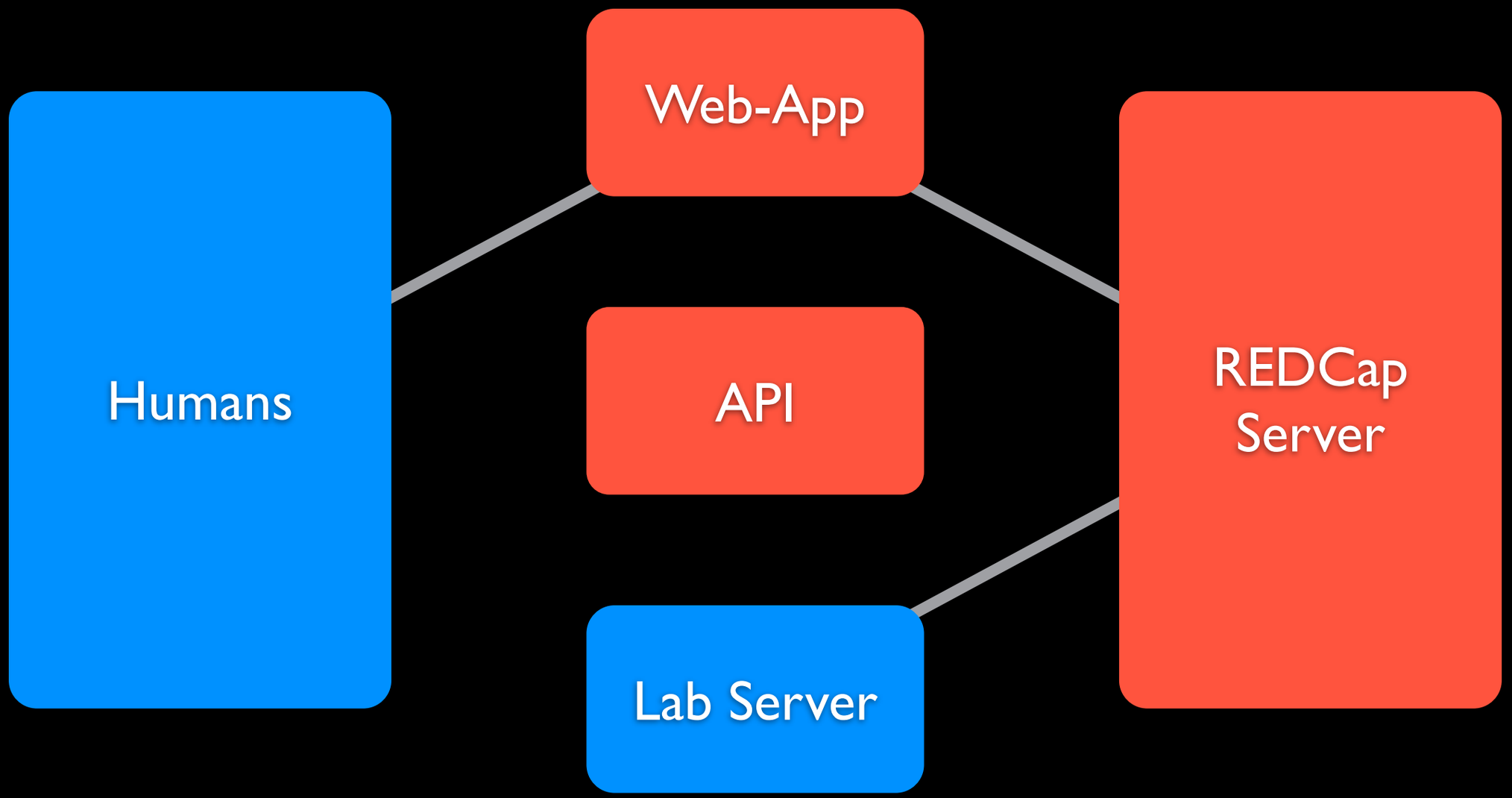
Better idea about **when** to run analyses?

Data Entry Triggers

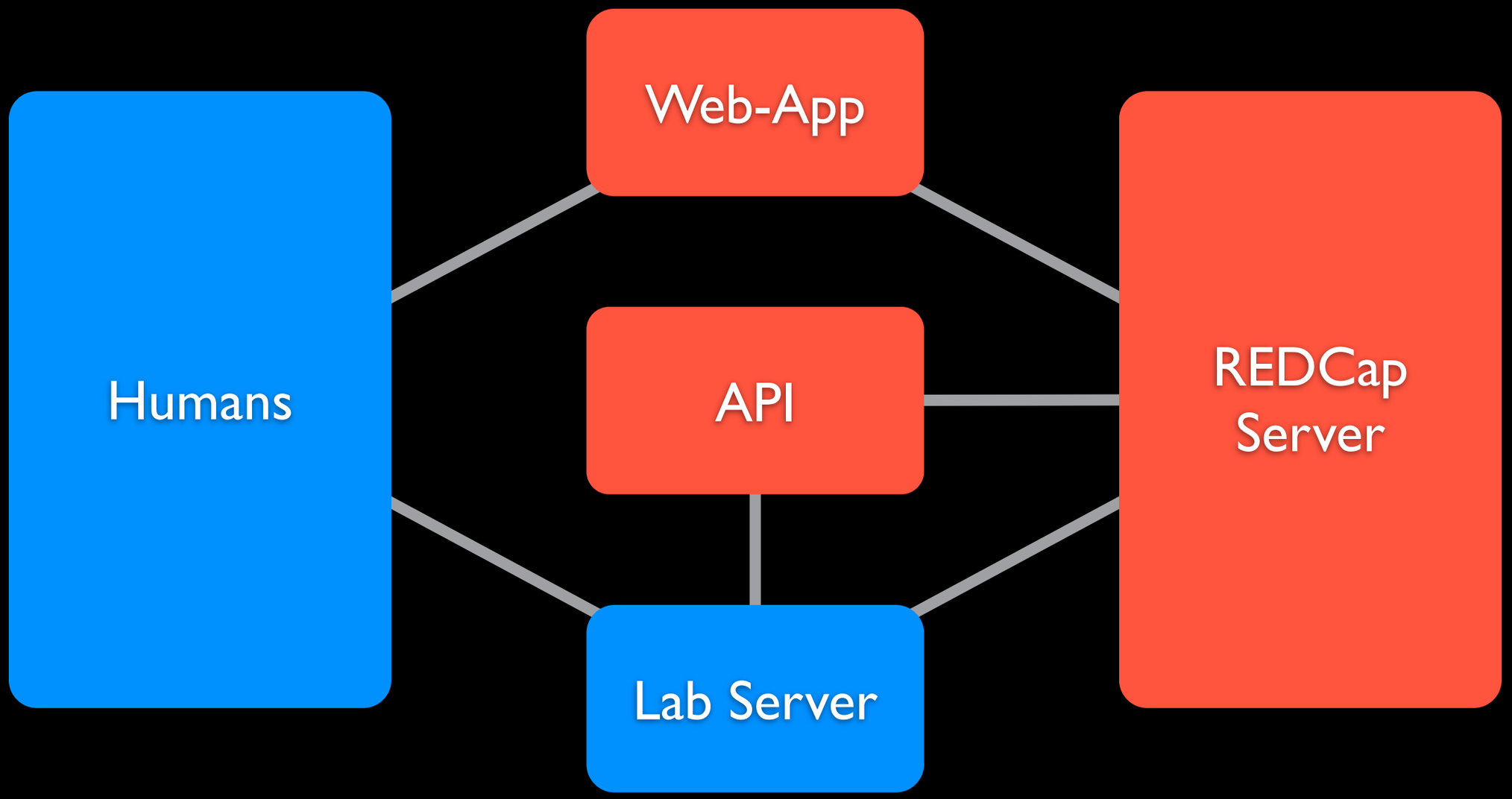
- Independent of but complimentary to API
- Register a URL to your project
- Internet notification when data is saved
- Notification contains context of the save

Data Entry Triggers

Data Entry Triggers



(Super) Data Entry Triggers



Data Entry Triggers: Pitfalls

- Not every research group:
  - Can setup, maintain & secure a web server
  - Has the resources to write the web-app

But every lab should have access to this infrastructure!

Switchboard

- I wrote a web-app to:
  - Parse incoming REDCap requests
  - Execute functions that “match” the request
- **In production** for our lab

(http://github.com/sburns/switchboard)

Data Entry Triggers

In a perfect world, we all share a KC-wide web-server

- Just one server to maintain & protect
- Sharing is good
- Remove excuses for buy-in
- Everyone benefits from optimization

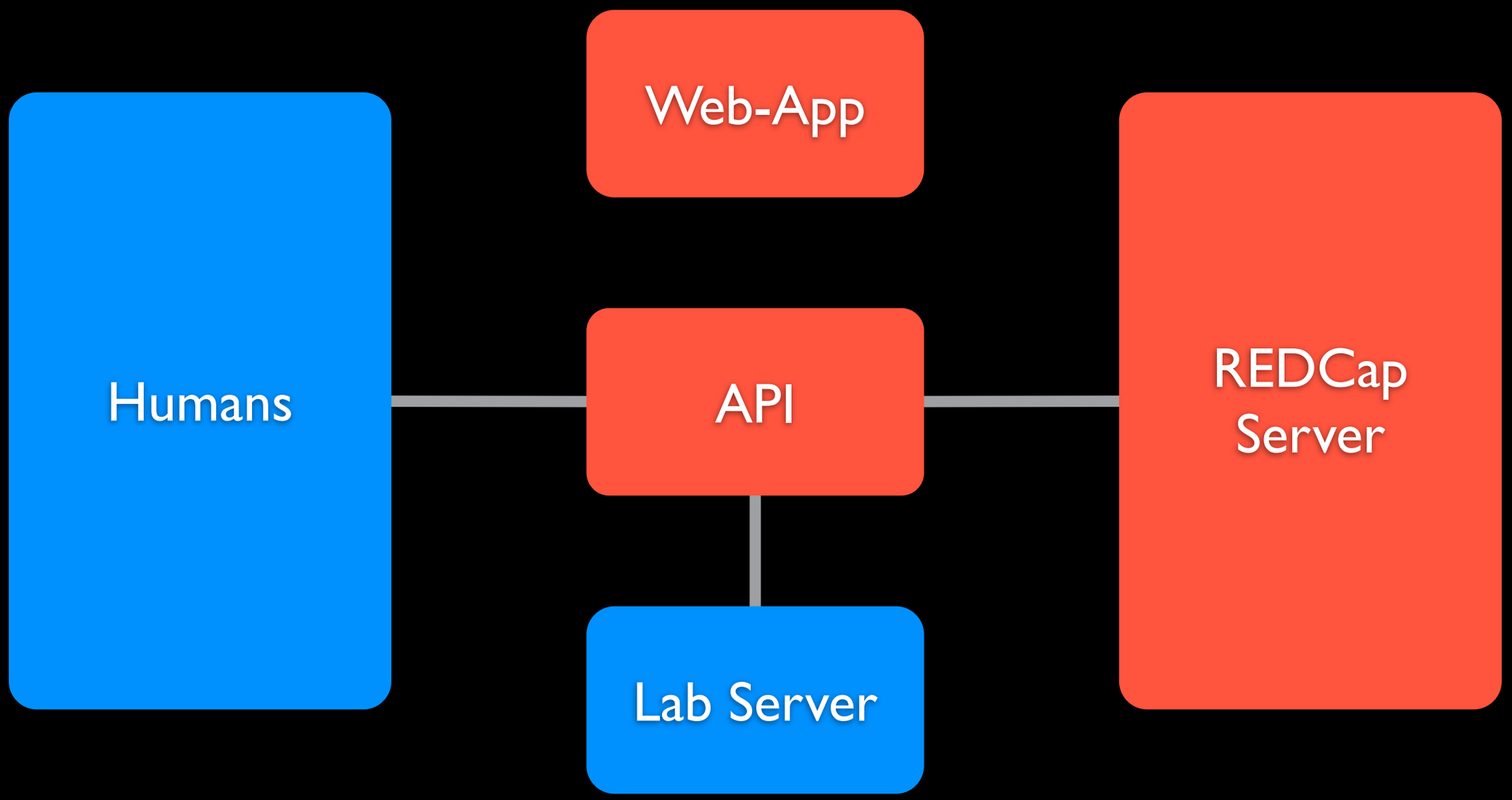
Conclusions

Engineering Better Science

- All the pieces exist to offload a massive amount of data-management work from humans to machines
- Cost-effective and improves work through improved accuracy and reproducibility
- Let machines do that which can be defined
- Let humans do the hard work

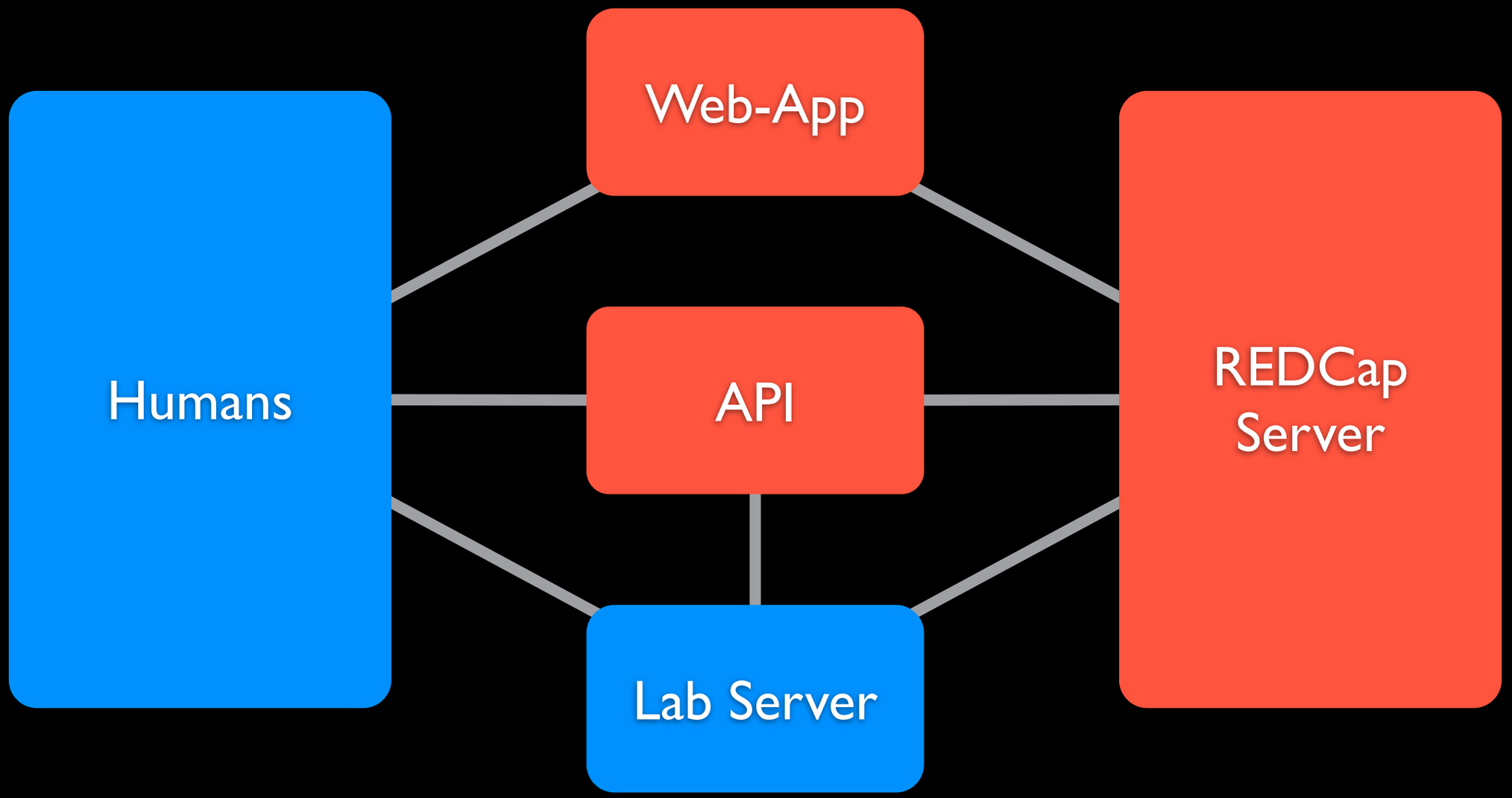
Automation improves research

(Easier to automate machines than humans)



Automate the automation

(Machines don’t make excuses)



Thank you

Laurie Cutting, Ph.D.

Nikki Davis, Ph.D.

Sheryl Rimrodt, M.D.

REDCap Team (Paul Harris, Rob Taylor, etc)

Email: scott.s.burns@vanderbilt.edu

Github: http://github.com/sburns

API& PyCap Tutorial: http://bit.ly/pycap-tutorial

Questions & Comments?