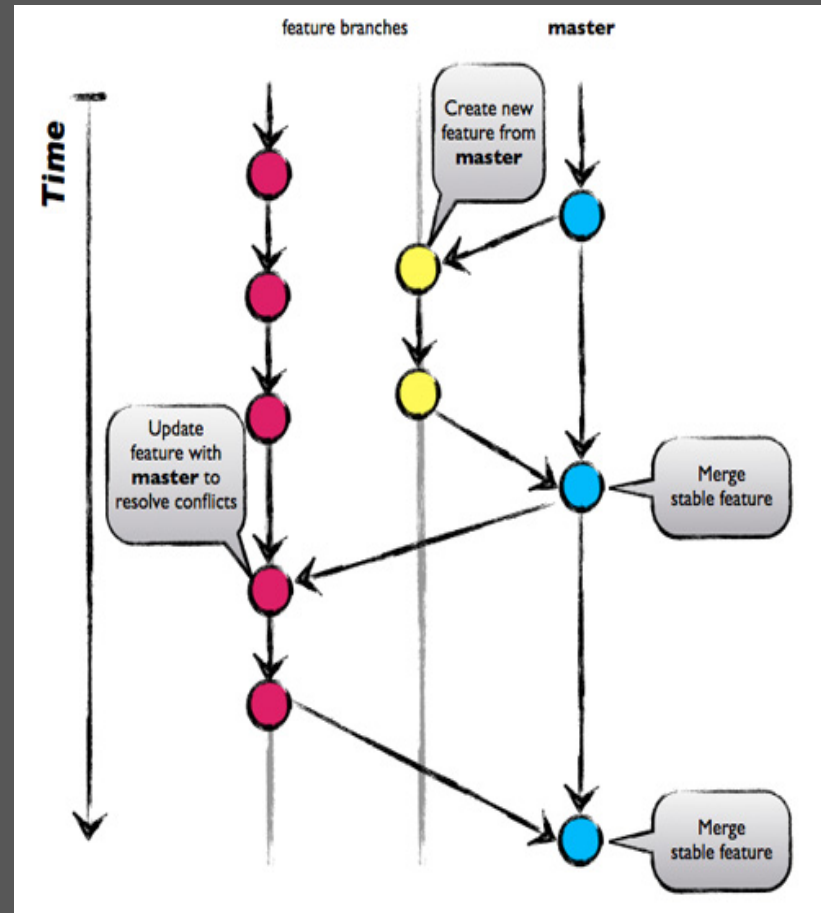


# Best Practices for Review of Contributions to Open Source Projects



July 2017



Open Source Overview

Development Process / Governance

Licenses

Community Engagement

# Open Source Overview

- Public agencies moving to an open source approach.
- [government.github.com](https://government.github.com)
- Department of Defense, NASA, Consumer Financial Protection Bureau, National Park Service



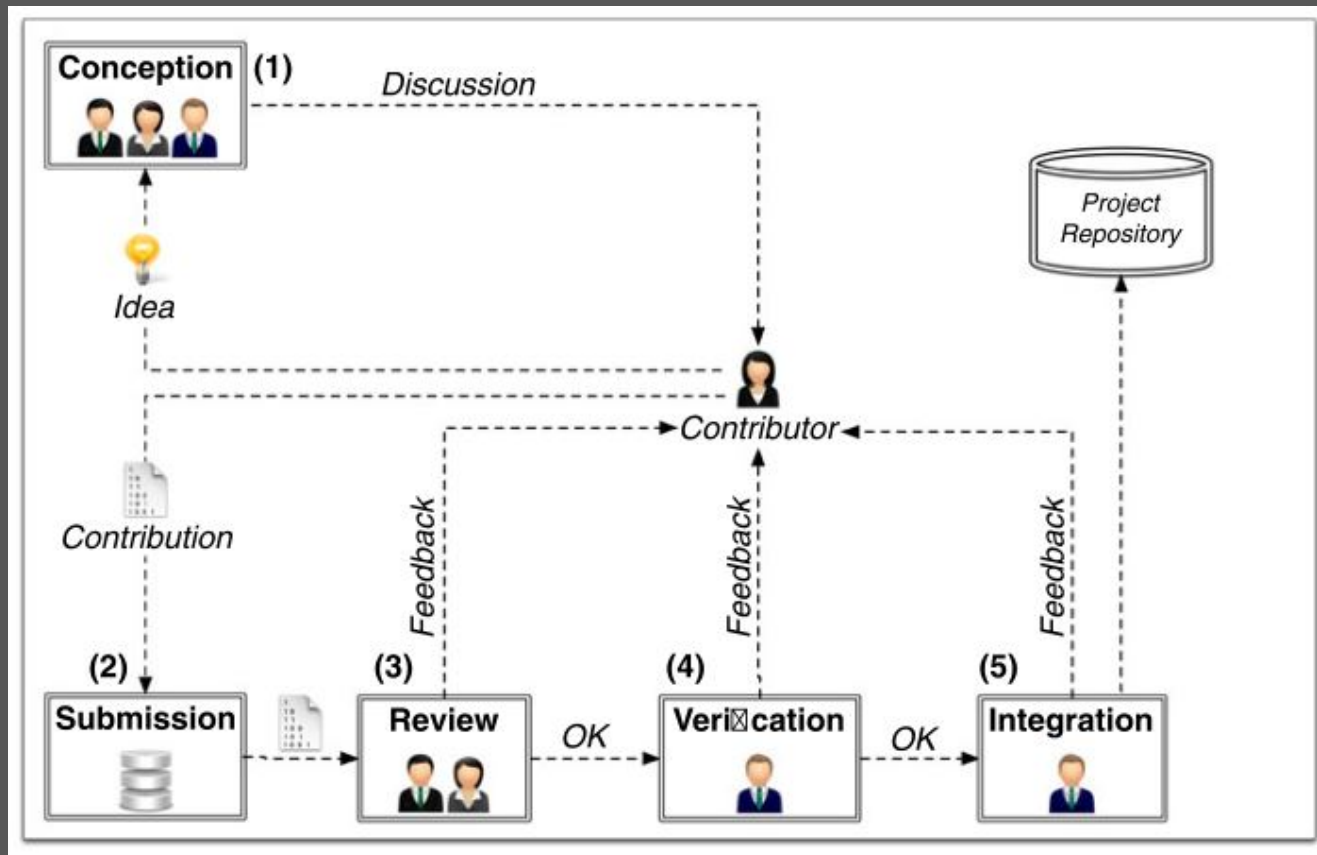
# Open Source Overview

Ideal for projects supported by public agencies.

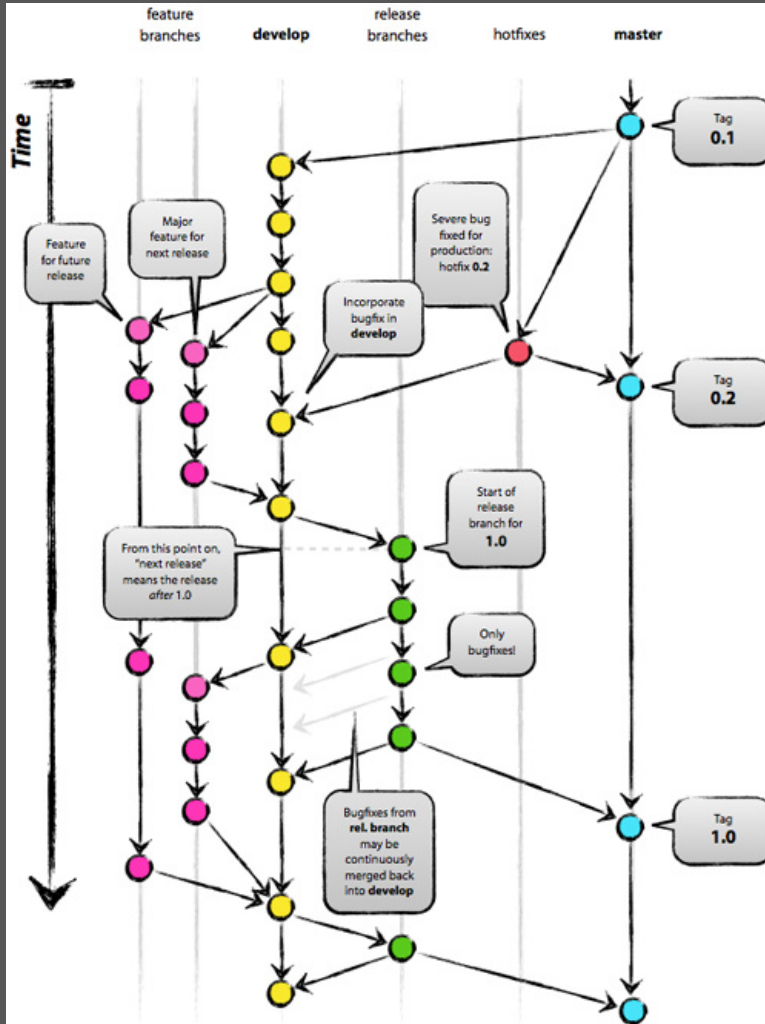
- **Transparency:** code readily available to the community of users.
- **Cost-effectiveness:** builds on common tools, share modules
- **Longevity:** community effort, not tied to single developer

# Development Process

Tied to governance, since need to clearly specify authority for review of code



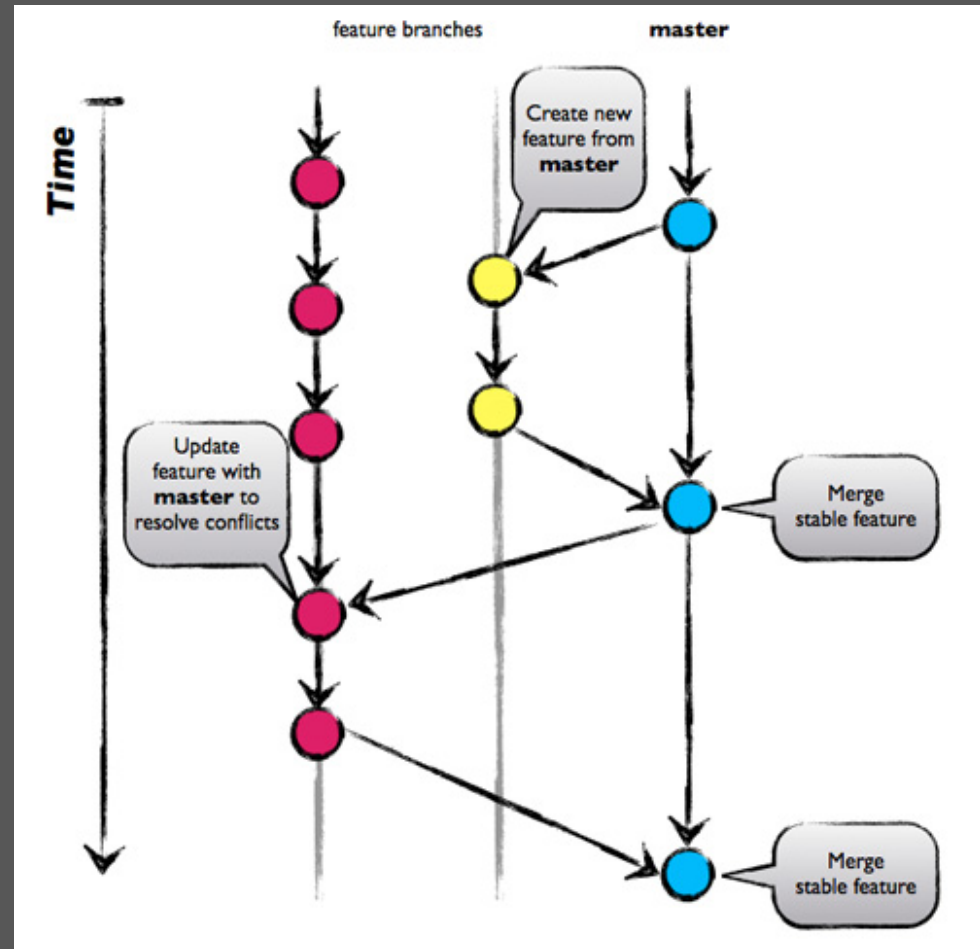
# Development Process



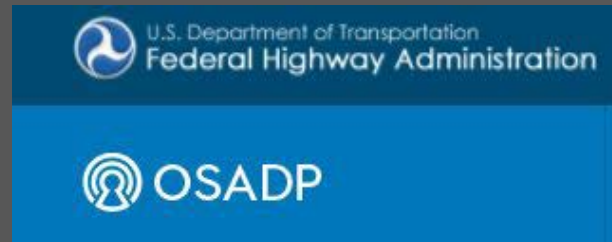
- Stable master
- Active development branch with frequent commits
- Periodic releases

# GitHub for Code Review

- **Issues** to initiate discussion of new features
- **Pull Requests** to propose code for review. Iterative process, merge to dev.
- **Wiki** can help organize priority features for development



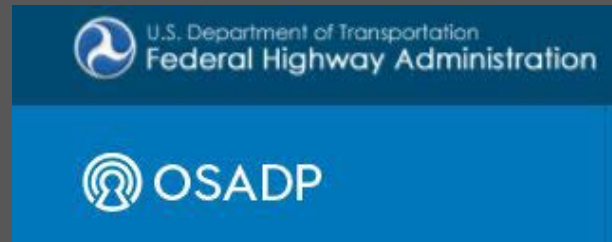
# Code Review Example: OSADP



- Funded by USDOT Intelligent Transportation Systems – Dynamic Mobility Applications program
- **Upload** access request form. Review by OSADP administrator
- **Staging** to private GitHub repository



# Code Review Example: OSADP



- **Review** by OSADP staff, following release checklist
- **Public Release** on itsforge.net
- **Collaboration** only by designated project members on private GitHub repository

# Code Review Example: OSADP

- **Release Checklist (partial)**
  - Has the source code been verified and deemed functional and stable?
  - Is the source code technically reviewed and inspected to ensure that no malicious code is embedded?
  - Have you verify to ensure no executable files or personal information are included?
  - Is the source code scanned for virus and no infected file found?
  - Is complete open source code base or assets (code, files, images, tables, data...) included?
  - Is a LICENSE.txt file included with a open source license terms and conditions such as Apache 2.0?

# Code Review Example: Node.js



- Very large project, >8,000 pull requests.
- **Users** use the JavaScript tool, may add Issues
- **Collaborators** review code, rights to commit and merge pull requests. Designated by CTC
- **Core Technical Committee (CTC)** meet weekly, consensus seeking. Membership by CTC nomination.

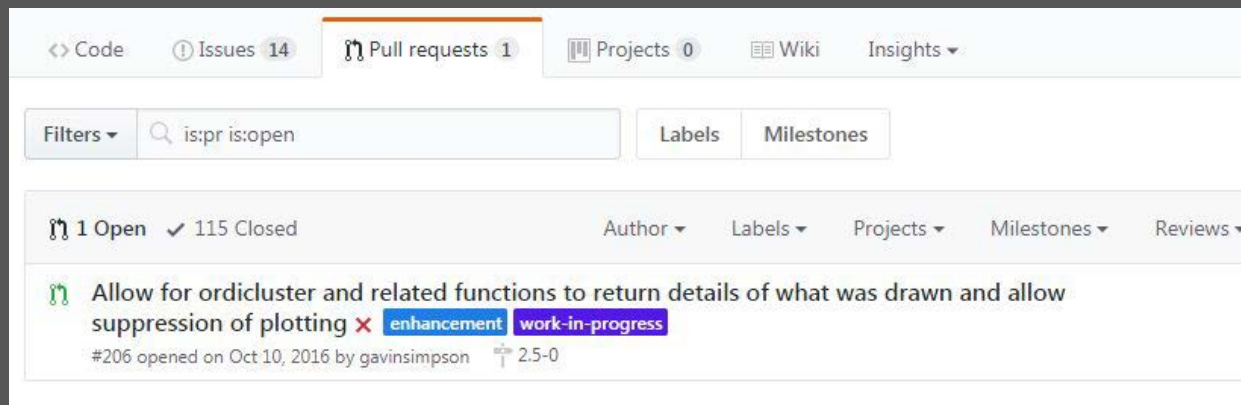
# Code Review Example: Node.js



- Development originally within one company (Joyent)
- **Node.js Foundation** sponsors some of the time of the Technical Committee, but Collaborators not funded
- Collaborators may be volunteer or given company time to support the open source project

# Code Review Example: vegan

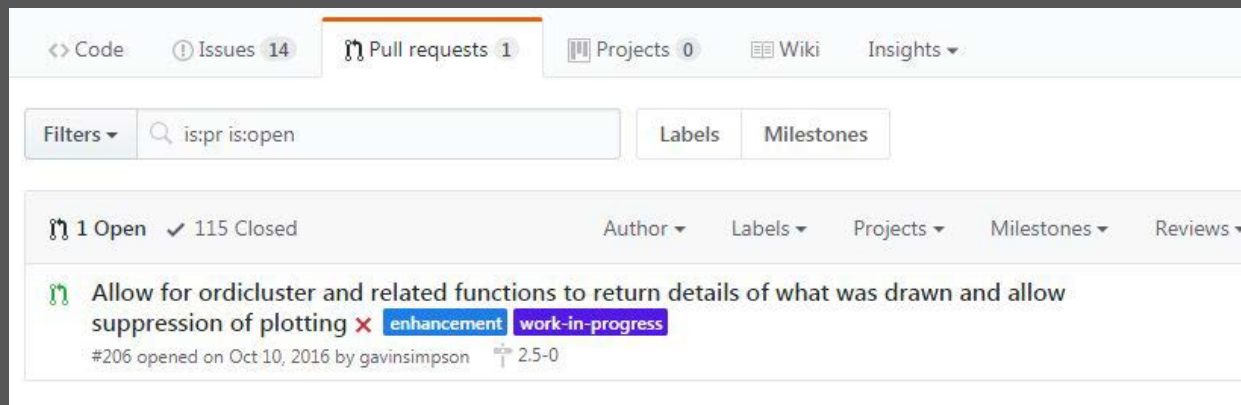
- R package with small group of developers, relatively large number of users (800 weekly downloads)
- No formal governance structure, no explicit code review process
- Users engage via Issues, R listserv, Stackoverflow



The screenshot shows the GitHub interface for the 'vegan' repository. At the top, there are navigation tabs for Code, Issues (14), Pull requests (1), Projects (0), Wiki, and Insights. Below this is a search bar with the query 'is:pr is:open' and buttons for Filters, Labels, and Milestones. The main content area shows a list of pull requests, with the first one selected. This pull request is titled 'Allow for ordcluster and related functions to return details of what was drawn and allow suppression of plotting' and is labeled as an 'enhancement' and 'work-in-progress'. It was opened on Oct 10, 2016, by gavinsimpson, and is associated with version 2.5-0.

# Code Review Example: vegan

- Developers are academics, working on personal time
- Strictly volunteer code review process



The screenshot shows the GitHub interface for a pull request in the 'vegan' repository. The navigation bar at the top includes links for Code, Issues (14), Pull requests (1), Projects (0), Wiki, and Insights. Below the navigation bar, there is a search bar with the query 'is:pr is:open' and buttons for Filters, Labels, and Milestones. The main content area shows a list of pull requests, with the first one selected. This pull request is titled 'Allow for ordcluster and related functions to return details of what was drawn and allow suppression of plotting' and is labeled as an 'enhancement' and 'work-in-progress'. It was opened on Oct 10, 2016, by gavinsimpson, and has a target version of 2.5-0.

# Code Review Example: [ActivitySim](#)

- **Users** use the modeling tool
- **Developers** contribute code / documentation
- **Committers** review code, designated by PMC
- **Project Management Committee (PMC)** composed of funders, provide oversight. PMC Contractor crucial.
- **Funders** may be part of PMC

# Licenses

## Choose an open source license

Which of the following best describes your situation?



### I want it simple and permissive.

The **MIT License** is a permissive license that is short and to the point. It lets people do anything they want with your code as long as they provide attribution back to you and don't hold you liable.

**jQuery**, **.NET Core**, and **Rails** use the MIT License.



### I'm concerned about patents.

The **Apache License 2.0** is a permissive license similar to the MIT License, but also provides an express grant of patent rights from contributors to users.

**Android**, **Apache**, and **Swift** use the Apache License 2.0.



### I care about sharing improvements.

The **GNU GPLv3** is a copyleft license that requires anyone who distributes your code or a derivative work to make the source available under the same terms, and also provides an express grant of patent rights from contributors to users.

**Bash**, **GIMP**, and **Privacy Badger** use the GNU GPLv3.



# Contributor License Agreements (CLA)

- “Intellectual Property Assignment Agreement”
- States you are entitled to contribute code/documentation, willing for it to be used in distributions and derivative work
- Cannot withdraw permission
- Requires recordkeeping and maintenance of agreements
- May be more burdensome than needed for VisionEval



**CLAHub**  
**Contributor License Agreements  
on GitHub**

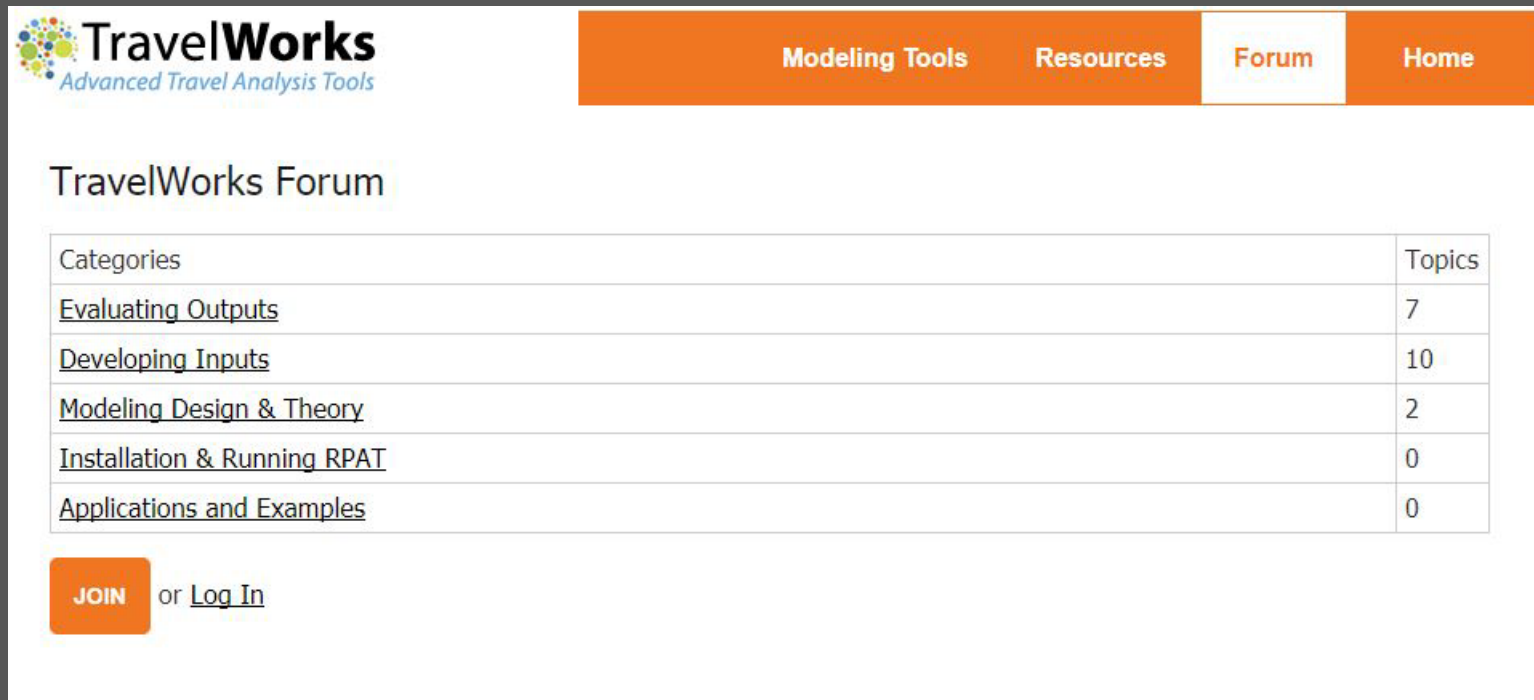
Easily add a [Contributor License Agreement](#) to your GitHub project, let contributors sign it, and see at a glance which pull request authors have signed your CLA, and which haven't.

# Summary

- **Engage** the community with clear wiki, active issues page, consider discussion board
- Develop **governance structure** which establishes the roles and responsibilities of users, developers, and collaborators/committers, as well as an oversight body composed of funders.
- Develop **code review** process which is clear, fair, and public
- Clearly **document** tools, provide guidance for developers
- Agree on a **license** which meets the needs of the VisionEval

# Community Engagement

## Discussion forums



**TravelWorks**  
Advanced Travel Analysis Tools

Modeling Tools Resources **Forum** Home

### TravelWorks Forum

Categories	Topics
<a href="#">Evaluating Outputs</a>	7
<a href="#">Developing Inputs</a>	10
<a href="#">Modeling Design &amp; Theory</a>	2
<a href="#">Installation &amp; Running RPAT</a>	0
<a href="#">Applications and Examples</a>	0

[JOIN](#) or [Log In](#)

# Community Engagement

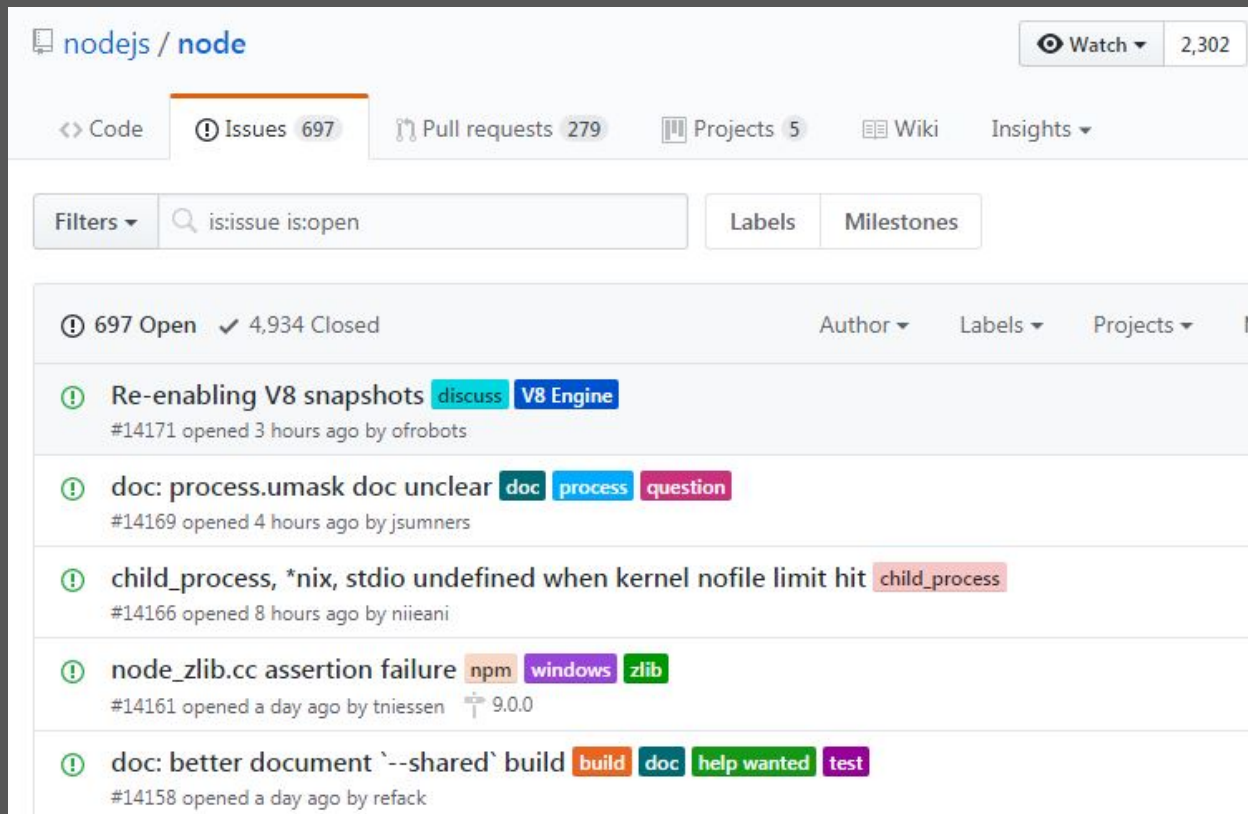
Listservs, developer and user

The screenshot shows a Google Groups interface. At the top, there is a search bar with the text "Search for messages". Below this, the "Groups" section is visible, with a navigation menu on the left containing "My groups", "Home", "Starred", "Favorites", and "Recently viewed". A yellow tooltip is present over the "Favorites" section, stating "Click on a group's star icon to add it to your favorites". The main content area displays the "Stan users" group, with a sub-header "35 of many topics". There are navigation tabs for "all categories", "all tags", "Latest" (which is selected), "Top", and "Categories". A "general" tab is also visible. The message list is organized into columns: "Topic", "Category", and "Users".

Topic	Category	Users
<input type="checkbox"/> We're M By Mich		
<input type="checkbox"/> rstan a By Ben		
<input type="checkbox"/> bayesp By Jona		
<input type="checkbox"/> shinyst By Jona		
<input type="checkbox"/> Guideli By Jona		
<input type="checkbox"/> new_CppObject_xp no valid constructor in 2.16	RStan	[User A] [User B]
<input type="checkbox"/> Von_mises documentation suggestion	General	[User A] [User B] [User C]
<input type="checkbox"/> Sampler error in rstan on Mac that doesn't reproduce on other computers	General	[User S] [User B] [User C]
<input type="checkbox"/> Schema for callback writers: one step towards retiring CSV	Developers	[User A] [User B] [User C] [User S]
<input checked="" type="checkbox"/> Periodic Gaussian Process - covariance matrix not positive semidefinite	General	[User A] [User B] [User C]
<input type="checkbox"/> Installation problems on Mac--no bin folder created	General	[User F] [User B] [User R]

# Community Engagement

## Issues page



nodejs / node Watch 2,302

[Code](#) **Issues** 697 [Pull requests](#) 279 [Projects](#) 5 [Wiki](#) [Insights](#)

Filters  [Labels](#) [Milestones](#)

697 Open ✓ 4,934 Closed Author Labels Projects M

- Re-enabling V8 snapshots [discuss](#) [V8 Engine](#)  
#14171 opened 3 hours ago by ofrobots
- doc: process.umask doc unclear [doc](#) [process](#) [question](#)  
#14169 opened 4 hours ago by jsumners
- child\_process, \*nix, stdio undefined when kernel nofile limit hit [child\\_process](#)  
#14166 opened 8 hours ago by niieani
- node\_zlib.cc assertion failure [npm](#) [windows](#) [zlib](#)  
#14161 opened a day ago by tniessen [9.0.0](#)
- doc: better document `--shared` build [build](#) [doc](#) [help wanted](#) [test](#)  
#14158 opened a day ago by refack

# To VisionEval

