



ClearlyDefined

Community Meeting - April 6, 2026



# Agenda

- Open Source Summit NA 2026
- Roadmap 2026/2027/2028



# Open Source Summit NA 2026

- [Beyond SBOMs: Making License Data Actionable With ClearlyDefined - Jamie Magee, Microsoft](#)





## Roadmap (2026, 2027, 2028)

- Proposed [roadmap](#) prepared by AboutCode
- 3 year plan with specific yearly milestones and deliverables

### Goals:

- Lowering the overall cost of operations
- Increasing reach, efficiency and accessibility of CD data and code



## Roadmap (2026, 2027, 2028)

### Key points:

- Combine ClearlyDefined and AboutCode operations and backends when possible to enable scan data reuse and avoid duplicated, wasted compute resources across ecosystems
- Refactor ClearlyDefined backend and frontend to reduce the technical debts and increase its reach
- Enroll new sponsors to share the cloud compute costs
- Unlock data access for increased community adoption
- Enable new use cases around security and project health



# Roadmap (2026, 2027, 2028)

## Themes:

- 2026
  - Reduce technical debt and operating costs
  - Improve the curation and harvesting queue
  - Improve the UI to facilitate community contributions
- 2027
  - Migrate to a federated data architecture to further reduce hosting costs and unlock data access.
  - Initiate work on new use cases, vulnerabilities and project health and risks
- 2028
  - Deploy security and project health use cases



# Roadmap (2026, 2027, 2028)

## 1) Create better license and package scans, faster and more efficiently

- Enable reusing scans across ClearlyDefined harvests and AboutCode data collections
- Design a priority mechanism to scan packages based on actual usage
- Improve the performance of ScanCode for faster and more cost efficient scans
- Improve ScanCode license detection based from the analysis of ClearlyDefined curated data
- Adopt federated data design with data mirrored across multiple hosts to improve service availability and reduce costs, reusing existing scans



# Roadmap (2026, 2027, 2028)

## 2) Reduce technical debt

- Update dependencies to latest versions
- Refactor storage of scans and attachments
- Switch to the open source DocumentDB from the proprietary MongoDB
- Maintain and improve the ClearlyDefined code base, CI/CD, and infrastructure
- Update project documentation to reflect new and ongoing development
- Facilitate outreach and onboarding for new contributors and sponsors



# Roadmap (2026, 2027, 2028)

## 3) Shared scans for shared costs

- Expand the pool of organizations hosting scans to share the scanning costs
- Design new prioritized scan queue to scan packages in an order that aligns with business needs
- Integrate distributed scanning with federated data access with on-premises, private data usage
- This also includes direct collaboration with open source foundations, and large projects such as Linux distributions to help them scan and share scans of their upstream and downstream packages.



## Roadmap (2026, 2027, 2028)

### 4) Integrate PURL and SBOM for easier interoperability

- Create new API endpoints to access ClearlyDefined data keyed by PURL
- Create a new SBOM API endpoint to expose reference package data in SBOM formats
- Adopt PURL as a core identifier in the database and file store



## Roadmap (2026, 2027, 2028)

### 5) Improve curation processes and user experience

- Develop and deploy a new and improved UI for efficient data curation
- Make it easier for organization to contribute their curations back in bulk
- Improve the data models for curation storage
- Create a new specification for multi-stakeholder data curations exchange



# Roadmap (2026, 2027, 2028)

## 6) Support for more use cases

- Security with vulnerability data
- Data for project health and lifecycle events and analytics
- Usage to find the most used open source components and drive scan priorities



ClearlyDefined

Thank you