

# Building scientific communities - Lessons learned with AIRR

Christian E. Busse<sup>[1,2,3]</sup>

[1] Div. of B Cell Immunology, German Cancer Research Center

[2] ORCID: [0000-0001-7553-905X](https://orcid.org/0000-0001-7553-905X)

[3] Github: [bussec](https://github.com/bussec)

# Outline

- **What:** AIRR data and the AIRR Community
- **How:** Beneficial patterns for community building
- **Why:** Interaction RSE ↔ domain-specific communities

# What is AIRR?

- AIRR = Adaptive Immune Receptor Repertoire<sup>[1]</sup>
  - Next-generation sequencing data describing the state of the adaptive immune system
  - Neither classical genome nor transcriptome data
  - Meaningful interpretation requires rich metadata

# The AIRR Community

- Grass-roots initiative, started 2015
- Community of practice: "To foster and support the generation, sharing and use of FAIR AIRR data"
- Vision of a set of shared resource, the AIRR \* Commons
- Current & future products
  - MiAIRR<sup>[1]</sup>: Minimal metadata standard for AIRR data
  - AIRR Standardized Representation<sup>[2]</sup>: Format for analysis results
  - IARC/OGRDB<sup>[3]</sup>: Review and submission system for inferred allele
  - REST API for AIRR repositories (*under development*)



[1] Rubelt F *et al.*, Nat Immunol 18:1274 (2017) [\[DOI\]](#)

[2] Vander Heiden JA *et al.*, Front Immunol 9:2206 (2018) [\[DOI\]](#)

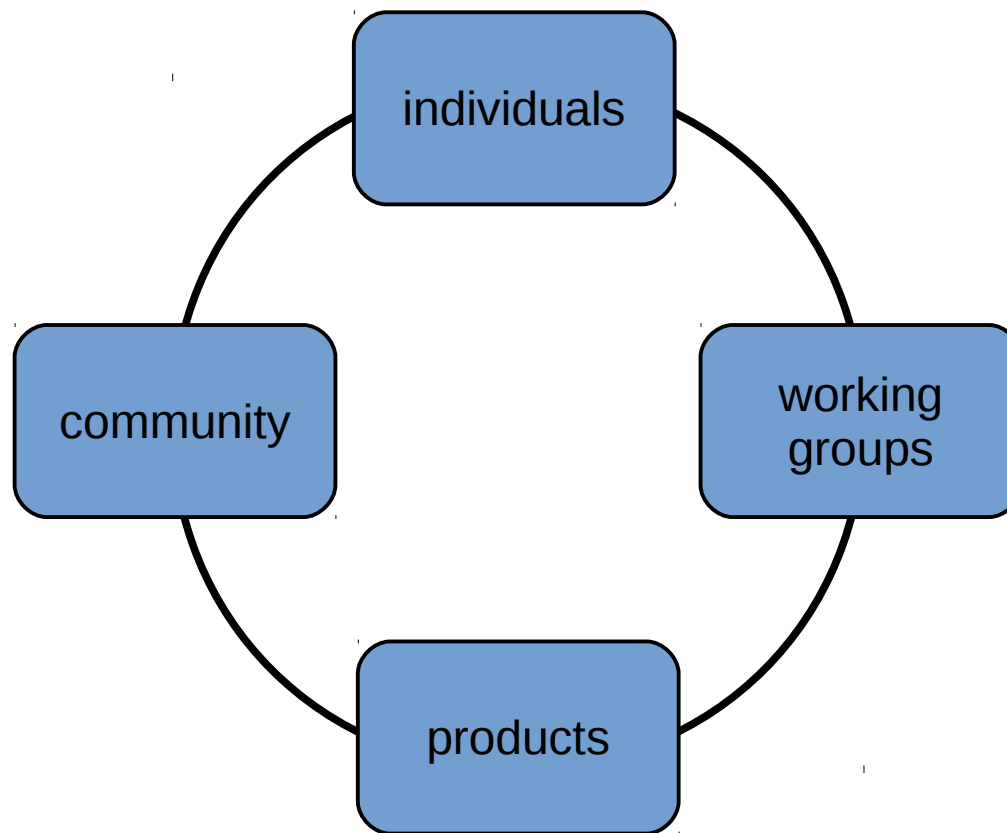
[3] Ohlin M *et al.*, Front Immunol 10:435 (2019) [\[DOI\]](#)

# Disclaimer



- This is a retrospective and  $n=1$  assessment
- We were young and naive...
- There are names for some of the presented patterns, they were deliberately omitted

# Entities



## Beneficial patterns: Individuals

- Do all the work, but mostly volunteers → “What is in it for me?”
- Community should:
  - Provide platforms for professional exchange
  - Provide opportunities for (horizontal) collaboration
  - Foster collaborations beyond the scope of the community

## **Beneficial patterns: Working groups**

- Have high level of autonomy (within mandate)
- Define attainable goals (within ~ 12 months time span)
- Have 5-8 active members
- Work agreement-based
- Communicate frequently with other working groups
- Document decisions and activities



## Beneficial patterns: Products

- Are an offer to the scientific domain to reach goals (e.g. FAIR AIRR data)
- Scratch someone's itch (in a generic fashion)
- Need to be good, not perfect
- Evolve over time, need maintenance
- Fit into a bigger picture → “Ecosystem” of tools and resources
- Are accepted (and potential endorsed) by the community

## **Beneficial patterns: Community**

- Be an open community → low threshold of participation
- Define shared values and a common vision
- Identify problems of common interest
- Focus on products first, do organization later
- Approach stakeholders early (but not too early)
- Commit to maintain products → sustainability

## **Not-so-beneficial patterns**

- We are all just volunteers, we can all just walk away
- The community
  - must not interfere with working group autonomy
  - should never require unanimous decisions
- Do not waste your time on a community journal

## Why should RSEs care?

- Definition of RSE as software engineer with domain expertise
- Domain-specific communities can support the cause of RSE
- RSE should try to understand domain-specific practices and demands

# Acknowledgements



<https://github.com/airr-community/>

<https://www.ireceptor-plus.com>