A practical guide to bring your project open source

Luciano Resende

June 13th, 2018

1
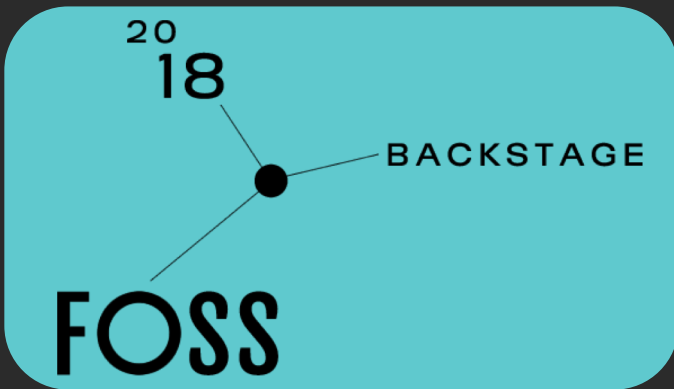
Agenda

Open Source Community Leadership

Center for Open Source Data & AI Technologies (CODAIT)

Items to consider when making your project open source

Proposing a project to Apache

Q&A

IBM
CODE

About me - Luciano Resende

Data Science Platform Architect – IBM – CODAIT
- Have been contributing to open source at ASF for over 10 years
- Currently contributing to : Jupyter Notebook ecosystem, Apache Bahir, Apache Toree, Apache Spark among other projects related to AI/ML platforms

lresende@apache.org

@lresende1975

https://github.com/lresende

https://www.linkedin.com/in/lresende

Open Source Community Leadership
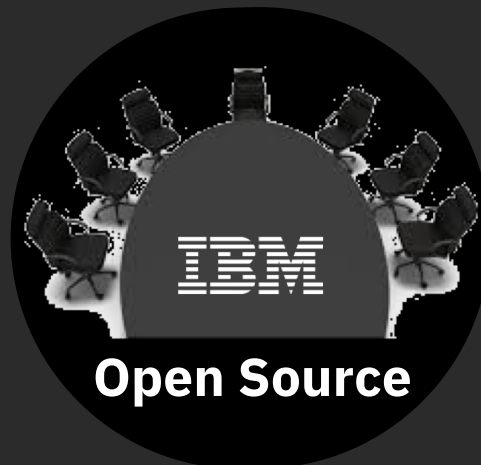

Founding Partner

188+ Project Committers

77+ Projects

Key Open source steering committee memberships

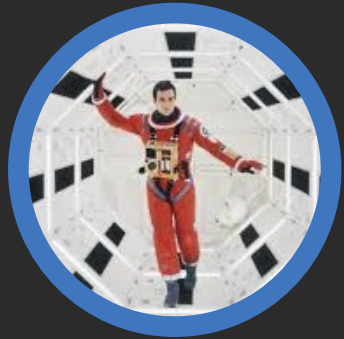**Open Source**

OSS Advisory Board

C O D A I T

IBM CODE

IBM's history of strong AI leadership

1968, 2001: A Space Odyssey

- IBM was a technical advisor

- HAL is "the latest in machine intelligence"

2011: Jeopardy!

- Watson beat two top Jeopardy! champions

1997: Deep Blue

- Deep Blue became the first machine to beat a world chess champion in tournament play

2018: Open Tech, AI & emerging standards

- New IBM centers of gravity for AI
- OS projects increasing exponentially
- Emerging global standards in AI

Center for Open Source Data and AI Technologies

CODAIT aims to make AI solutions dramatically easier to create, deploy, and manage in the enterprise

Relaunch of the Spark Technology Center (STC) to reflect expanded mission

CODAIT

codait.org

codait (French)

= coder/coded

https://m.interglot.com/fr/en/codait

IBM CODE

Why making your project open source ?

Assert eminence on particular area/technology
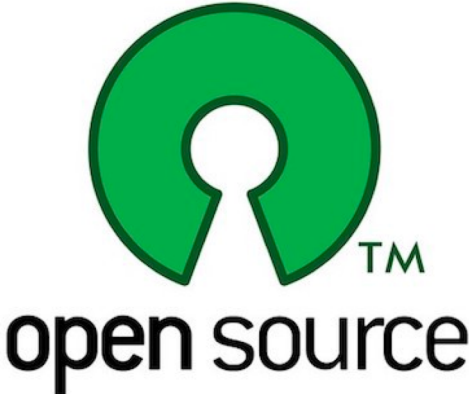
–Improve Personal/Organization brand recognition

Promote new Technology

–Build a community and extend the scope of the project

Open source becoming the "de facto" standard

–Apache Spark for distributed computing

IBM
CODE

Public Code is not Open Source

What makes your project open source?

When we call a project open source, most of the times we are implying that the project license complies with the <u>open source definition</u>—in brief, they allow software to be freely used, modified, and shared.

— https://opensource.org/osd

The <u>Open Source Initiative—OSI</u> has a list of <u>approved open source licenses</u>.

— https://opensource.org/licenses

— https://opensource.org/licenses/category
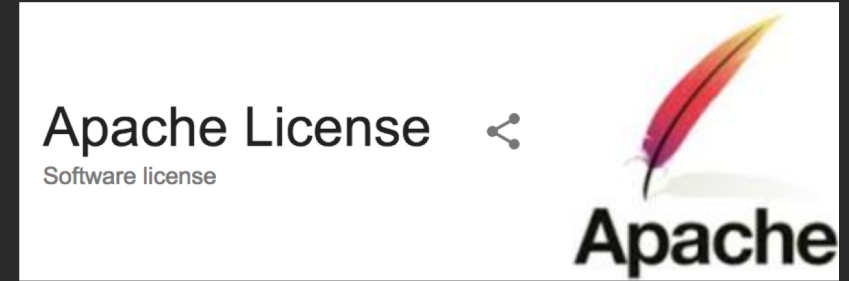
IBM
CODE

Evaluating an open source project

Project License

—The license provides a legal framework for collaboration

Project Governance Model

—The governance model provides a social framework for collaboration

Choose a Project License

Apache License
Software license
Apache

Mainly three license categories

– Permissive: The permissive licenses, as the name describe, are very permissive permitting that the software even become proprietary (e.g. Public Domain, MIT, BSD-new, Apache License 2.0).

– Strongly-protective or strong-copyleft: These licenses will not permit the software to become proprietary (e.g. GPL, GPL2, GPL3, AGPL3)

– Weakly-protective or weak-copyleft: These licenses are a compromise between the two categories described above, restricting the software to become proprietary but enabling it to be part of a larger proprietary work.

IBM CODE

Applying the Apache License

—Add license to a given project by creating a `LICENSE` file with the contents of the actual license.

—Add License headers to each file that accept comments


```
Copyright [yyyy] [name of copyright owner]

Licensed under the Apache License, Version 2.0 (the "License");
you may not use this file except in compliance with the License.
You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License.
```

IBM
CODE

Is that all ?

Dependencies and License Compatibility

Image by David A. Wheeler—https://www.dwheeler.com/essays/floss-license-slide.html
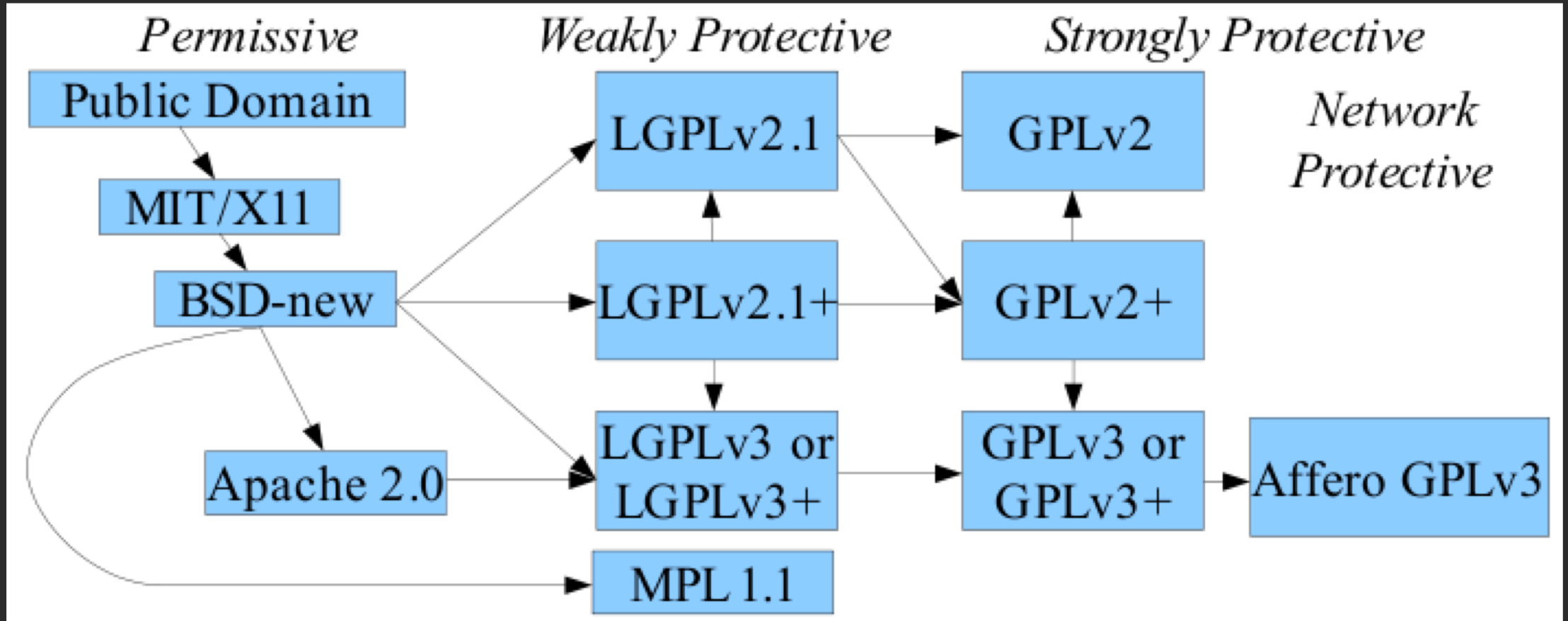
Accepting Contributions

For each contribution, assure

– Contribution precedence

– Permission to package and redistribute according to project license

Apache CCLA/ICLA

– https://www.apache.org/licenses/icla.pdf

Developer Certificate of Origin

– https://developercertificate.org/
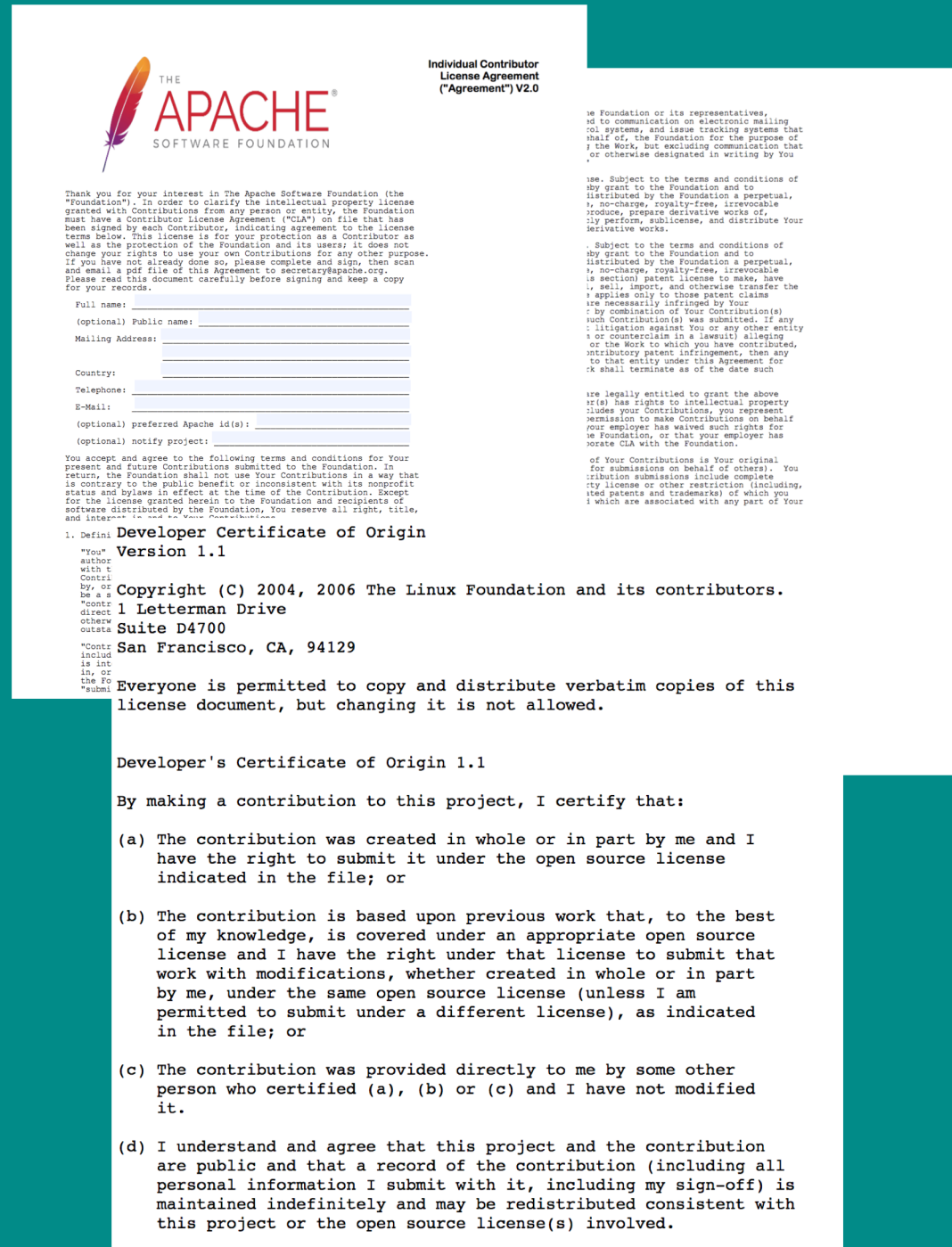
Building a Community

Attract users

– Getting started guides

– Documentation

– Tests

Attract contributors

– Be welcoming, and mentor new members

– Have a list of 'help wanted' or 'beginners' task

IBM
CODE

Documentation

Suggested topics

– Build and Installation steps

– Contribution guidelines

– API Documentation

Tools

– GitHub pages/wiki

Tools for Python

– Read the Docs - https://readthedocs.org/

Apache Software Foundation

Why Apache?

Apache License 2.0

– Permissive license with great enterprise acceptance/adoption

Meritocratic Governance Model

– Governance model is well understood and accepted by enterprises

– Individuals are recognized based on their contributions

  • Committers, PMC members, Foundation Members, etc

Branding

IBM
CODE

Apache Incubator

The Apache Incubator is the entry path for projects and codebase wishing to become part of the Apache Software Foundation

- Ensure all donations are in accordance with the ASF legal standards

- Develop new communities that adhere to our guiding principles

Incubator PMC provides oversight for new projects on the following areas :

- Responsible for the acceptance and oversight of new incubator projects

- Ensures that the projects develop products according to the Foundation's philosophy and guidelines for collaborative development

- Legal – ensures donation are in accordance with legal standards

- Community – helps develops communities that adhere to guiding principles

- Regularly evaluates products to determine if the product should be abandoned, continue to receive guidance and support, or proposed to the board for promotion to full project status as part of an existing or new Foundation PMC;

Apache Incubator – Project proposal

Champion

– A candidate project shall be sponsored by an Officer or Member of the Foundation. The Champion assists the candidate on their initial submission to a Sponsor.

Mentors

– Mentors are chosen by the Sponsor to actively monitor the podling, guide the podling in the Apache Way, and report its status to the Sponsor and the Incubator PMC.

– All Mentors must be members of the Incubator PMC.

Apache Incubator – Best practices

Have a detailed description of your project

Highlight dependency on other Apache projects
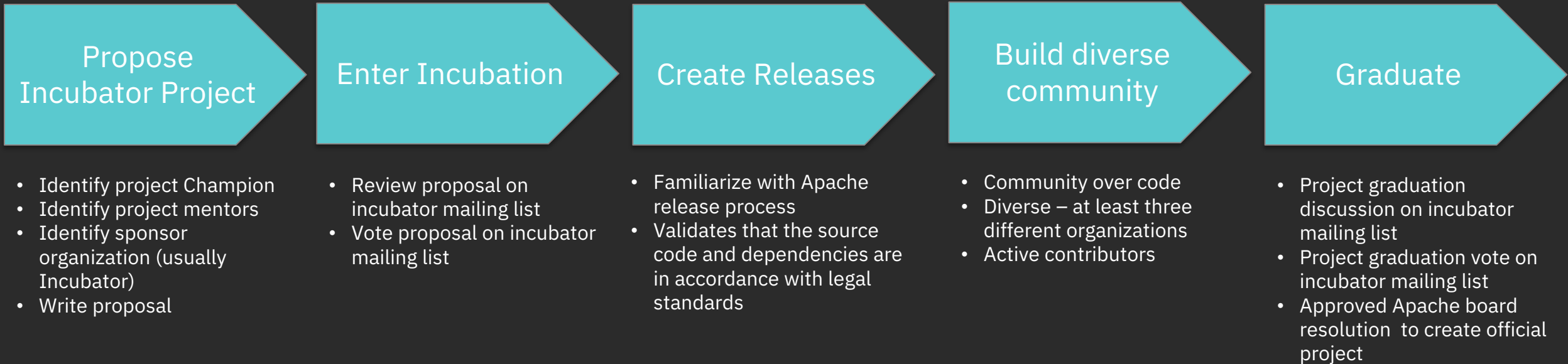
– Or potential areas of collaboration

Identify any possible dependency risks

– E.g. dependencies that are incompatible with AL 2.0

Have a diverse set of initial committers

– Partner with external members to create the proposal

Apache Incubator – Project Lifecycle

Propose Incubator Project

- Identify project Champion
- Identify project mentors
- Identify sponsor organization (usually Incubator)
- Write proposal

Enter Incubation

- Review proposal on incubator mailing list
- Vote proposal on incubator mailing list

Create Releases

- Familiarize with Apache release process
- Validates that the source code and dependencies are in accordance with legal standards

Build diverse community

- Community over code
- Diverse – at least three different organizations
- Active contributors

Graduate

- Project graduation discussion on incubator mailing list
- Project graduation vote on incubator mailing list
- Approved Apache board resolution to create official project

Summary

Public code is not open source

It's all about License (and governance)

Community over Code

And you now know how to bringing your project to Apache

IBM
CODE