

Cornell University

Roper Center: Where are we now? Where are we going?

Tim Parsons, Software Engineer, Roper Center

IASSIST & CART  2018



Cornell University



Overview

- Review of the Roper Center
- Accomplishments since 2016
- Discovery Process
- Roadmap
- Q & A



About the Roper Center

- Largest archive of public opinion data in the world
- Founded in 1947, with data going back to the 1930s
- iPOLL Search – subscription service to search the Roper Center Archive
- IT Staff of 3
 - Two developers
 - One Systems Administrator



Recap from 2016

- 3 phased approach
 - 0.0 – Current state
 - 1.0 – Rebuilt applications with same functionality as 0.0
 - 2.0 – Additional features and enhancements



Why are we doing this?

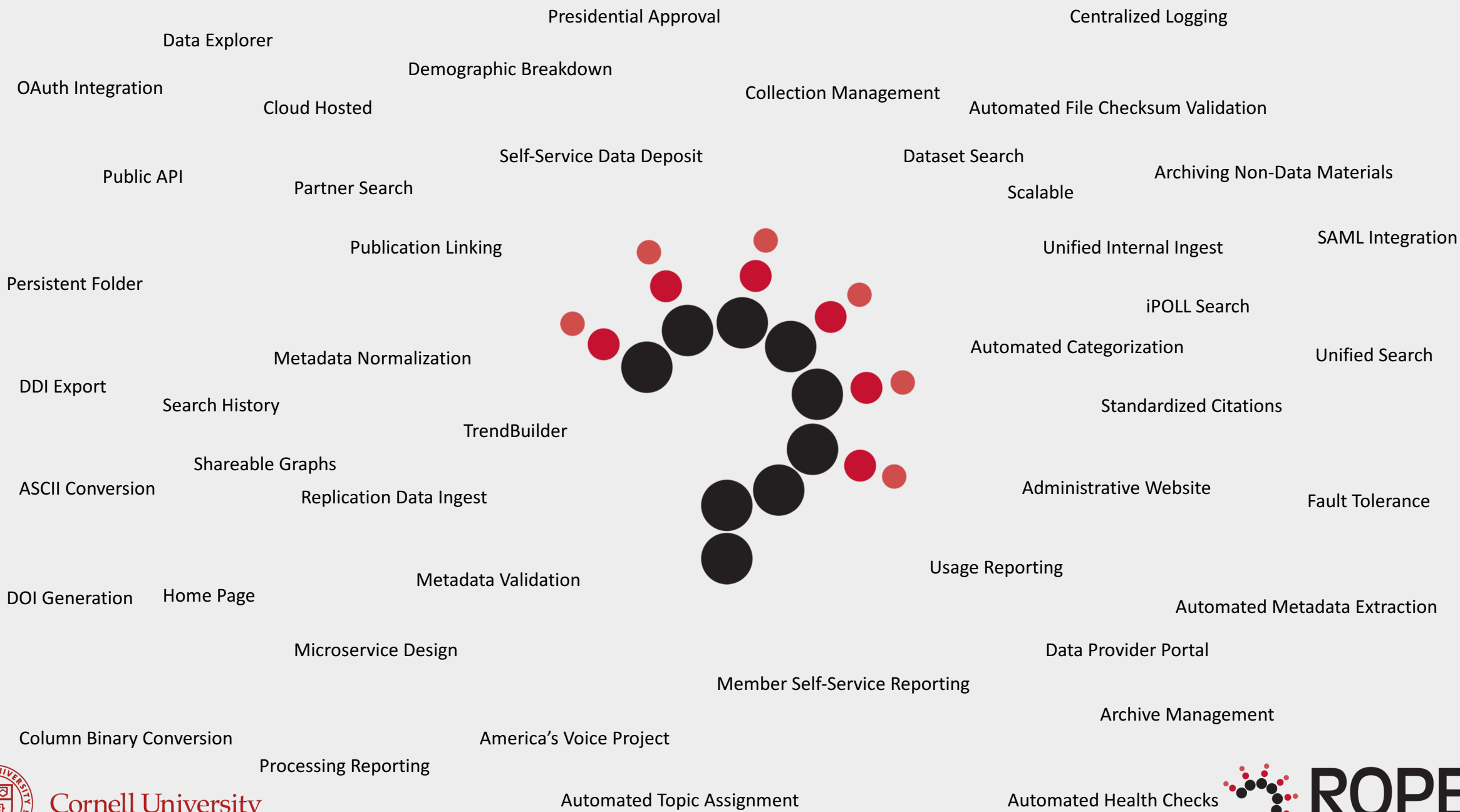
- Update the look and feel
- Provide better search functionality
- Provide better metadata
- Increase security
- Increase speed at which we can deliver new features



Accomplishments

- Built a new archive management application
 - Guided workflow style
 - Unified system for ingest of questions and datasets
- Normalized and migrated old archive into new archive
- Migrated current architecture to the cloud
- Automated column binary conversion





Cornell University



ROPER
CENTER
For Public Opinion Research

Projects and Versions

iPOLL

v1.0

- Public API
- Unified Search
- SAML Integration
- OAuth Integration
- Demographic Breakdown
- Standardized Citations
- COUNTER Usage Reports
- Search History
- Persistent Folders
- Partner Search

v2.0

- Member Self-Service Reporting
- TrendBuilder
- Data Explorer
- DDI Export

Archive Management

v1.0

- Unified Ingest
- Metadata Normalization
- Metadata Validation
- Self-Service Data Deposit
- Central Archive Management
- DOI Generation
- Column Binary Conversion
- Automated File Integrity Validation

v2.0

- Automated Categorization
- Automated Topic Assignment
- Automated Metadata Extraction
- ASCII Conversion
- Replication Data Ingest

Presidential Approval

v1.0

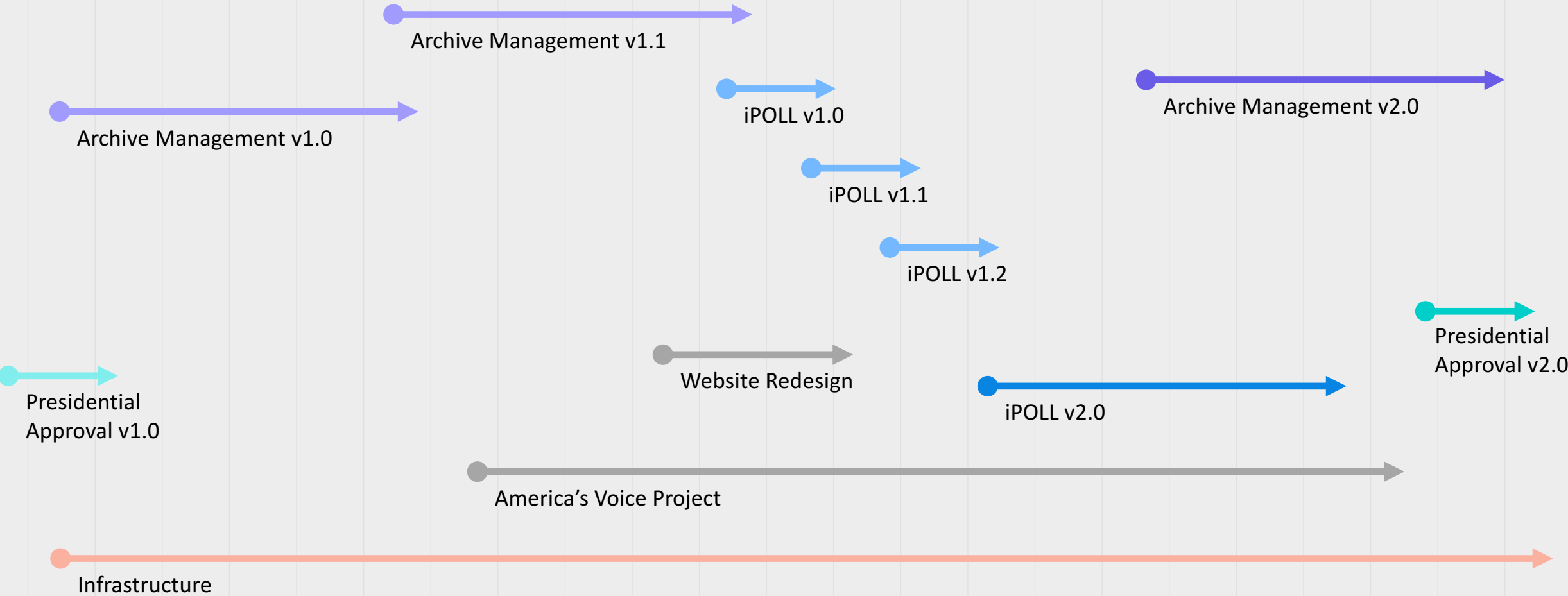
- View presidential approvals
- Graph Builder
- Shareable Graphs

v2.0

- Global 'Executive' approval

Infrastructure

- Centralized Logging
- Cloud Hosted
- Fault Tolerant
- Automated System Health Checks
- Microservice Architecture
- Scalable

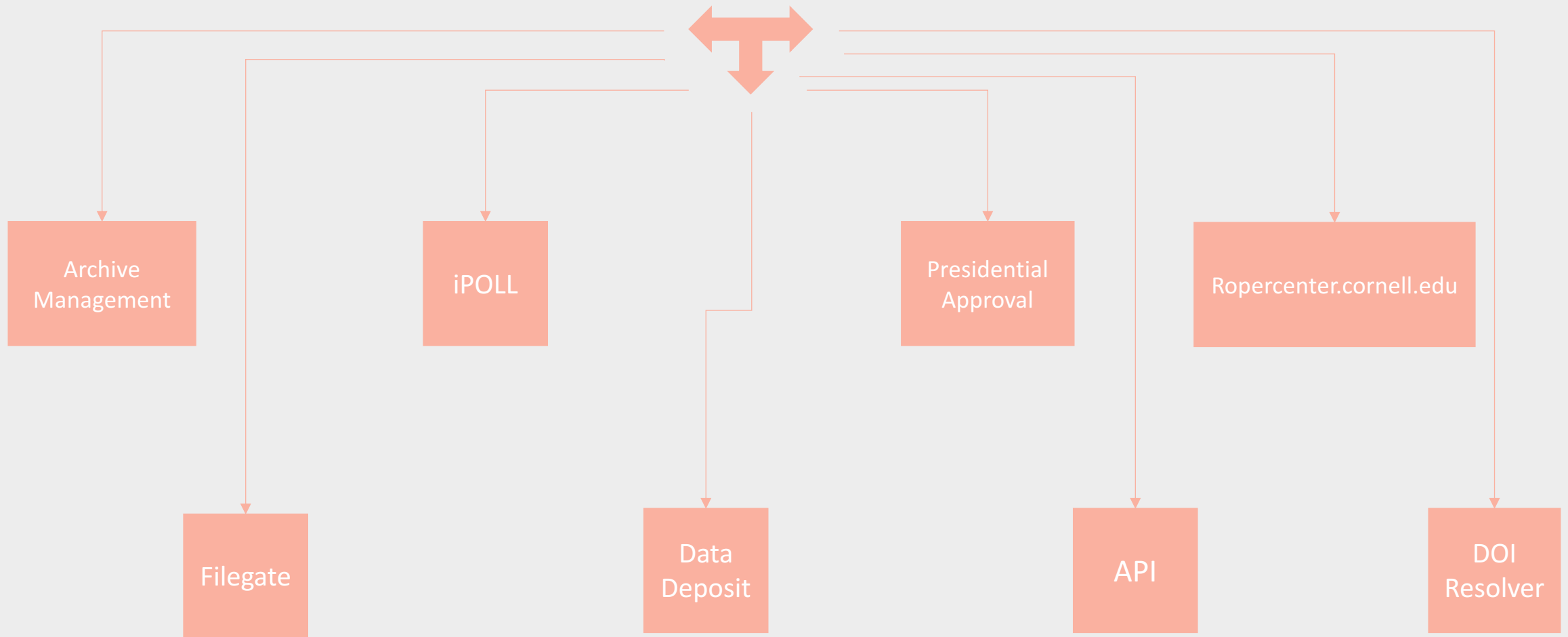


Release Methodology

- Phased Approach
 - New systems live alongside old systems
 - Shut off specific systems when new systems are rebuilt
- Migration Service
 - Allows Archive Management system to edit data, then migrate to old systems for public searching.



Infrastructure



In Closing

- We've been busy!
- Building a new archive is hard
- There's a lot more to do
 - Finish iPOLL
 - Build v2 features for iPOLL, Archive Management, Presidential Approval

Questions?



Cornell University



Thank you!

Tim Parsons - tim.parsons@cornell.edu



Cornell University

