



# Uncharted Territory



## Navigating a Community for the First Time

Leah Cole, Google Cloud DevRel  
October 2019



Hi, I'm Leah!



Hi, I'm Leah!



Hi, I'm Leah!



# Leah + Google + OSS

	<b>Leah Cole, November 2018</b>
OSS	Casual consumer
Apache	I think I've heard of it?
Airflow	Is it wind?

# Leah + Google + OSS

	<b>Leah Cole, November 2018</b>	<b>Google</b>
OSS	Casual consumer	Consumer, Contributor, Producer
Apache	I think I've heard of it?	Consumer, Contributor, Critical to some Cloud Products
Airflow	Is it wind?	Cloud Composer

# Unknown unknowns



~~Unknown~~  
unknowns

# How?

# How? Community Assessment







# Assessing a Community

# Assessing a Community



# Assessing a Community

## What Did I Look At?

- Community Taxonomy 
- Usage 
- Issues 
- Contribution 
- Releases 
- Similar Tools/Products 



# Community Taxonomy



# Community Taxonomy



- Structure



# Community Taxonomy



- Structure
- Committers

# Community Taxonomy



- Structure
- Committers
  - Who?

# Community Taxonomy



- Structure
- Committers
  - Who?
  - How many?

# Community Taxonomy



- Structure
- Committers
  - Who?
  - How many?
- Companies?

# Community Taxonomy

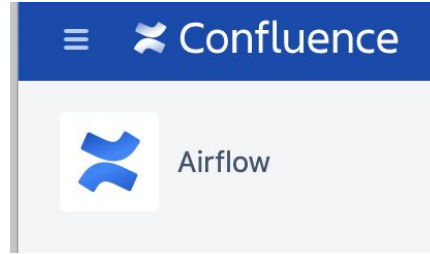


- Structure
- Committers
  - Who?
  - How many?
- Companies?
- Origin

# Community Taxonomy



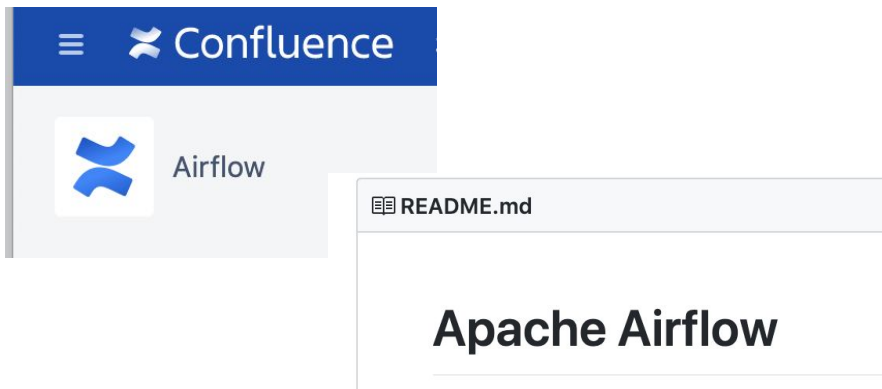
- Structure
- Committers
  - Who?
  - How many?
- Companies?
- Origin



# Community Taxonomy



- Structure
- Committers
  - Who?
  - How many?
- Companies?
- Origin



# Community Taxonomy



- Structure
- Committers
  - Who?
  - How many?
- Companies?
- Origin

A collage of images related to Apache Airflow. At the top left is the Confluence logo (three horizontal lines) and the word "Confluence" in white on a blue background. Below it is the Airflow logo (two blue curved lines) and the word "Airflow". To the right is a document icon and the text "README.md". Below that is the Apache Airflow logo (a colorful pinwheel) and the text "Apache Airflow". At the bottom is the text "Apache Airflow Documentation".



# Community Taxonomy



- Structure
- Committers
  - Who?
  - How many?
- Companies?
- Origin

The screenshot shows a Confluence page for Apache Airflow documentation. At the top, there's a blue header with the Confluence logo and the word 'Confluence'. Below that, there's a section for 'Airflow' with the Airflow logo. The main content area has the title 'Apache Airflow' and 'Apache Airflow Documentation'. At the bottom, there's a search bar and a list of recent discussions. The discussions include:

- SCUSS] Change all example\_dags's start\_date to specific datetime** (3 people, 2 replies, 2019-05-13 11:03)
- SCUSS] Change all example\_dags's start\_date to specific datetime** (2 people, 2 replies, 2019-05-13 10:57)
- : Bay Area Airflow meetup** (4 people, 7 replies, 2019-05-10 22:54)
- ification request from Season of doc Mentor (SDOC2019)** (2 people, 1 reply, 2019-05-10 22:45)

Usage



# Usage



- Individuals/Companies

# Usage



- Individuals/Companies
- How many?

# Usage



- Individuals/Companies
- How many?
- Dependencies?

# Usage



- Individuals/Companies
- How many?
- Dependencies?

## Statistics

View statistics for this project via [Libraries.io](#), or by using [Google BigQuery](#)


# Usage



- Individuals/Companies
- How many?
- Dependencies?

## Statistics

View statistics for this project via [Libraries.io](#), or by using [Google BigQuery](#)

 Used by ▼ 508

 Watch ▼ 662

 Star 14.3k

 Fork 5.4k

Issues





# Issues



- # open issues

# Issues



- # open issues
- Date of last resolution

# Issues



- # open issues
- Date of last resolution
- How long?

# Issues



- # open issues
- Date of last resolution
- How long?
- How to submit?

# Issues



- # open issues
- Date of last resolution
- How long?
- How to submit?
- How easy for newbies to fix?

# Issues



- # open issues
- Date of last resolution
- How long?
- How to submit?
- How easy for newbies to fix?

## Open issues Switch filter ▾

Order by Priority ▾

🔴 AIRFLOW-4355

Externally triggered DAG is marked as 'succes...

🔴 AIRFLOW-4089

pytz.utc has no attribute name

🔴 AIRFLOW-4470

RBAC Github Enterprise OAuth provider callba...

🟢 AIRFLOW-4390

manage\_slas() needs refactoring

🔴 AIRFLOW-2898

Task not entering queued state for pool

🔴 AIRFLOW-4424

Scheduler does not terminate after num\_runs ...

🟢 AIRFLOW-2290

Include CVE references in changelog

🔴 AIRFLOW-4346

Kubernetes Executor Fails for Large Wide DAGs

🔴 AIRFLOW-4363

Encounter JSON Decode Error when using do...

🔴 AIRFLOW-3036

Upgrading to Airflow 1.10 not possible using G...

🔴 AIRFLOW-4176

[security] webui shows password - admin/log/?...

🟢 AIRFLOW-2289

Add additional quick start to INSTALL

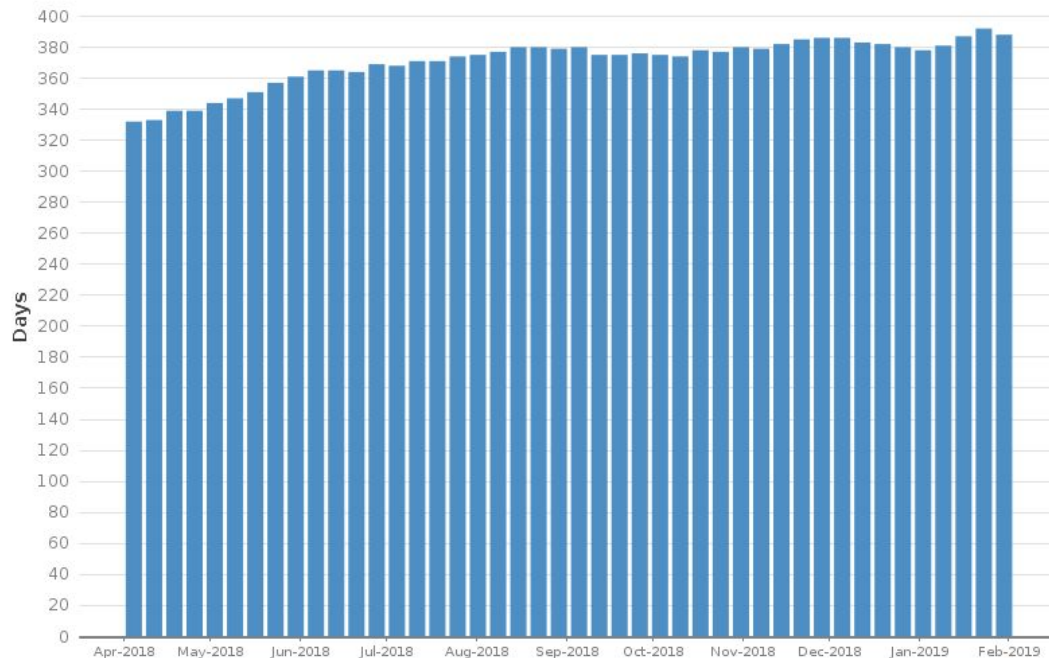
🔴 AIRFLOW-2222

# Issues



- # open issues
- Date of last resolution
- How long?
- How to submit?
- How easy for newbies to fix?

This chart shows the average number of days issues were unresolved for on a given day over the past 300 days.

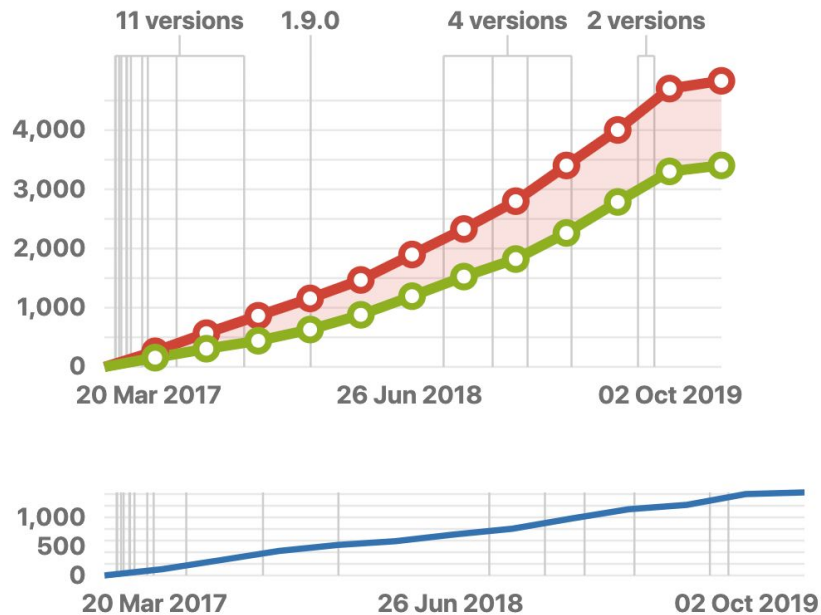


# Issues



- # open issues
- Date of last resolution
- How long?
- How to submit?
- How easy for newbies to fix?

## Created vs Resolved Chart: Apache Airflow



**Issues in the last 1,000 days** (grouped quarterly) [View in Issue Navigator](#)

- Created issues (4833)
- Resolved issues (3404)



# Contribution



# Contribution



- Who contributes?

# Contribution



- Who contributes?
  - Individuals

# Contribution



- Who contributes?
  - Individuals
  - Companies

# Contribution



- Who contributes?
  - Individuals
  - Companies
- How to Start?

# Contribution



- Who contributes?
  - Individuals
  - Companies
- How to Start?
- PR merge frequency

# Contribution

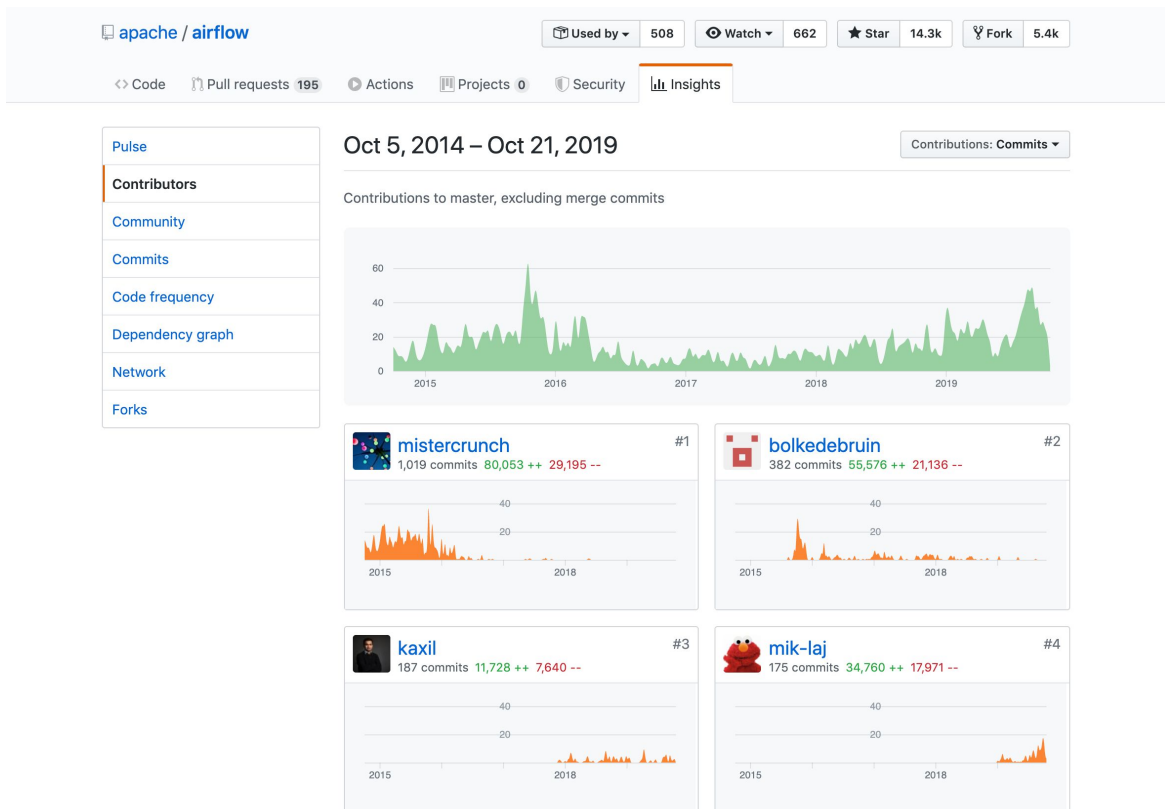


- Who contributes?
  - Individuals
  - Companies
- How to Start?
- PR merge frequency
- Whose PRs get merged?

# Contribution



- Who contributes?
  - Individuals
  - Companies
- How to Start?
- PR merge frequency
- Whose PRs get merged?

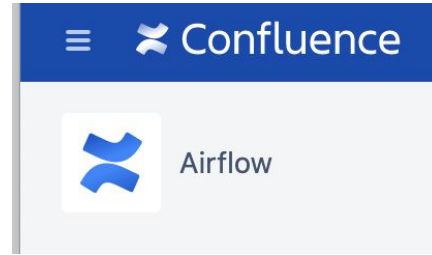




# Contribution



- Who contributes?
  - Individuals
  - Companies
- How to Start?
- PR merge frequency
- Whose PRs get merged?



# Contribution



- Who contributes?
  - Individuals
  - Companies
- How to Start?
- PR merge frequency
- Whose PRs get merged?

The screenshot shows the GitHub repository for Apache Airflow. At the top, it displays the repository name 'apache / airflow' and various statistics: 'Used by 508', 'Watch 662', 'Star 14.3k', and 'Fork 5.4k'. Below this, there are navigation tabs for 'Code', 'Pull requests 195', 'Actions', 'Projects 0', 'Security', and 'Insights'. The current branch is 'master' and the file being viewed is 'airflow / CONTRIBUTING.rst'. A commit by 'potiuk' is highlighted, with the message 'Google Season of Docs updates to CONTRIBUTING doc (#6283)' and a link to the pull request. Below the commit information, there are 5 contributors. The file content shows a table of contents for the 'CONTRIBUTING.rst' file, listing sections such as 'Contributions', 'Documentation', and 'Pre-commit Hooks'.

Branch: master | airflow / CONTRIBUTING.rst

potiuk [AIRFLOW-XXXX] Google Season of Docs updates to CONTRIBUTING doc (#6283) 10d39ef 8 days ago

5 contributors

684 lines (459 sloc) | 26.3 KB

Raw Blame History

- [Contributions](#)
  - [Report Bugs](#)
  - [Fix Bugs](#)
  - [Implement Features](#)
  - [Improve Documentation](#)
  - [Submit Feedback](#)
- [Documentation](#)
- [Pull Request Guidelines](#)
- [Development Environments](#)
  - [Local virtualenv Development Environment](#)
  - [Breeze Development Environment](#)
- [Pylint Checks](#)
- [Pre-commit Hooks](#)
  - [Prerequisites for Pre-commit Hooks](#)
  - [Enabling Pre-commit Hooks](#)
  - [Using Docker Images for Pre-commit Hooks](#)

# Releases



# Releases



- How often?

# Releases



- How often?
- Apache Way?

# Releases



- How often?
- Apache Way?
- Non-binding opinions?

# Releases



- How often?
- Apache Way?
- Non-binding opinions?
  - Are they listened to?

# Releases



- How often?
- Apache Way?
- Non-binding opinions?
  - Are they listened to?
- Semantic Versioning?



# Releases



- How often?
- Apache Way?
- Non-binding opinions?
  - Are they listened to?
- Semantic Versioning?

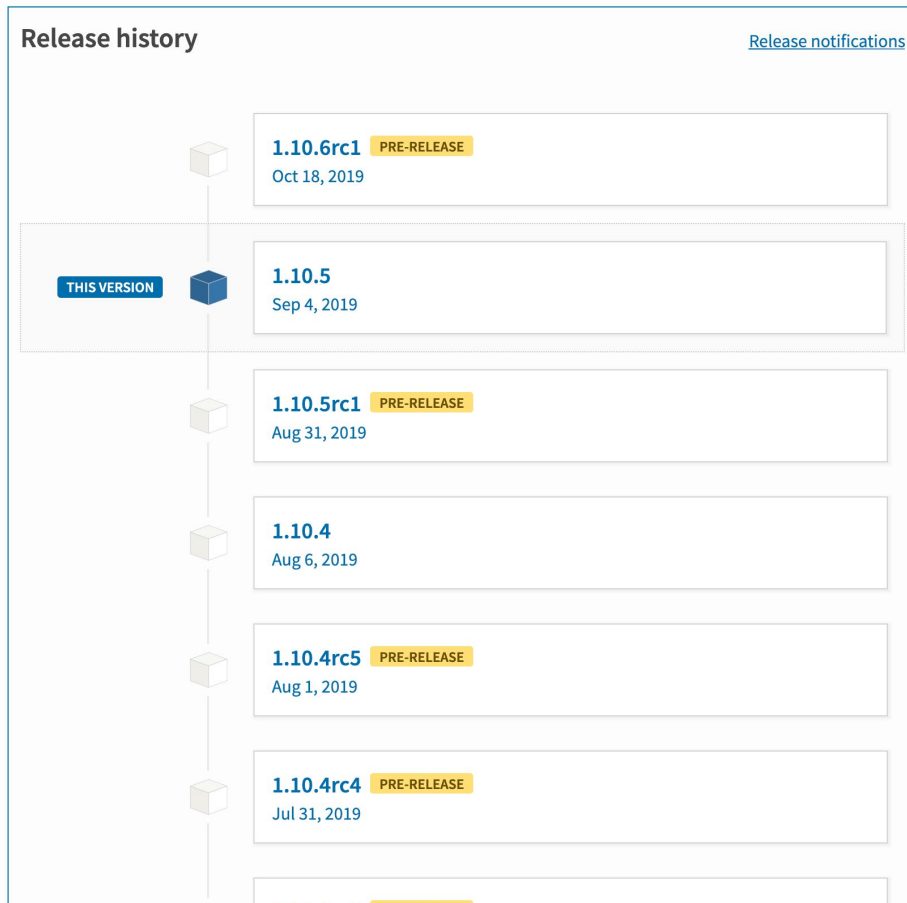
A screenshot of an Apache Airflow mailing list thread. The header shows the search bar with "Type search terms...", "Advanced" filter, and the current thread "dev@airflow.apache.org, last month". A "Show next 10" button is visible. The thread contains three messages:

- SCUSS] Change all example\_dags's start\_date to specific datetime** (2019-05-13 11:03): 3 people, 2 replies. The body text starts with "Airflower: I start this thread proposal to change all example\_dags's start\_date to specific datetime, including airflow/example\_dag and airflow/contrib..."
- SCUSS] Change all example\_dags's start\_date to specific datetime** (2019-05-13 10:57): 2 people, 2 replies. The body text starts with "Airflowers: I start this thread proposal to change all example\_dags's start\_date to specific datetime, including airflow/example\_dag and airflow/contri..."
- : Bay Area Airflow meetup** (2019-05-10 22:54): 4 people, 7 replies. The body text starts with "i've finalized the meetup logistics (details here <https://www.meetup.com/Bay-Area-Apache-Airflow-Incubating-Meetup/events/260712102/>). Pleas..."

# Releases



- How often?
- Apache Way?
- Non-binding opinions?
  - Are they listened to?
- Semantic Versioning?



# Similar Tools/Products

# Similar Tools/Products

Comparisons to draw

# Similar Tools/Products

Comparisons to draw

- Community vs. Company-driven

# Similar Tools/Products

Comparisons to draw

- Community vs. Company-driven
- Age of project

# Similar Tools/Products

Comparisons to draw

- Community vs. Company-driven
- Age of project
- Tooling used

# Similar Tools/Products



## Comparisons to draw

 [spotify / luigi](#)

 Used by ▾

822

 Watch ▾

507

 Star

12.4k

 Fork

2k

 Code

 Issues 42

 Pull requests 20

 Actions

 Projects 0

 Security

 Insights

Luigi is a Python module that helps you build complex pipelines of batch jobs. It handles dependency resolution, workflow management, visualization etc. It also comes with Hadoop support built in.

[python](#)

[luigi](#)

[orchestration-framework](#)

[scheduling](#)

[hadoop](#)



# Similar Tools/Products



## Comparisons to draw

spotify / luigi

Used by ▾

822

Watch ▾

507

★ Star

12.4k

Fork

2k

<> Code

! Issues 42

🔗 Pull requests 20

▶ Actions

📁 Projects 0

🛡 Security

📊 Insights

Luigi is a Python module that helps you build complex pipelines of batch jobs. It handles dependency resolution, workflow management, visualization etc. It also comes with Hadoop support built in.

python

luigi

orchestration-framework

scheduling

hadoop

# Similar Tools/Products



## Comparisons to draw

spotify / luigi

Used by ▾

822

Watch ▾

507

★ Star

12.4k

Fork

2k

<> Code

! Issues 42

🔗 Pull requests 20

▶ Actions

📁 Projects 0

🛡 Security

📊 Insights

Luigi is a Python module that helps you build complex pipelines of batch jobs. It handles dependency resolution, workflow management, visualization etc. It also comes with Hadoop support built in.

python

luigi

orchestration-framework

scheduling

hadoop

# Similar Tools/Products



## Comparisons to draw

spotify / luigi

Used by ▾

822

Watch ▾

507

★ Star

12.4k

Fork

2k

<> Code

Issues 42

Pull requests 20

Actions

Projects 0

Security

Insights

Luigi is a Python module that helps you build complex pipelines of batch jobs. It handles dependency resolution, workflow management, visualization etc. It also comes with Hadoop support built in.

python

luigi

orchestration-framework

scheduling

hadoop

# Similar Tools/Products



## Comparisons to draw

spotify / luigi

Used by ▾

822

Watch ▾

507

★ Star

12.4k

Fork

2k

<> Code

Issues 42

Pull requests 20

Actions

Projects 0

Security

Insights

Luigi is a Python module that helps you build complex pipelines of batch jobs. It handles dependency resolution, workflow management, visualization etc. It also comes with Hadoop support built in.

python

luigi

orchestration-framework

scheduling

hadoop

# What Do You Do Next?

- Save



# What Do You Do Next?

- Save
- Share kindly, if appropriate



# What Do You Do Next?

- Save
- Share kindly, if appropriate
- Do it again



TL;DR

- Take Your Time





TL;DR

- Take Your Time
- Do Your Homework



TL;DR

- Take Your Time
- Do Your Homework
- Give What You Can



Thank you!

