

# *Linking Publications to SDG's: An Area Based Approach*

Ed Noyons & Ismael Rafols, CWTS, Leiden University

Mapping Research related to the Sustainable Development Goals (SDGs)

September 4th, 2020

# Linking communities to SDGs rather than individual publications

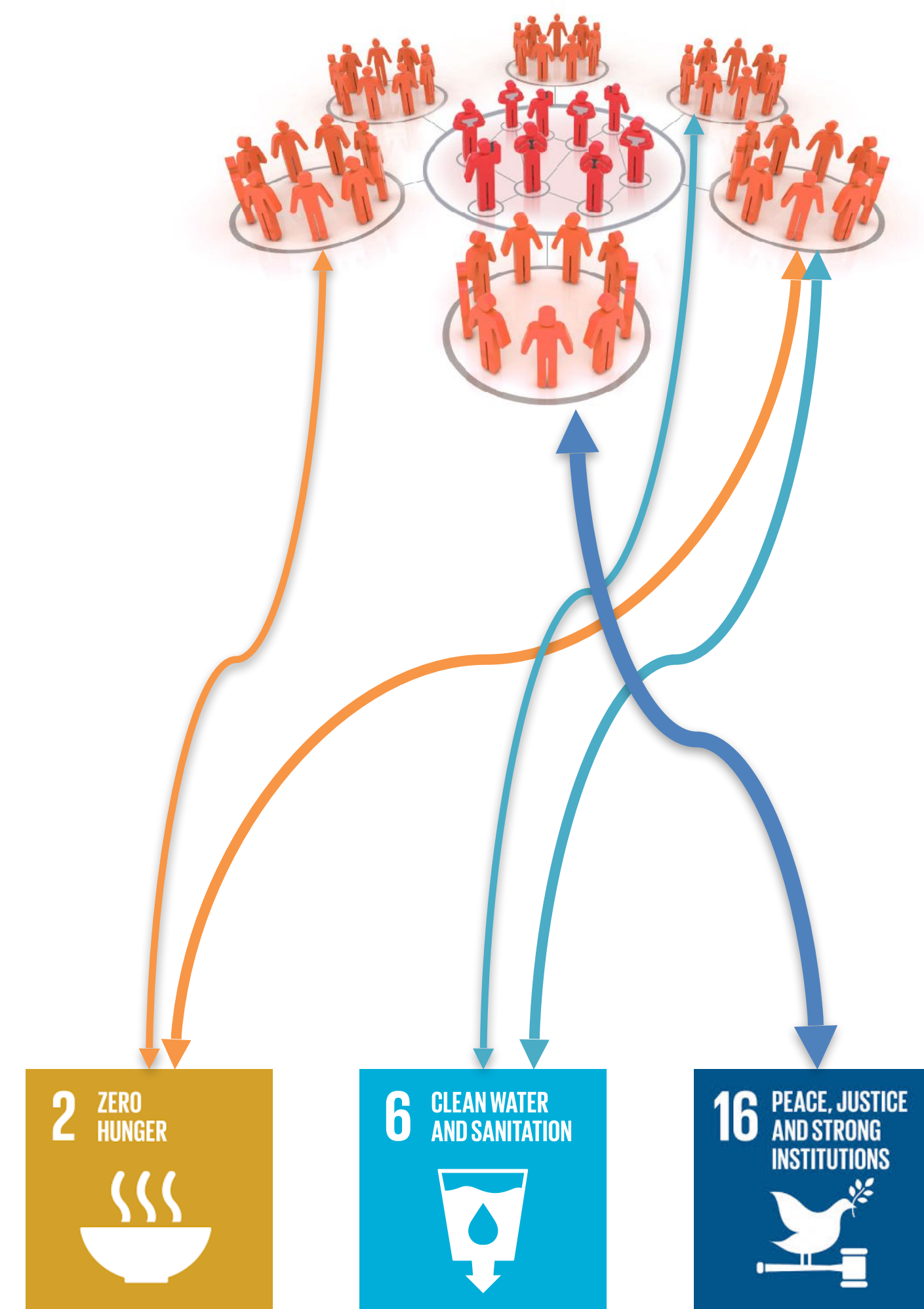
# Outline

- Approach
- Application/ results
- Implications/ opportunities/ conclusions

# Approach

## Key approach

- We identify communities (research areas)
- We link communities to SDGs (e.g., via keywords)
- We characterise the link:
  - Strength
  - Type



# Why not individual papers?

- Keyword approach
- Policy goals vs scientific goals
- Policy language vs scientific language
- Long term impact





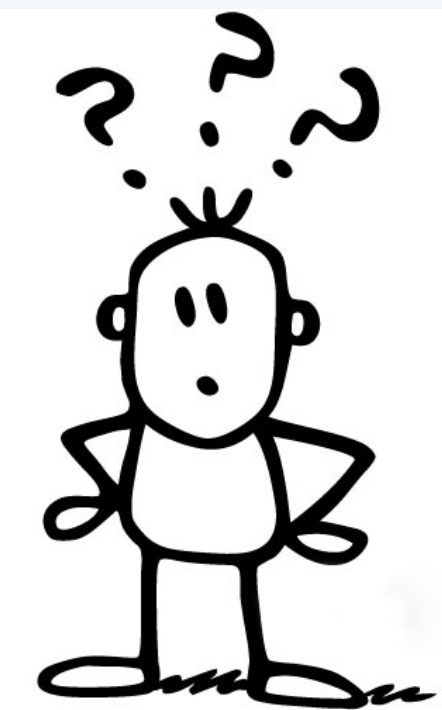
# Why communities?



- Science is a collaborative effort;
- Connection/ relation is more stable;
- There may be a reciprocal relation;
- Less dependent on current jargon;
- Easier to assess relevance, verify, validate;
- Using community/area characteristics.



# What communities?





# Communities within the WoS science landscape (publication based classification, 4000 communities/ clusters/ areas)

Social Sci & Hum.

Maths & CompSci

Life & Earth

Physical Sci & Engin.

Biomedical & Health



## Communities or research areas

- Clusters (bags of publications) created by direct citation relations
- No journal information/ classification used

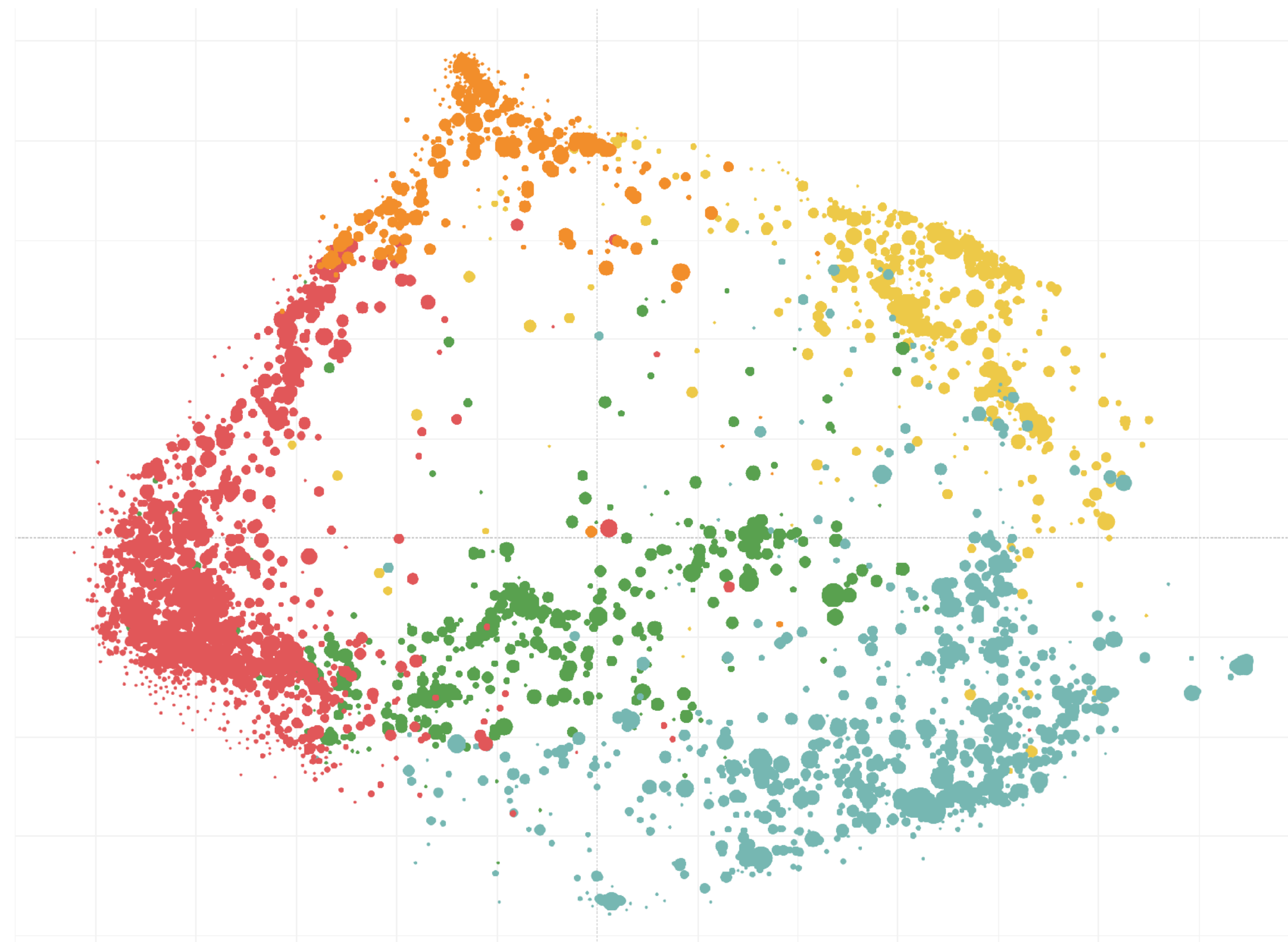
# Steps

- Select keywords from relevant (policy) documents
- Collect publications using these keywords (seeds)
- Distribute this collection (seeds) across publication-level classification (self-organized structure of science)
- Create links based on this info between SDGs and classification
  - Set thresholds
  - Validate/ characterise relationship

# Application

# Map of all science (4000 communities)

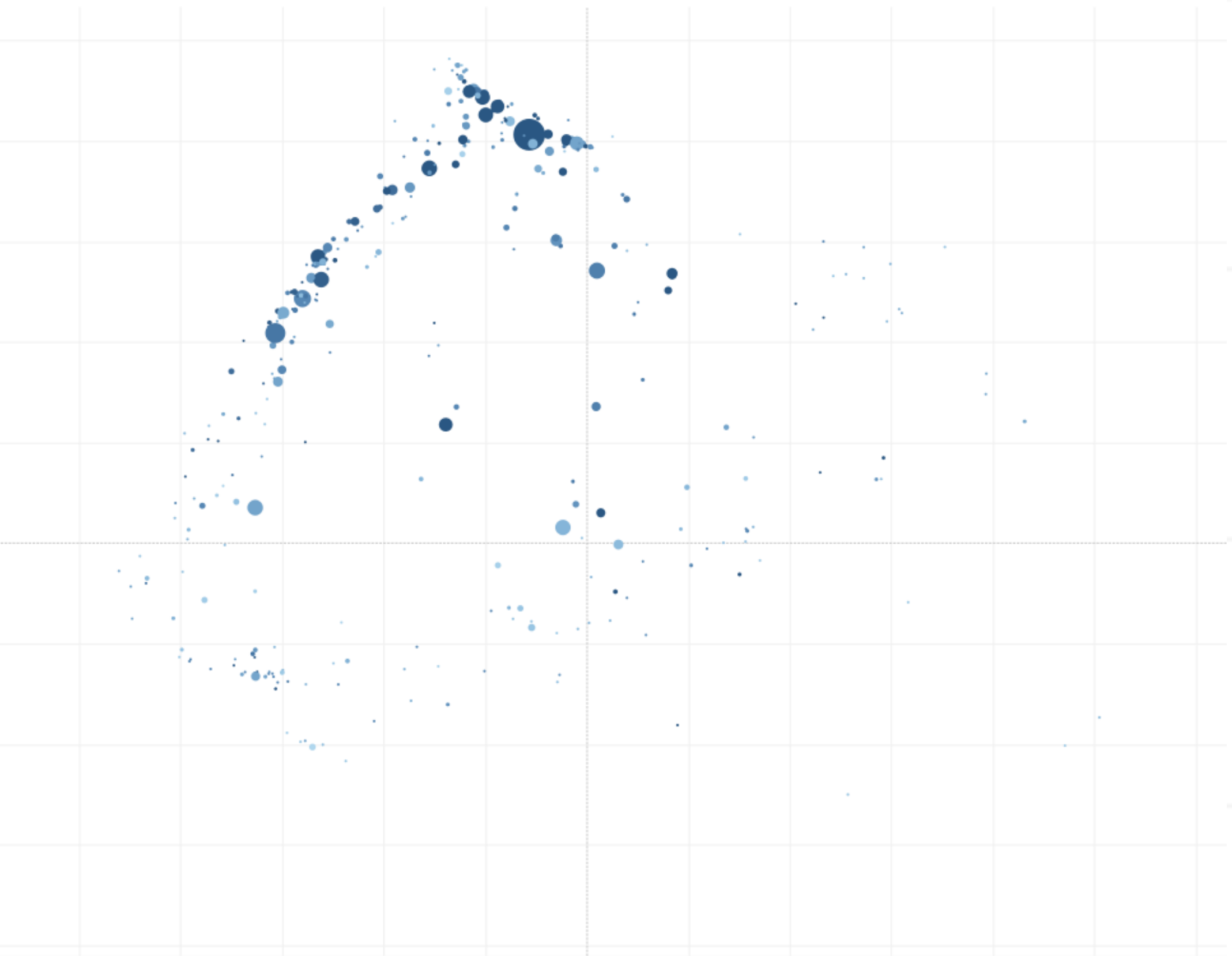
SDG 1 - End poverty in all its forms everywhere  
(thresholds: # seeds  $\geq 1,000$  or share seeds  $\geq 0.10$ )





# Example: SDG 1 (distribution

SDG 1 - End poverty in all its forms everywhere  
(thresholds: # seeds >= 0 or share seeds >= 0.00)



## Select SDG

SDG id

1

## Threshold

# seeds >=

1,000

OR

Share seeds >=

0.10

## Map layout

Param size

Size community



# Example: SDG 1, share seeds $\geq 0.1$

SDG 1 - End poverty in all its forms everywhere  
(thresholds:# seeds  $\geq 1,000$  or share seeds  $\geq 0.10$ )

Label community **1473**: multidimensional poverty; gini index; poverty dynamic; deprivation; income mobility

Top keywords: "employment " AND ("poverty") (23); "economic growth " AND ("poverty") (21); "vulnerable " AND ("poverty") (15); "income poverty" (15)

# seeds (2018): 131 (0.21653)  
# seeds selected (2018): 131  
# pubs in community (2018): 605

**Meso cluster info**  
Meso cluster: 83 (# micro: 10)  
Local agreement: 0.2000  
Meso labels: tax competition; intra industry trade; overeducation; multidimensional poverty; wage inequality

Select SDG

SDG id  
1

Threshold

# seeds  $\geq$   
1,000

OR

Share seeds  $\geq$   
0.10

Map layout

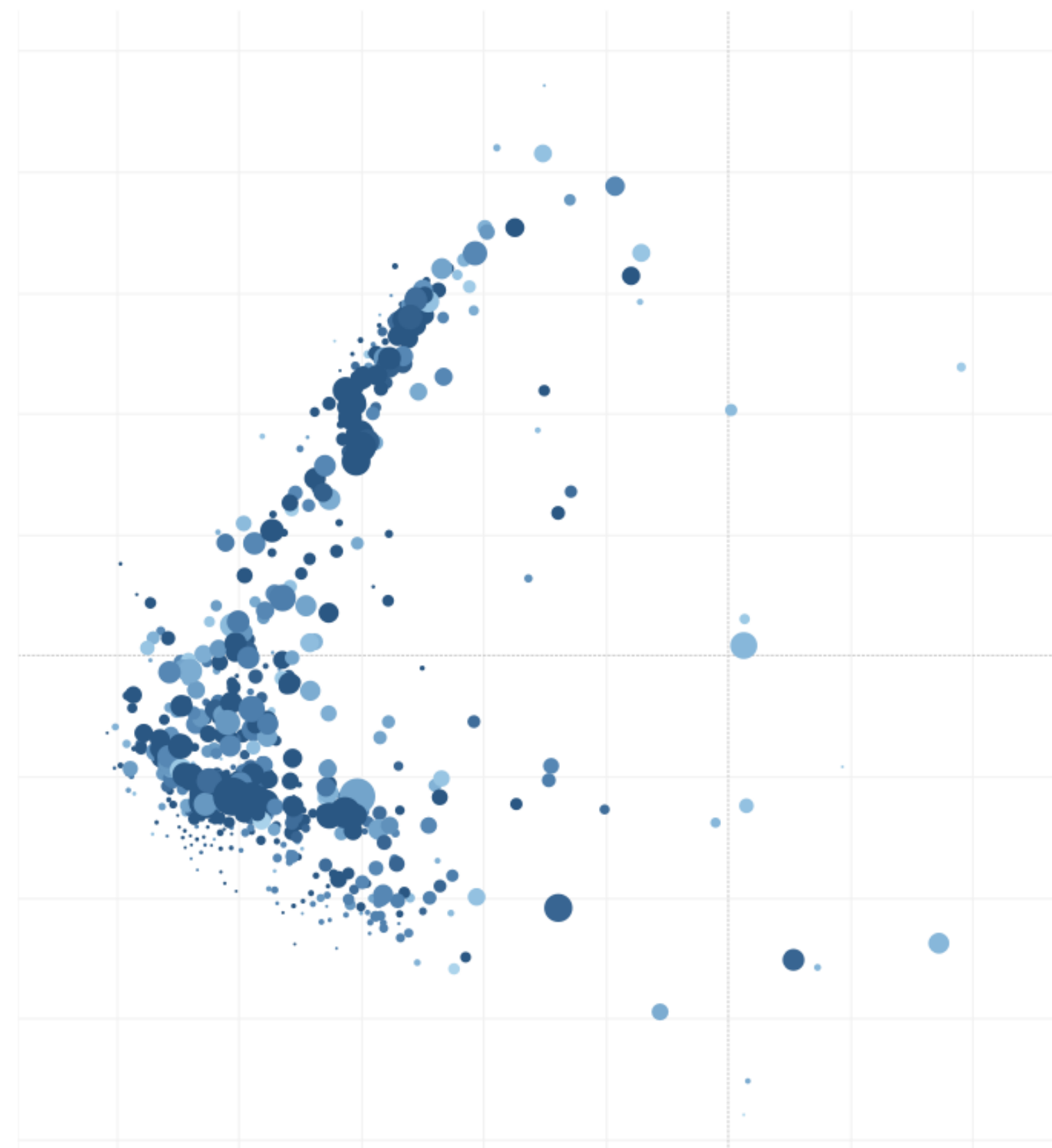
Param size  
Size community

Param color  
Local agreement



# Example SDG 3 (same threshold)

SDG 3 - Ensure healthy lives and promote well-being for all at all ages  
(thresholds: # seeds  $\geq 1,000$  or share seeds  $\geq 0.10$ )



Select SDG

SDG id  
3

Threshold

# seeds  $\geq$   
1,000

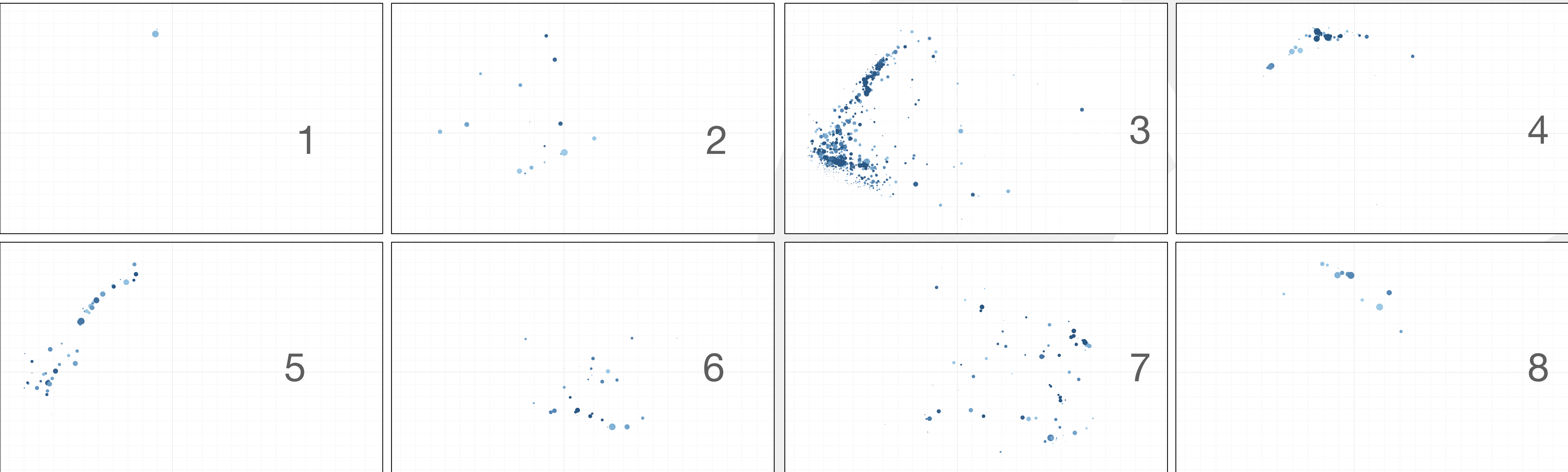
OR

Share seeds  $\geq$

A selection of related communities/ areas:

- Anticancer drug delivery
- MRI
- Ambient air pollution
- Epidemic model
- Bicycle, cycling
- Internet addiction
- Religiosity, mental health
- ...

# SDG 1-8 linked to R&D communities



# Implications/ opportunities/ conclusions

# Using results

- Less biased towards subjective choice of keywords
- Easier maintenance
- Identifying indirect impact or relatedness
- Picking up recent relevant developments
- Stable results



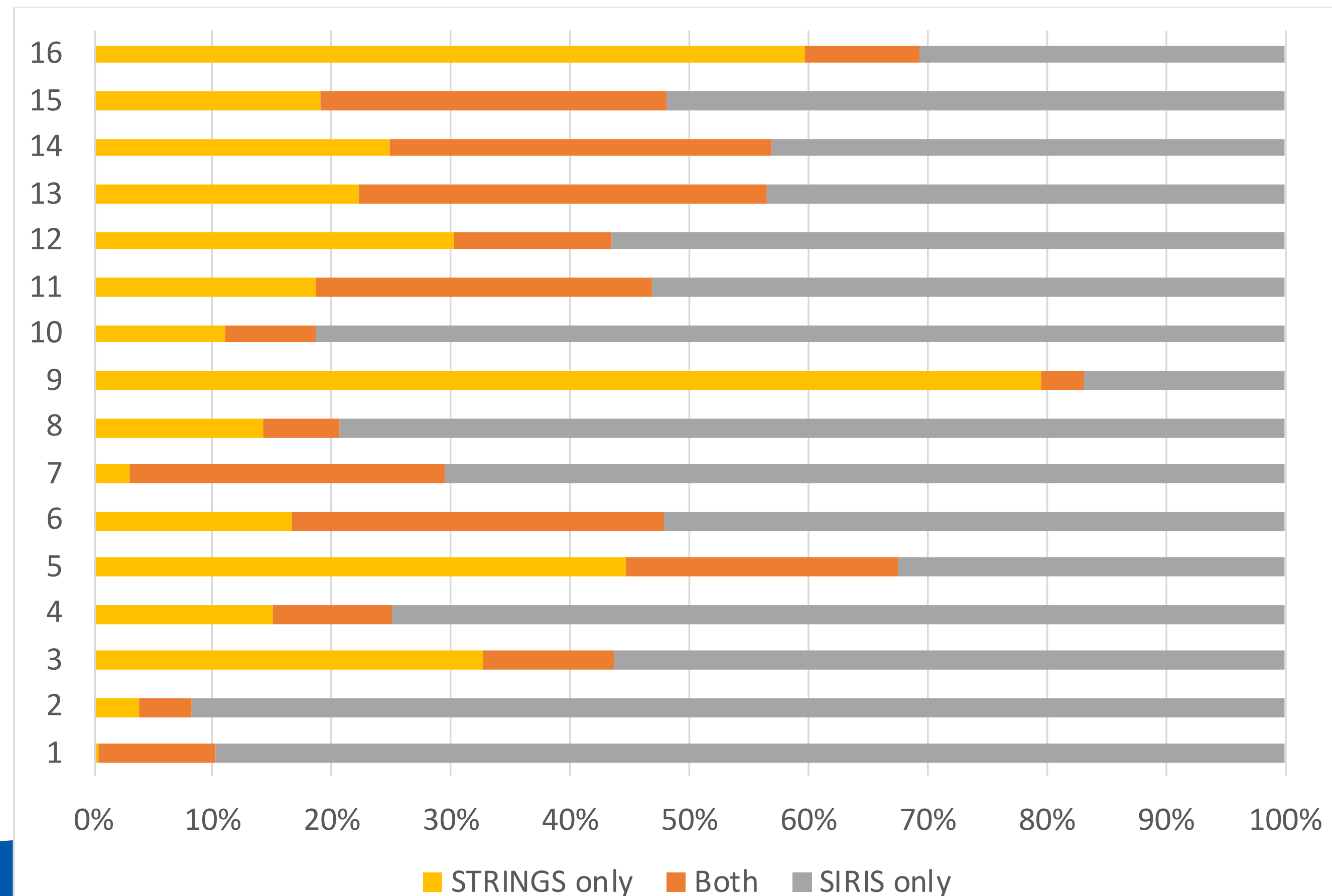
## Validation of linking output to SDG

- Check clusters rather than individual papers
- Opportunity to characterise link
- Check recall
- Compare different approaches
- Validate results by locating research cited in policy documents

# Comparing two approaches

# Collecting R&D output using keywords

- Comparing results STRINGS approach with SIRIS
- Similar collection procedures lead to little overlap in most SDGs

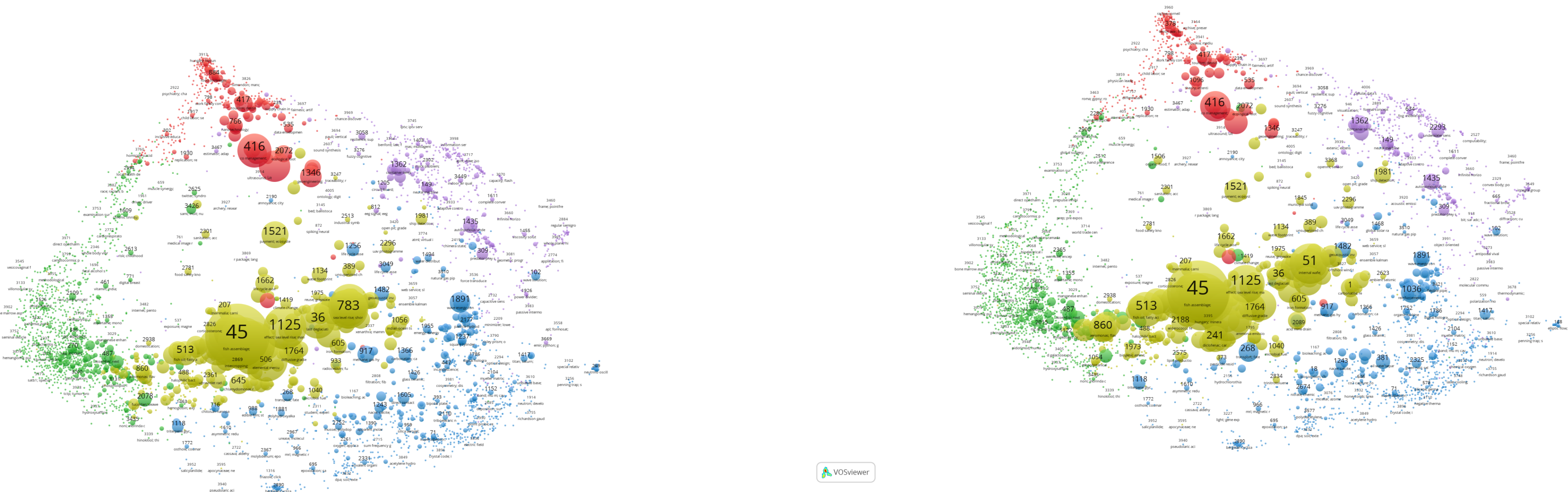


# SDG14 positioning STRINGS and SIRIS publications

Modest publications overlap, large community overlap

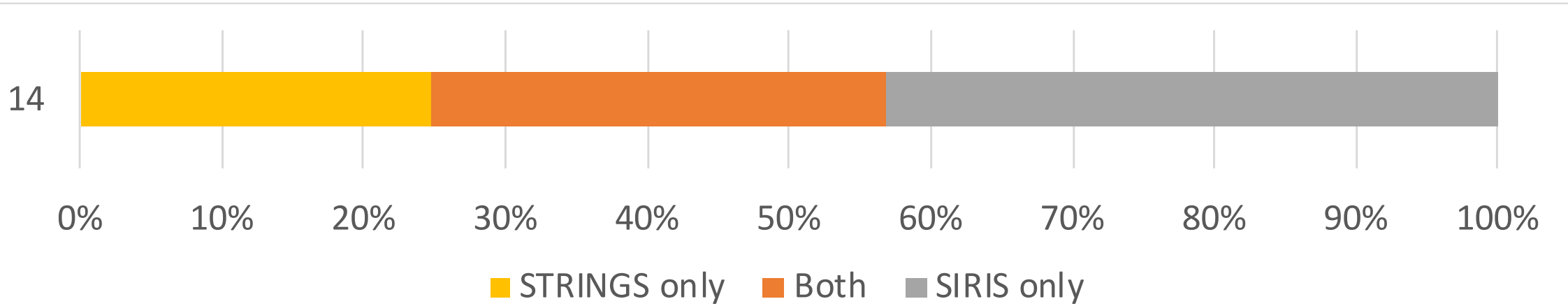
• STRINGS

• SIRIS



VOSviewer

VOSviewer



# Opportunities

- Room for different interpretations of an SDG and its relation to R&D
- Beyond output and scientific impact