Centre for Health Informatics

Andy Rae

22 October 2025

Discovery

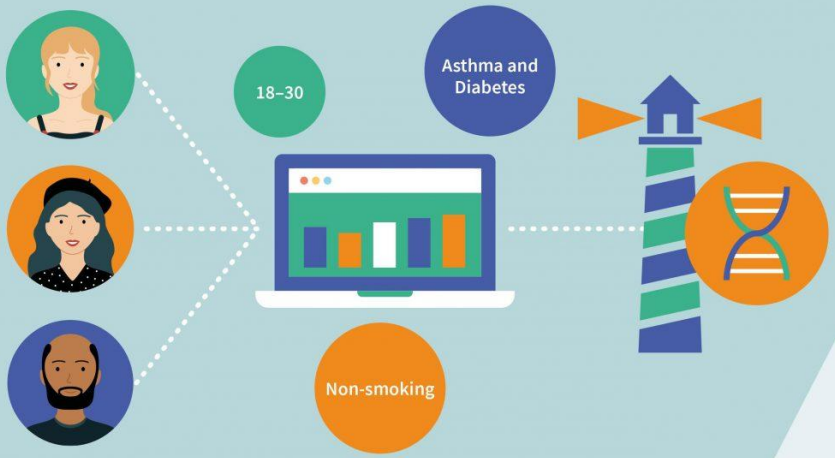
Cohort Discovery: meeting specific data needs

The Innovation Gateway's Cohort Discovery search tool allows researchers to search, assess and request access to potential datasets that precisely match information on particular populations of interest – for example data across multiple datasets relating to non-smoking patients aged 18 to 30 that have asthma and diabetes.

Health Data Research Innovation Gateway

HDRUK Health Data Research UK

The tool enables multiple dataset searches to be done in a fast, secure, de-identified and ethical way. At no point is there any individual level data transferred from the custodian.



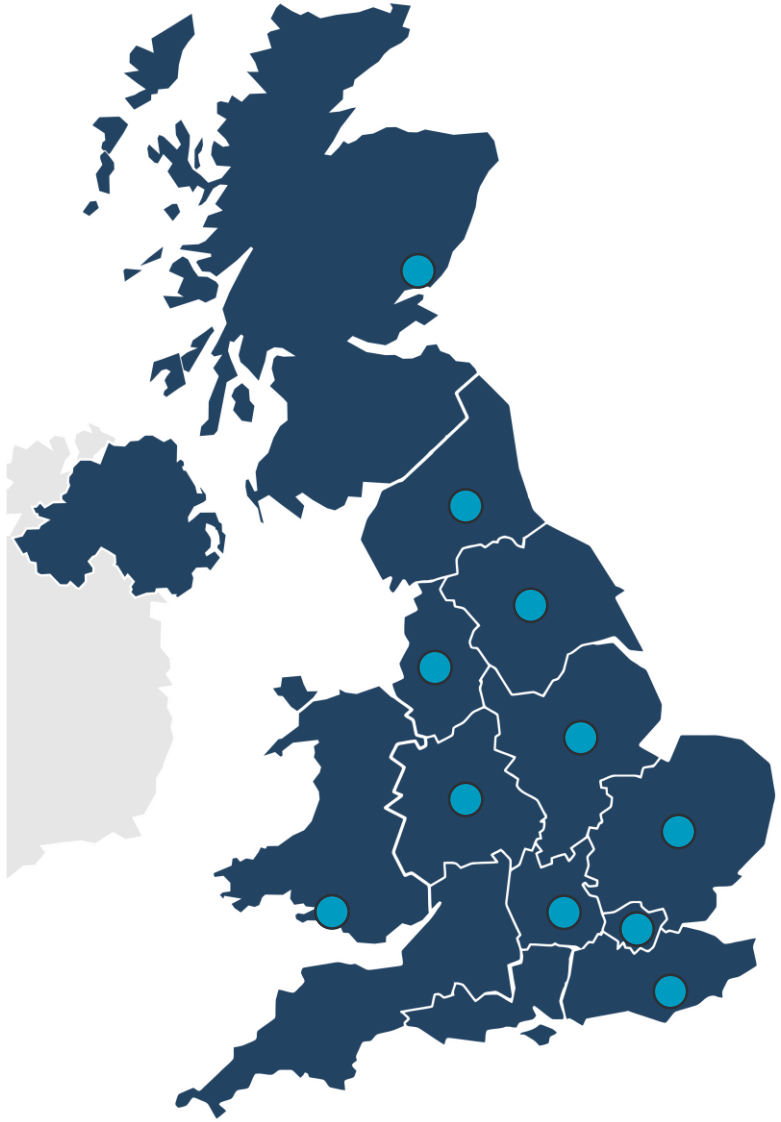
How can Cohort Discovery help you?

Gateway users, including researchers, can carry out more specific searches and assessments on datasets listed in the Gateway, so that potential datasets which match the requirements of a research project are easier to find.

For data custodians, Cohort Discovery means requests to access data are more likely to result in success, whilst still retaining control over who has access to the data.

HDR UK has extensively tested the cybersecurity of the Cohort Discovery tool, and had this independently verified, providing confidence in its privacy and security.

Discovery



CPRD

NHS

BC Platforms

Research Secure Data Environment NETWORK

Bunny

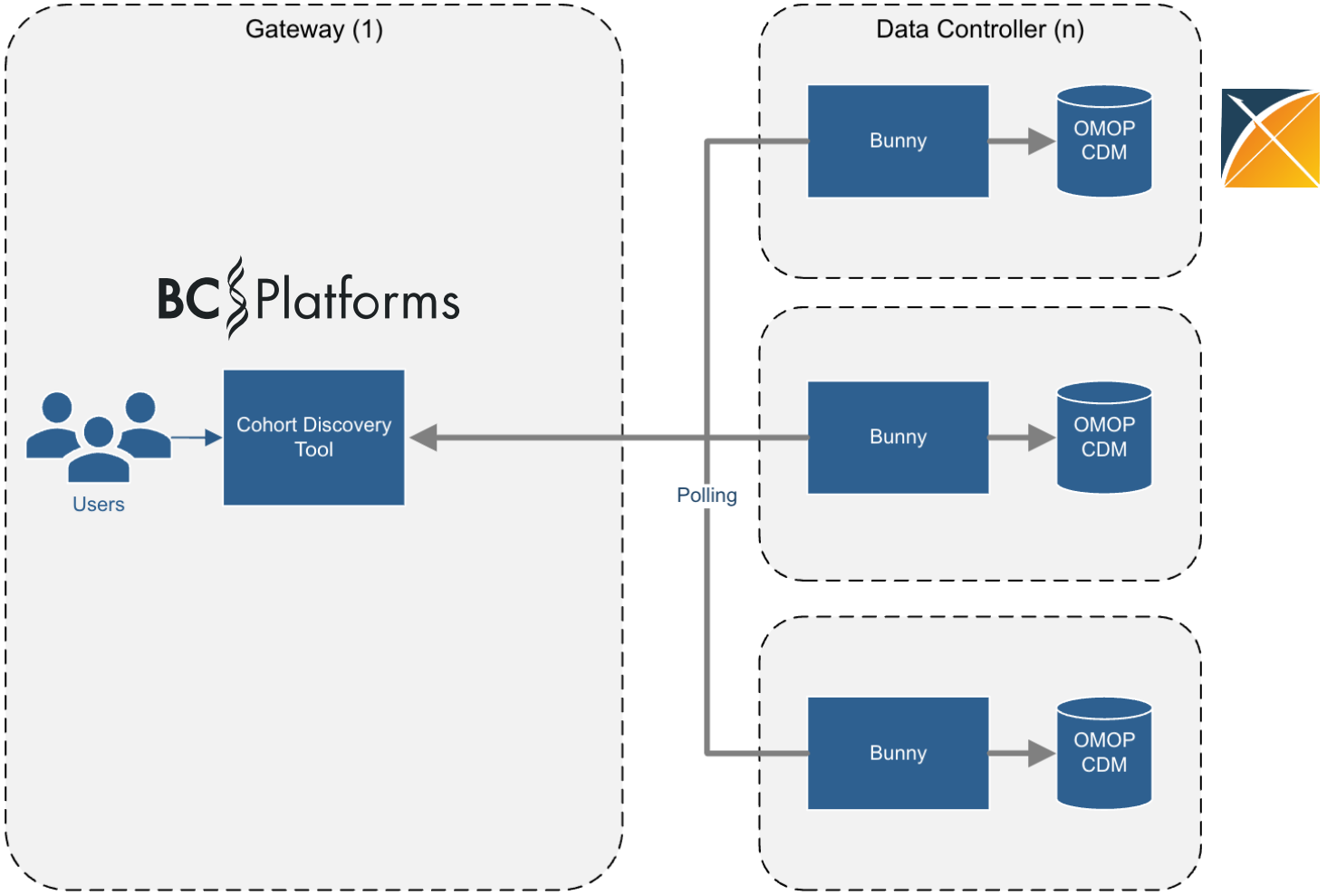
HDRUK

Health Data Research UK

Discovery

HDRUK
Health Data Research UK

Bunny

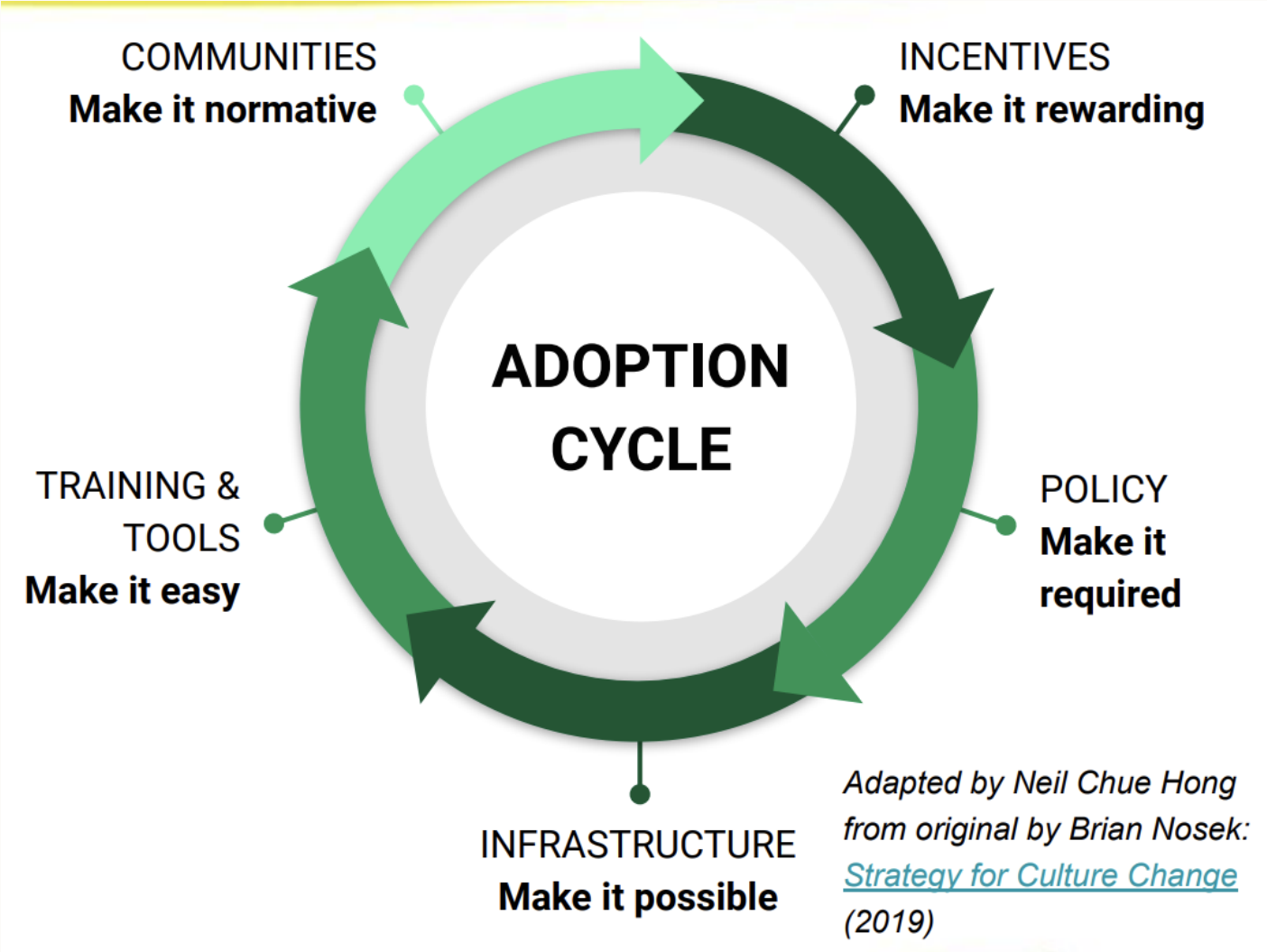


Discovery

- Secure by Design
- Fast Query Resolution
- Built in Disclosure Control
- MIT Open Source
- Python Microservice

Bunny

Adoption Cycle



Security Model

Security overview

Security policy • **Enabled**
View how to securely report security vulnerabilities for this repository — View security policy

Security advisories • **Enabled**
View or disclose security advisories for this repository — View security advisories

Private vulnerability reporting • **Enabled**
Allow users to privately report potential security vulnerabilities — See reported vulnerabilities

Dependabot alerts • **Enabled**
Get notified when one of your dependencies has a vulnerability — View Dependabot alerts

Code scanning alerts • **Enabled**
Automatically detect common vulnerability and coding errors — View alerts

Secret scanning alerts • **Enabled**
Get notified when a secret is pushed to this repository — View detected secrets

Test Suite

- Continuous Integration
- Unit, integration, end to end
- Meaningful coverage

https://hutch.health/bunny/deployment/security

Coverage 88%

▼ Coverage Report

File	Stmts	Miss	Cover	Missing
src/hutch_bunny				
cli.py	34	3	91%	26, 32–33
daemon.py	18	18	0%	1–32
src/hutch_bunny/core				
db.py	32	12	62%	30–35, 44–56, 75–77
db_manager.py	103	24	77%	96, 110, 121, 132, 241–245, 249–251, 282–292, 295–300, 303–306, 309
execute_query.py	31	7	77%	50, 54, 65–71
results_modifiers.py	8	1	88%	36
src/hutch_bunny/core/rquest_dto				
activity_job.py	12	12	0%	1–36
base_dto.py	6	2	67%	3, 7
cohort.py	14	1	93%	17
group.py	14	1	93%	19
query.py	38	4	89%	32, 79, 109, 116
rule.py	50	17	66%	41–52, 88–99
src/hutch_bunny/core/solvers				
availability_solver.py	184	5	97%	273, 379, 469–474
query_solvers.py	149	5	97%	36, 367, 409–411
src/hutch_bunny/core/upstream				
task_handler.py	13	13	0%	1–41
TOTAL	**1068**	**125**	**88%**	

Image Publishing

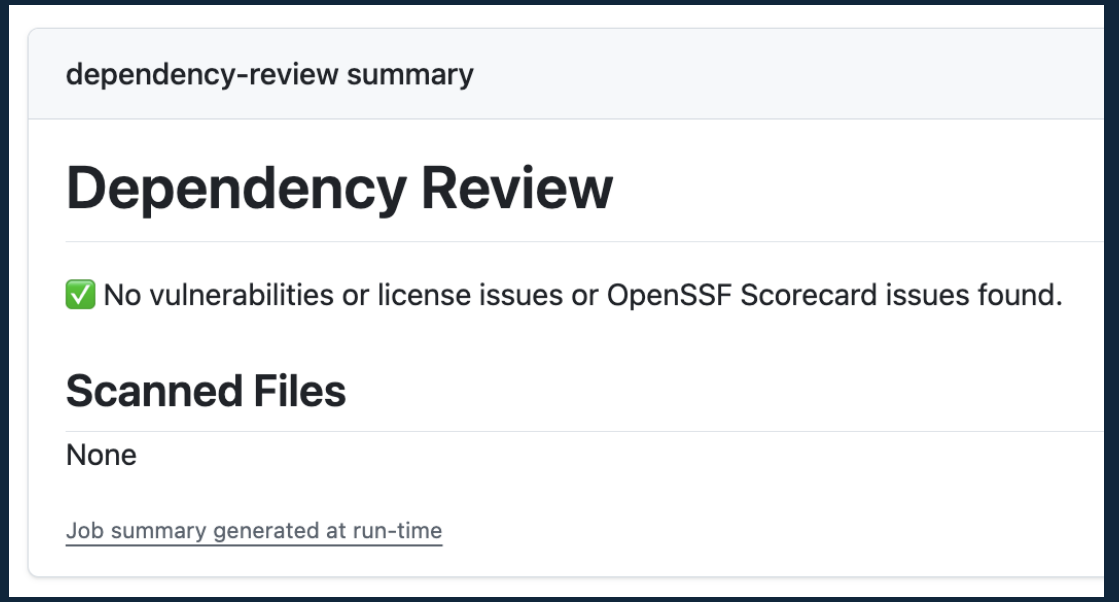
- Github Actions
- Github Container Registry
- Pinned workflow dependencies
- Pinned image hashes

https://hutch.health/bunny/deployment/security

```yaml
jobs:
  publish-bunny:
    runs-on: ubuntu-latest
    permissions:
      packages: write
    steps:
      - name: Check out the repo
        uses: actions/checkout@v4

      - name: Set up QEMU
        uses: docker/setup-qemu-action@29109295f81e9208d7d86ff1c6c12d2833863392 # v3.6.0

      - name: Set up Docker Buildx
        uses: docker/setup-buildx-action@f95db51fddba0c2d1ec667646a06c2ce06100226 # v3.0.0

      - name: Docker Login
        uses: docker/login-action@343f7c4344506bcbf9b4de18042ae17996df046d # v3.0.0
        with:
          registry: ${{ env.registry }}
          username: ${{ github.actor }}
          password: ${{ secrets.GITHUB_TOKEN }}

      - uses: SebRollen/toml-action@b1b3628f55fc3a28208d4203ada8b737e9687876 # v1.2.0
        id: read_version
        with:
          file: pyproject.toml
          field: project.version

      - name: Docker Metadata action
        id: meta
        uses: docker/metadata-action@902fa8ec7d6ecbf8d84d538b9b233a880e428804 # v5.7.0
        env:
          DOCKER_METADATA_ANNOTATIONS_LEVELS: manifest,index
        with:
          images: ${{ env.registry }}/${{ env.repo-owner }}/${{ env.image-name }}
          # Tag notes:
          # - RFC3339 is not suitable for docker tags, so we squash the date
          # - We tag both the short (7-char prefixed) and full sha commit hashes; both are useful
          # - `edge` represents latest main branch commit (potentially unstable)
          tags: |
            type=sha
            ${{ github.sha }}
            type=raw,value={{date 'YYYYMMDDHHmmss[Z]'}}
            edge
```

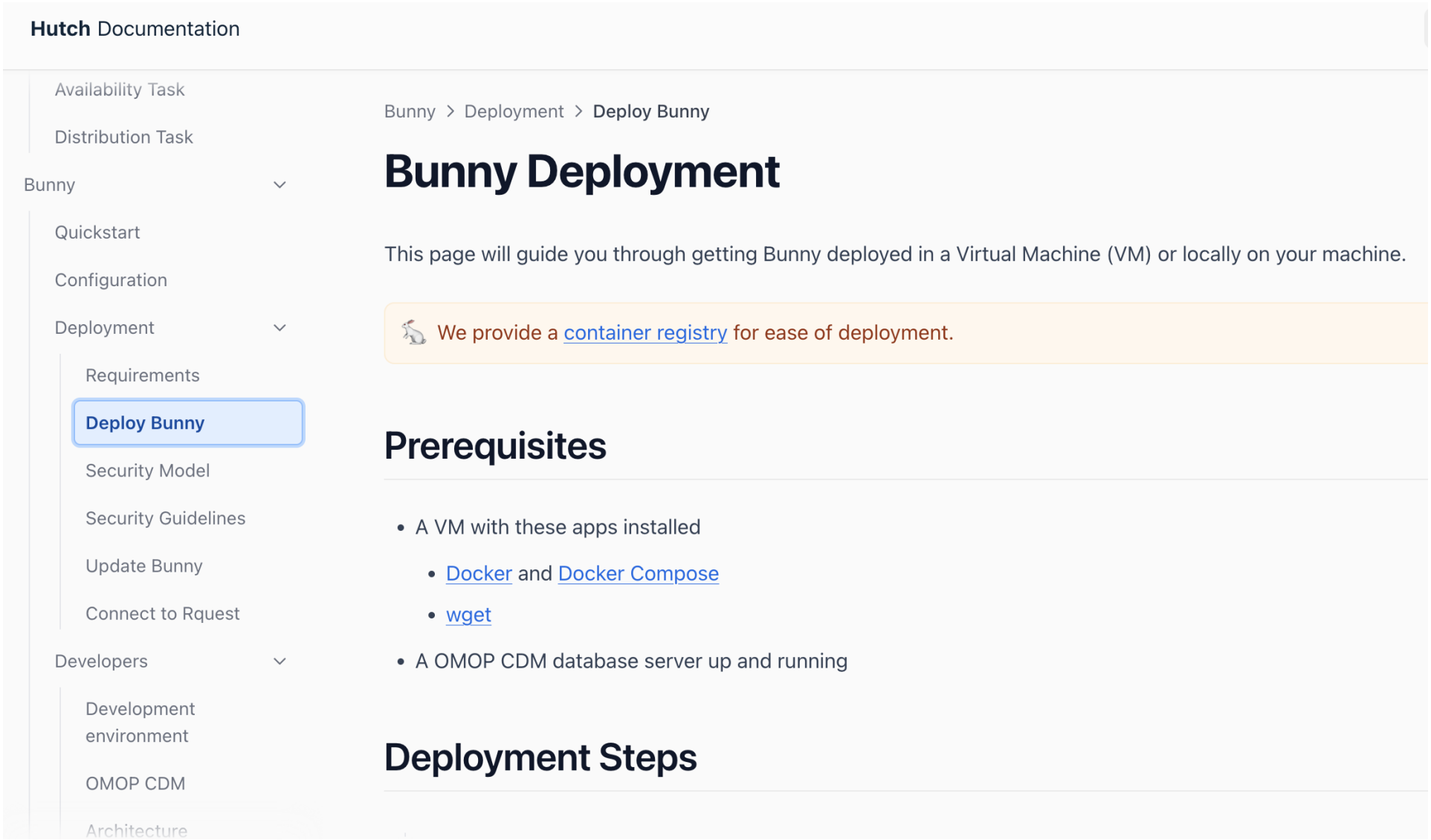
Supply Chain

- Github Actions
- dependency-review-action
- Dependabot
- Software Bill of Materials (SBOM)



https://hutch.health/bunny/deployment/security

Documentation



https://hutch.health/contributing

Contributions Policy

Thanks for your interest in contributing! We welcome contributions of all kinds, including:

- Bug reports and new issues
- Code improvements and new features
- Documentation updates and fixes

Submitting Contributions

To contribute:

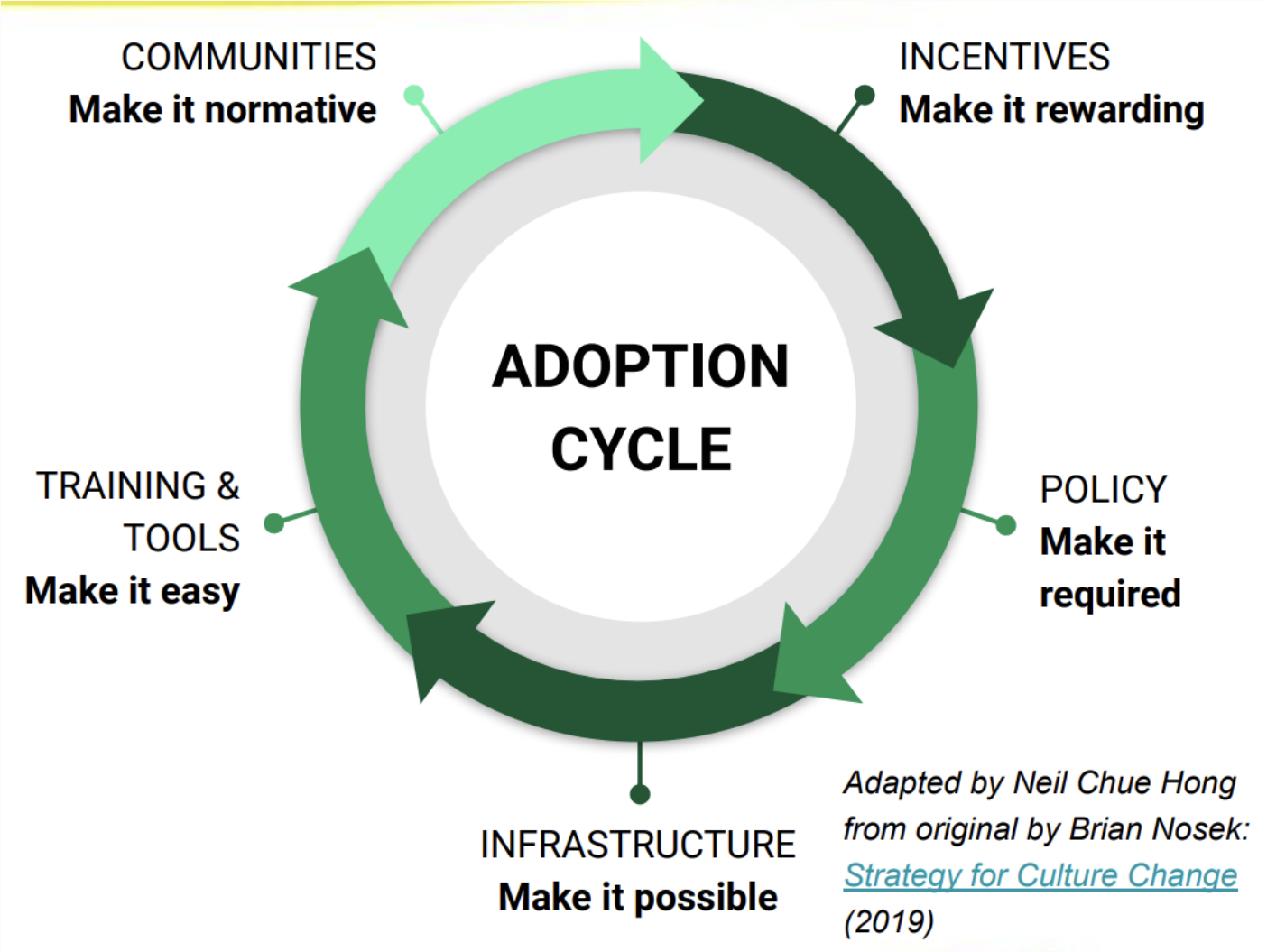
- Check open issues, and create or comment on an issue to volunteer to work on it.
- A maintainer will then assign the issue to you if it is available and appropriate to work on.
- Fork the repository and create a branch for your changes.
- Follow existing coding patterns and keep contributions focussed.
- Open a pull request (PR) with a clear description of your changes.

Copyright & Licensing

By contributing, you agree that your contributions are licensed under the same MIT License as the project. This means your work can be freely used, modified, and distributed under the same terms.

https://hutch.health/contributing

Adoption Cycle



University of Nottingham
UK | CHINA | MALAYSIA

Federated Research

Discovery

HDRUK
Health Data Research UK

DARE UK