

Better infrastructure  
Better science  
Better society



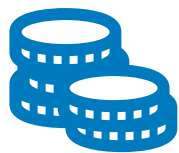
**ODISSEI**

Open Data Infrastructure for Social Science and Economic Innovations

# Current social data infrastructure



# Current social data infrastructure



Nationaal Kiesersonderzoek



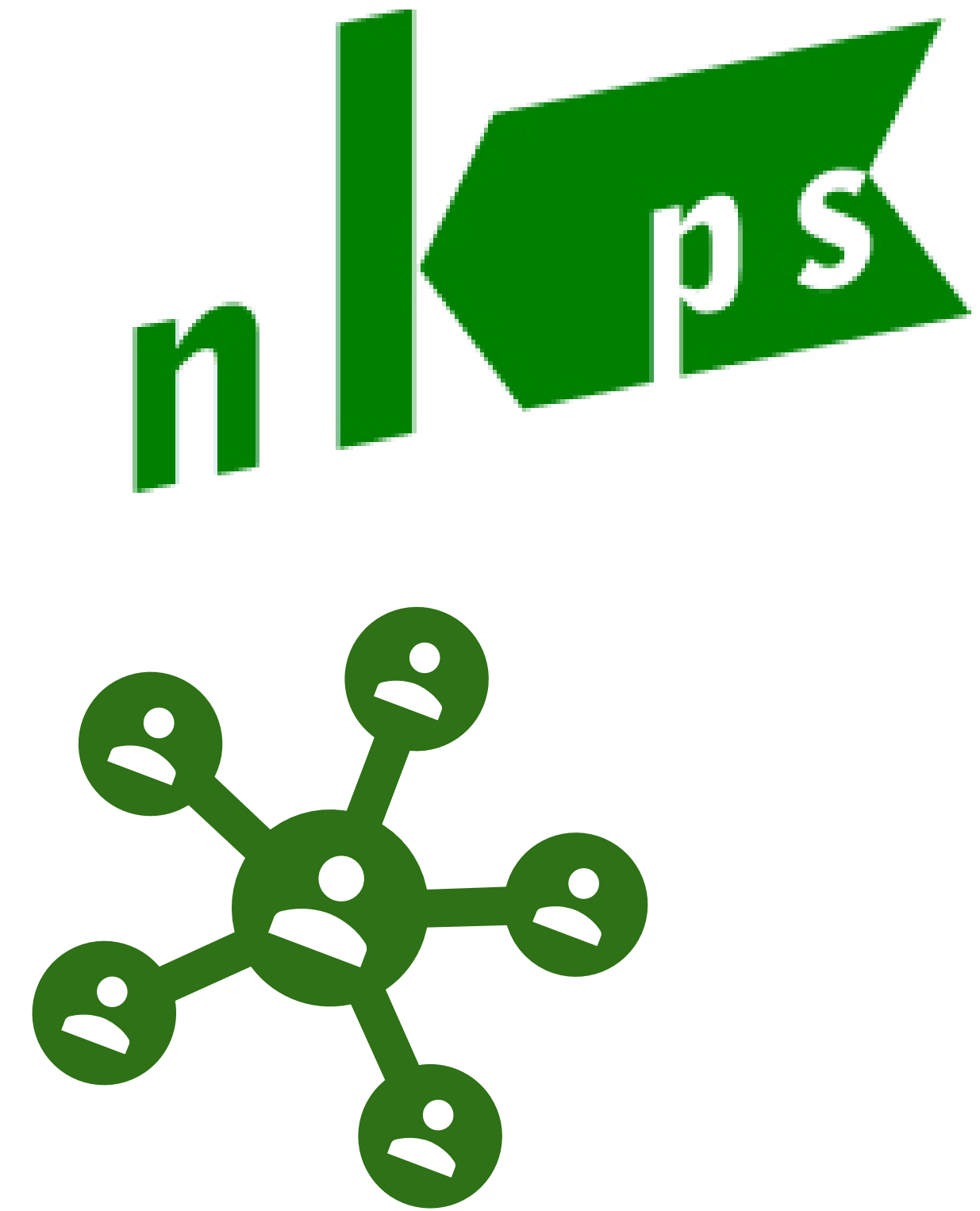
# Linked Lives



Pearl  
Dysktra



Families  
in Context





nsCr



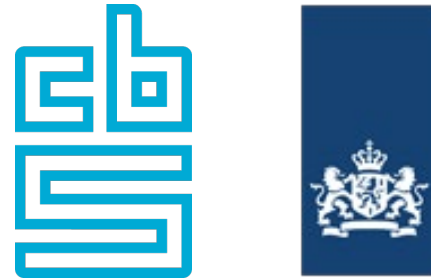
netherlands  
eScience  
center



# A unique collaborative consortium

Data Archiving and Networked Services  
**DANS**

N i D i



**TU Delft**

*Erasmus*



**CentERdata**  
Instituut voor dataverzameling en onderzoek

TILBURG UNIVERSITY

**Maastricht University**



UNIVERSITY OF TWENTE.



ODISSEI creates a data infrastructure for the social science in the Netherlands, on behalf of 32 member organisations.



# ODISSEI Secure Supercomputer (OSSC)



**ODISSEI**

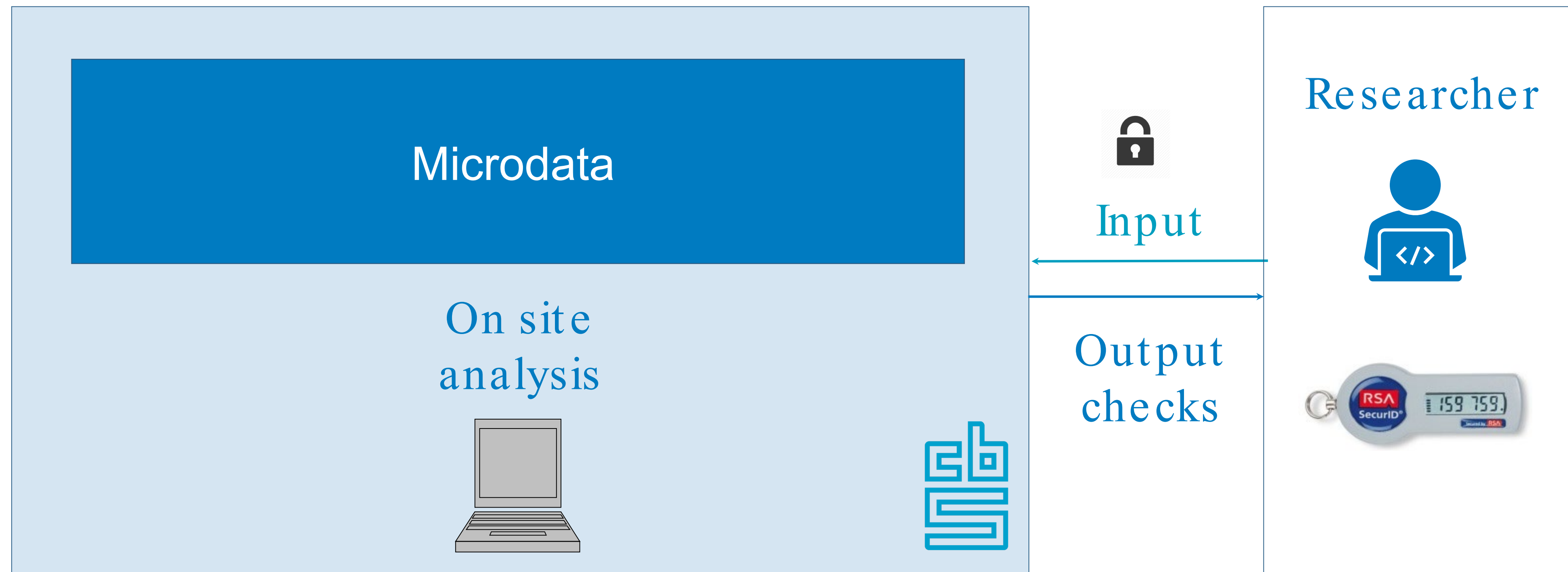
Open Data Infrastructure for Social Science and Economic Innovations

# CBS Microdata

**Microdata :**  
Linkable data on individual, companies, and addresses

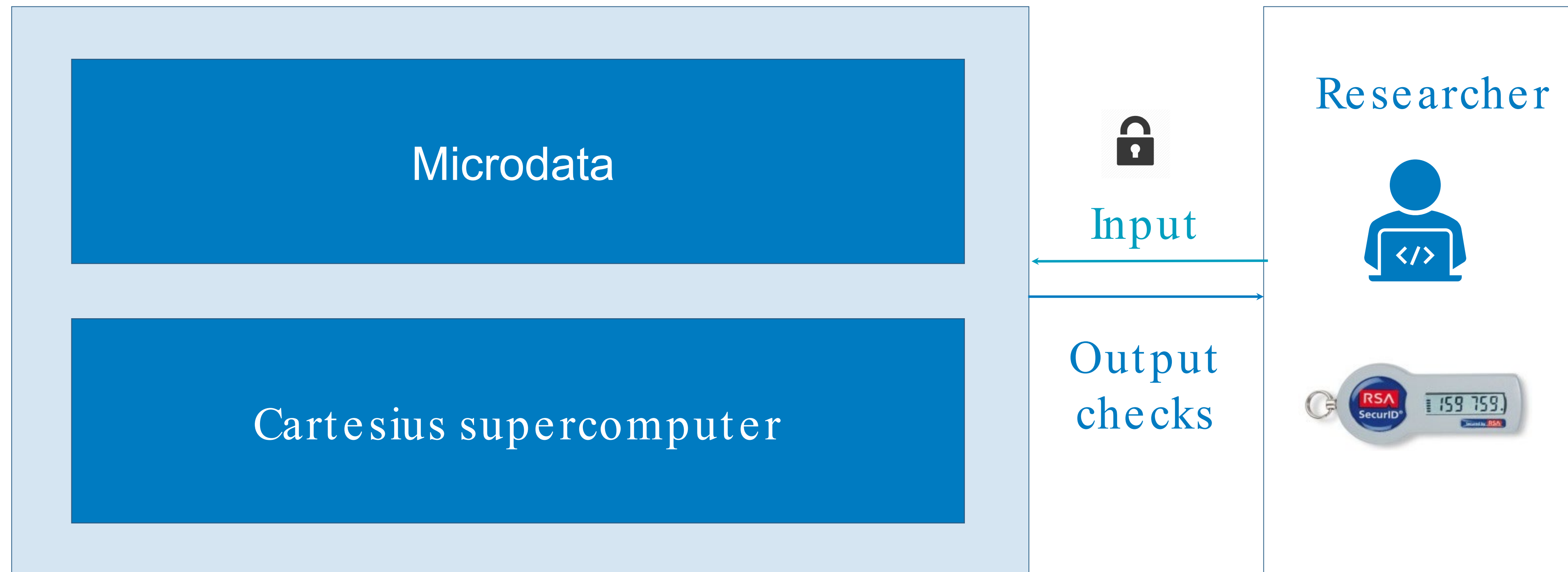


# CBS Remote Access environment

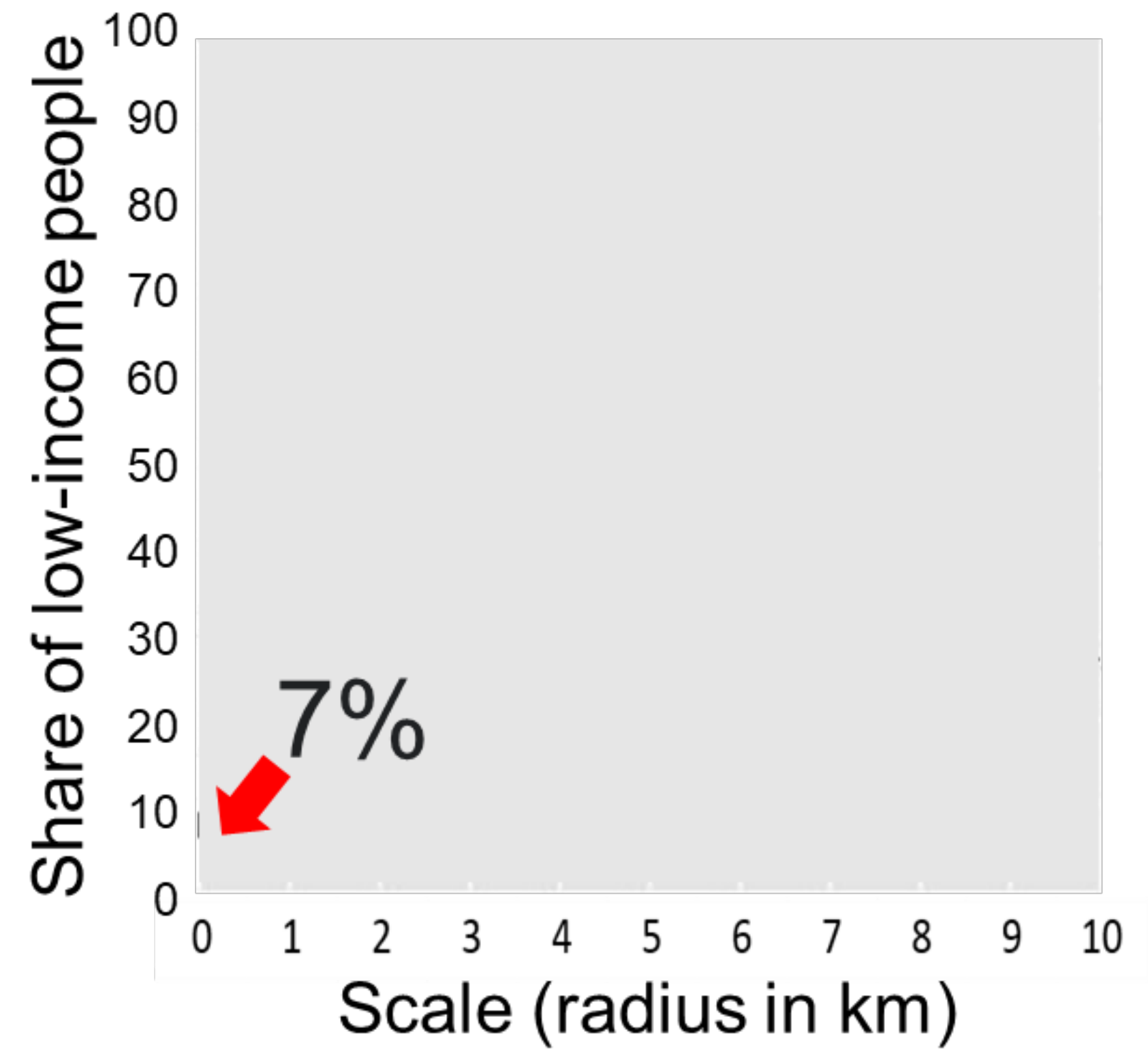
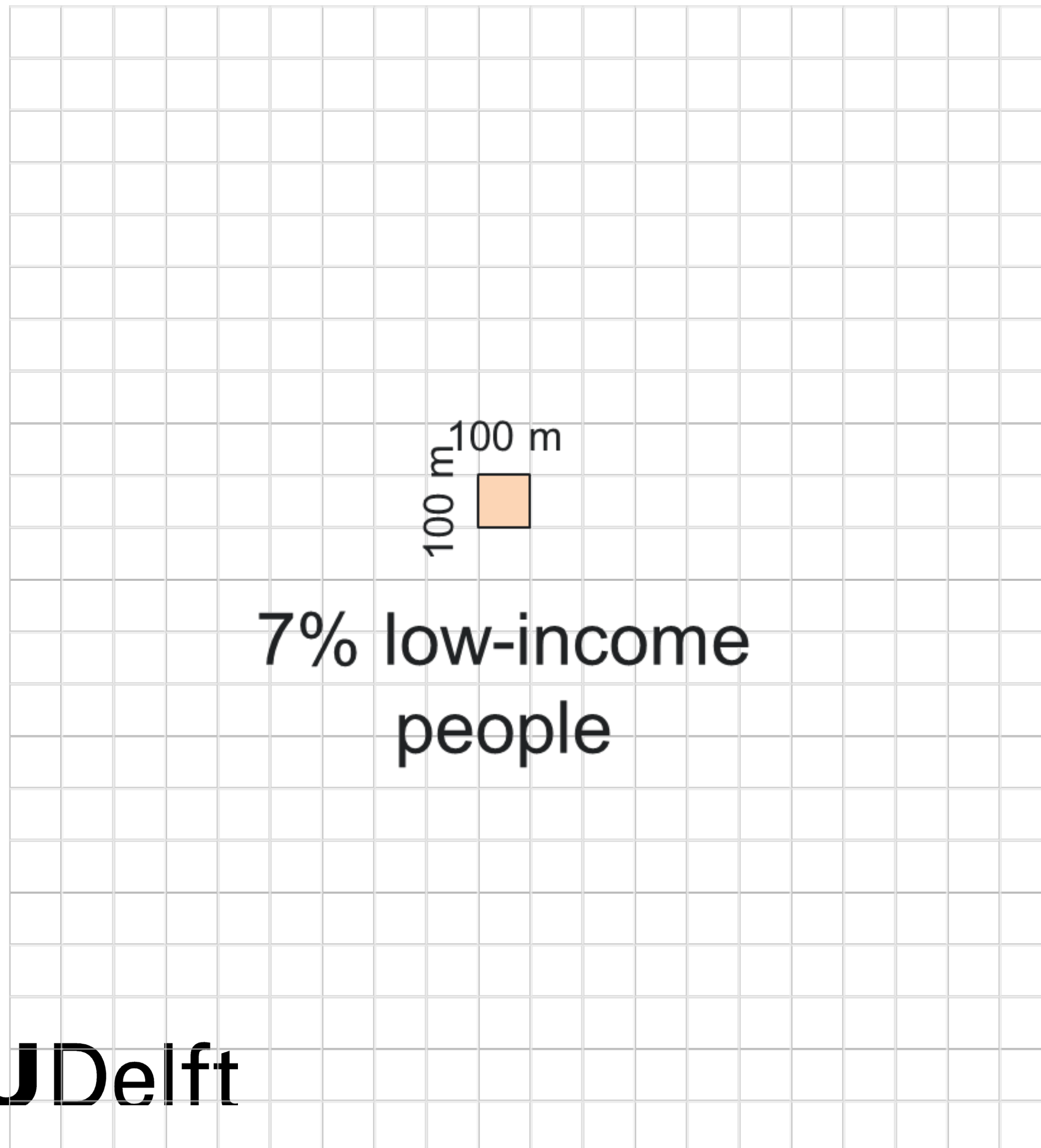




# The ODISSEI Secure Supercomputer (OSSC)

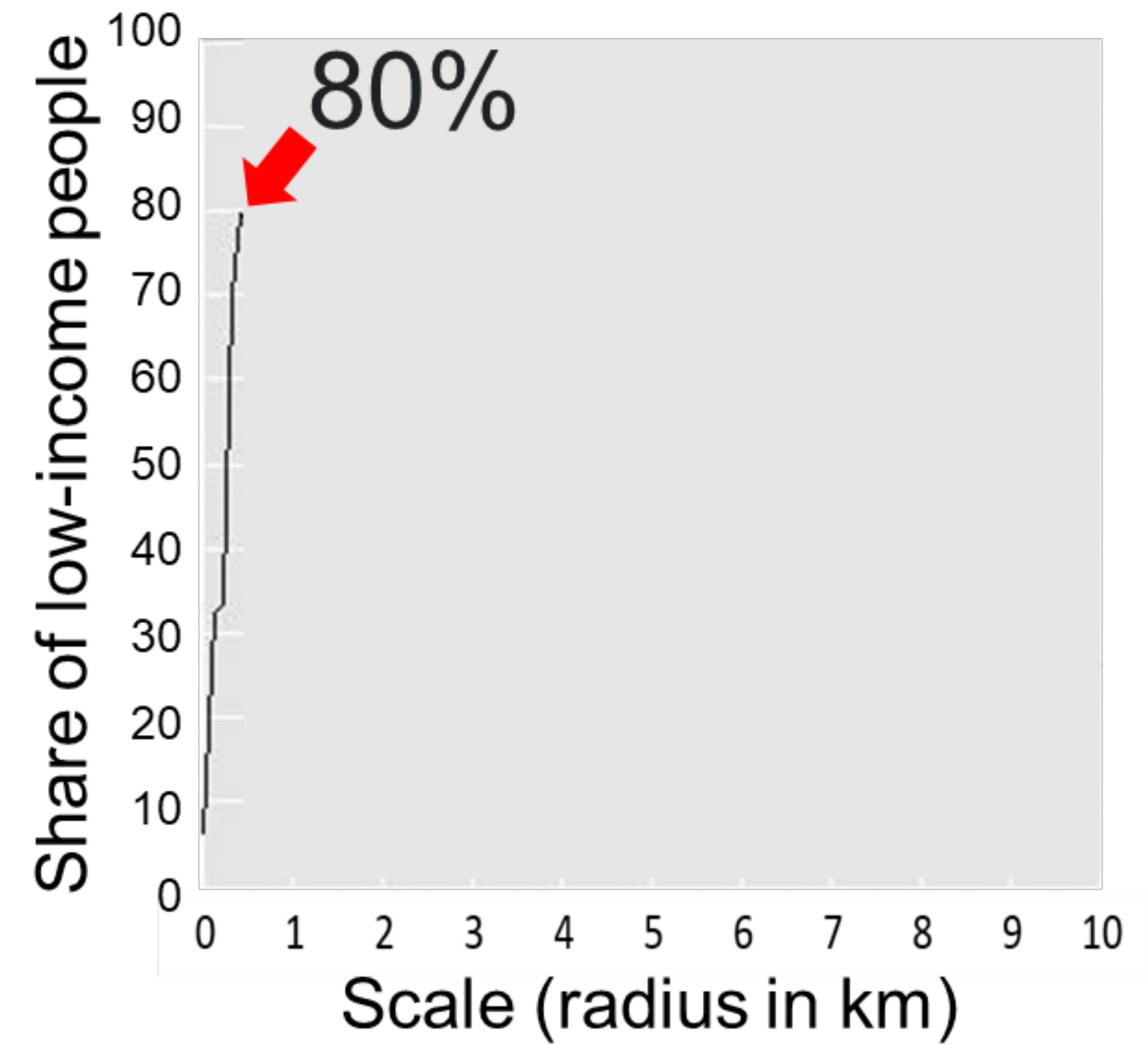
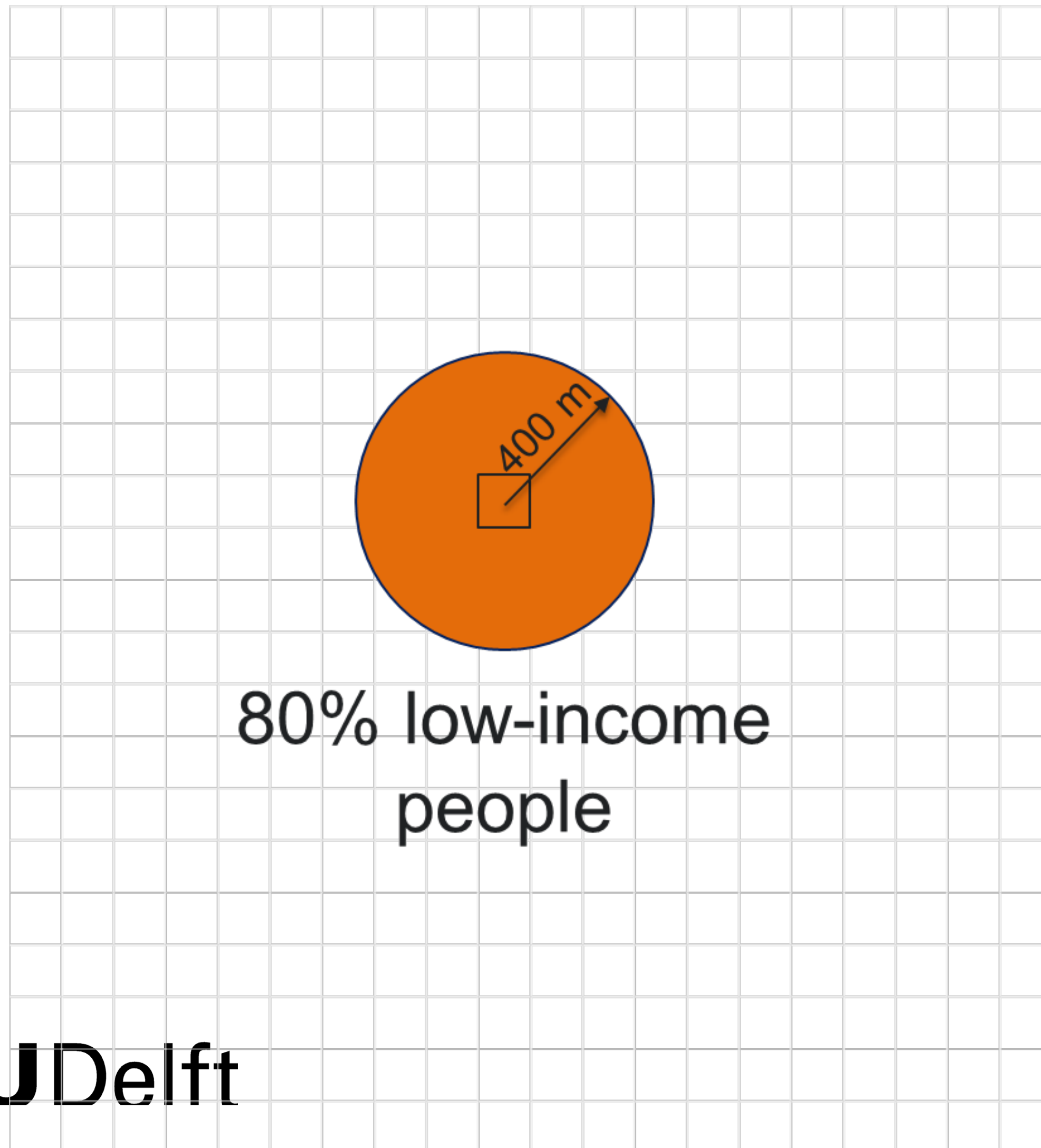


# Exposure to spatial context at 101 scales



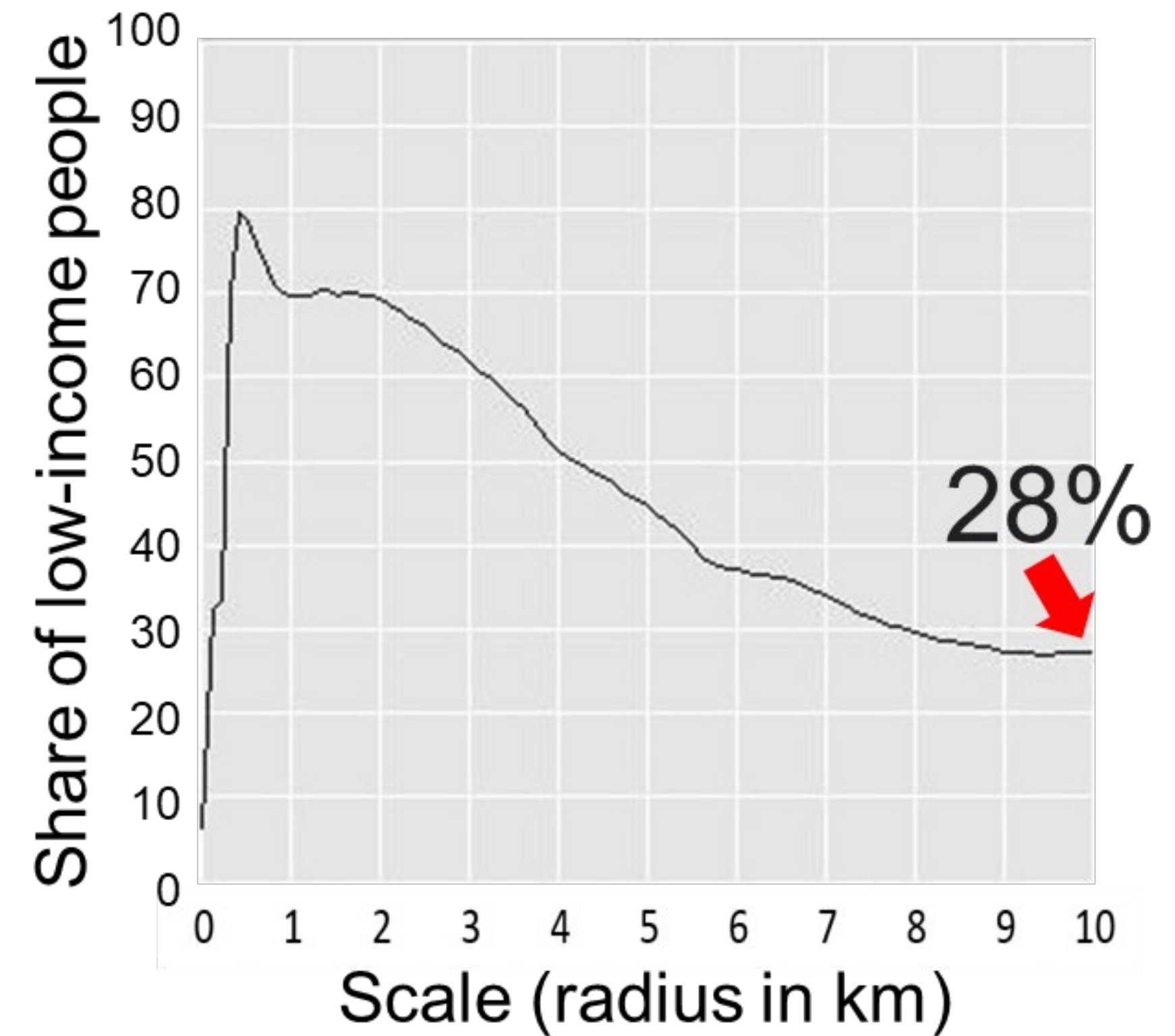
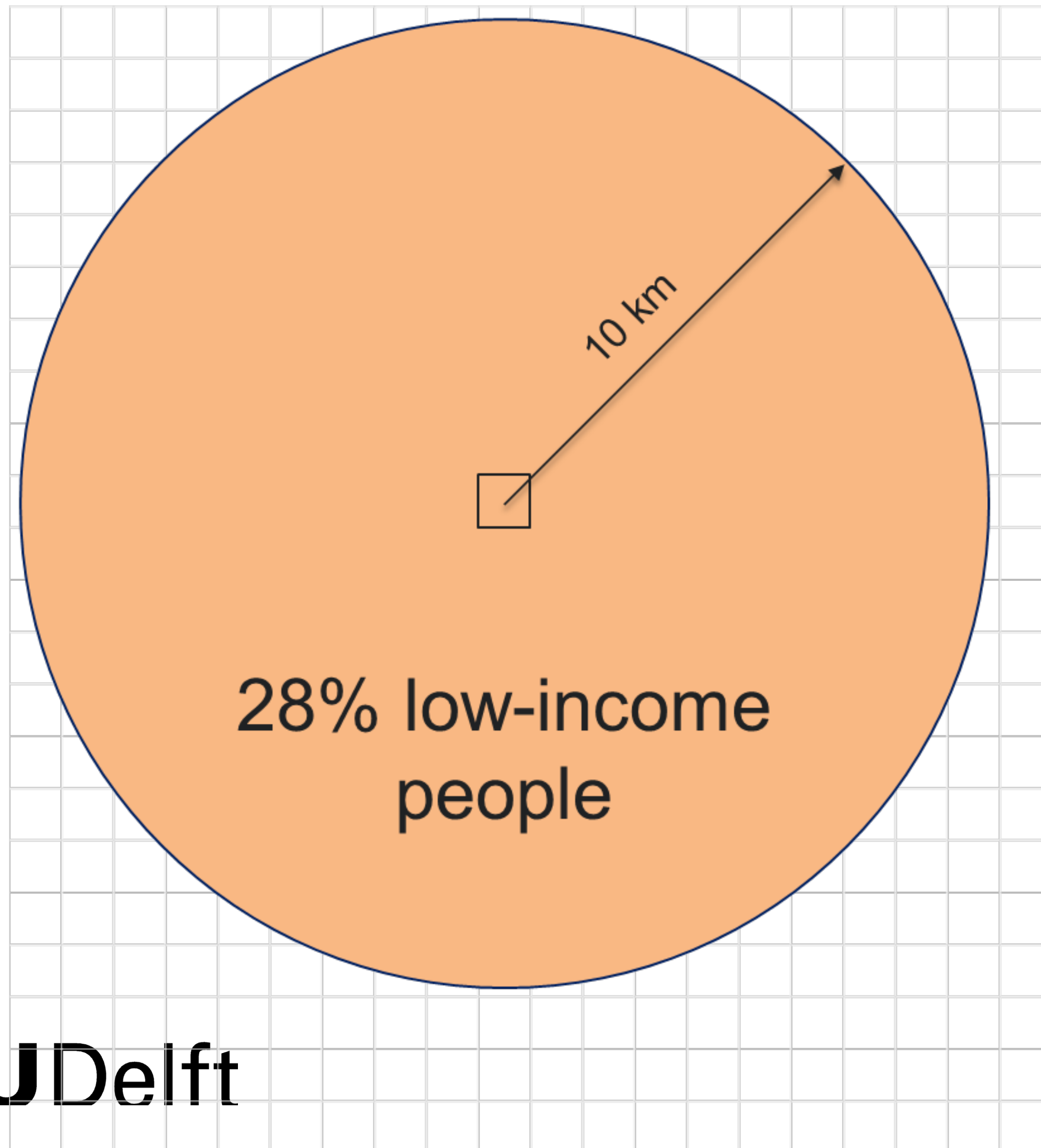
(see Petrović, van Ham, & Manley, 2018)

# Exposure to spatial context at 101 scales



(see Petrović, van Ham, & Manley, 2018)

# Exposure to spatial context at 101 scales



(see Petrović, van Ham, & Manley, 2018)

# Reduced computing time on ODISSEI Secure Supercomputer

585,000 populated cells × 101  
scales × 2 variables × 15 years =  
1.8 billion data points

→ 4 months continuous  
calculations

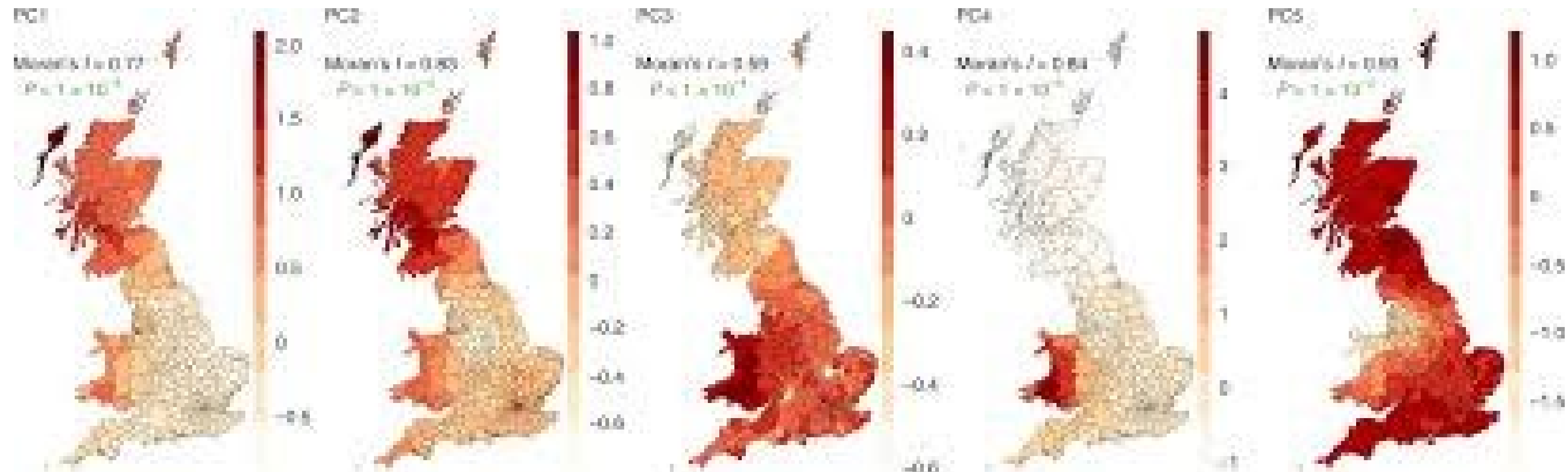
585,000 populated cells × 101  
scales × 3 variables × 19 years =  
3.4 billion data points

→ 1 week on 24 nodes

# Increasing the Dimensions in social analysis

- Increasing the number of time points (Months or Weeks)
- Replicating across indicators like educational level, poverty, unemployment, crime
- Examining interdependencies between indicators
- Correlates with geospatial data such as the availability of new facilities or infrastructure

# Case Studies – Genetic Analysis



## Case Studies – Genetic Analysis

Median genetic risk score for schizophrenia in Dutch Towns



Colodro-Conde, Lucía et al. 2018.  
“Association Between Population  
Density and Genetic Risk for  
Schizophrenia.” *JAMA Psychiatry*.  
Retrieved July 5, 2018  
(<http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/jamapsychiatry.2018.1581>).





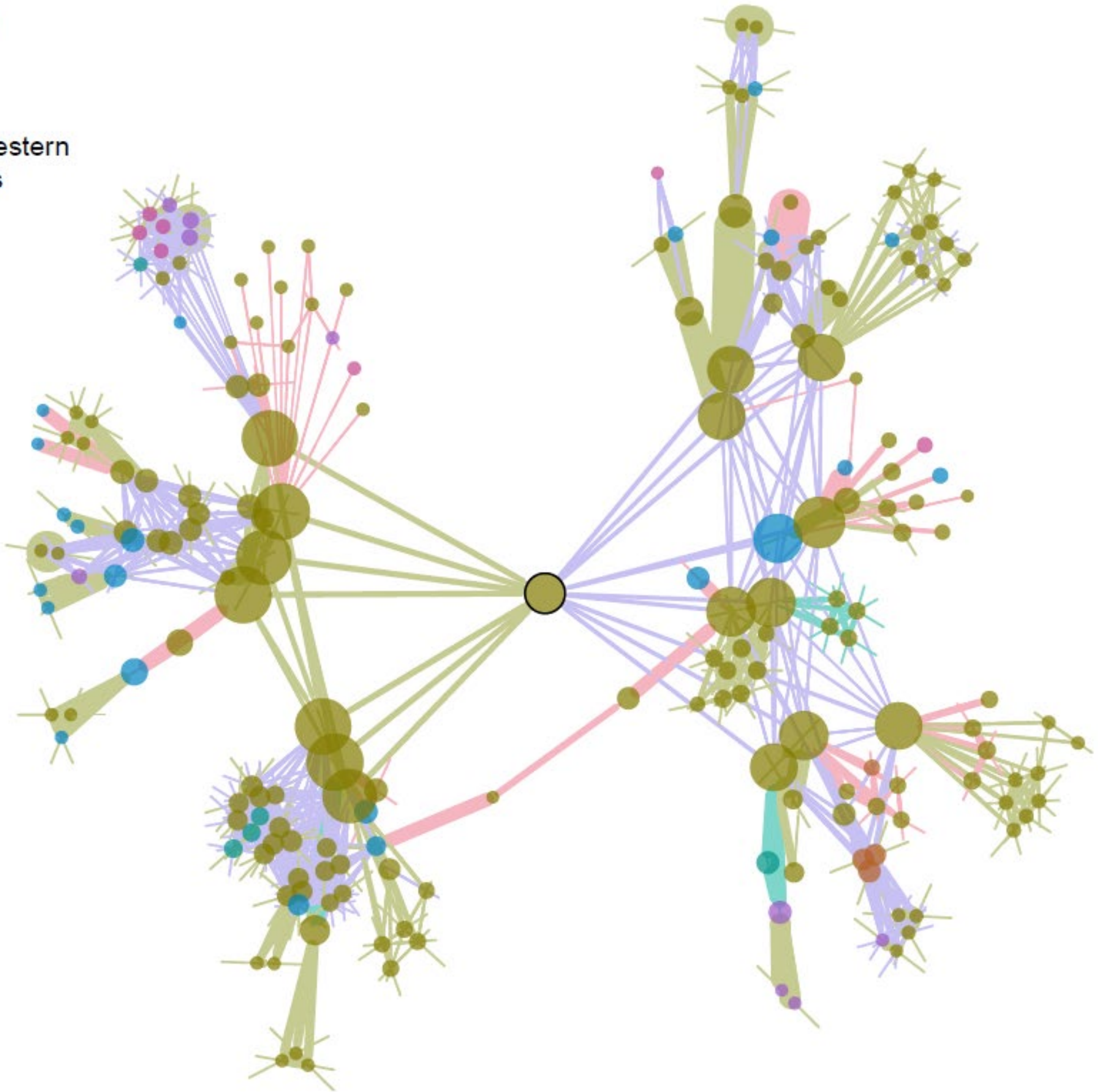
# Case Studies – Genetic Analysis

- How are specific genetic traits clustered across the Dutch population?
- To what extent do specific educational institutions support better outcomes for specific genetic traits?
- To what extent is behavior such as homosexuality associated with genetic traits, environmental cues, an interaction of both or neither?



# Case Studies – Network Analysis of CBS data

- Moroccan
- Native Dutch
- Turkish
- Surinam
- Western
- Other non-western
- Dutch Antilles



- werk
- familie
- school
- buurt



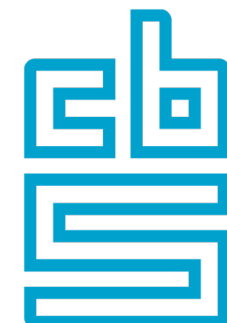
## Case Studies – Network Analysis of CBS data

- In which areas do immigrants become integrated and which do they remain isolated?
- How efficient are companies at matching workers to jobs in different sectors, regions or over time?
- How does energy efficient behavior spread through society? From work? Through neighborhoods? Through family?



# The ODISSEI Secure Supercomputer (OSSC)

Nowhere else is there an infrastructure that brings together register data, survey data and high-performance computing in a digital environment that is available to the entire community of social scientists.



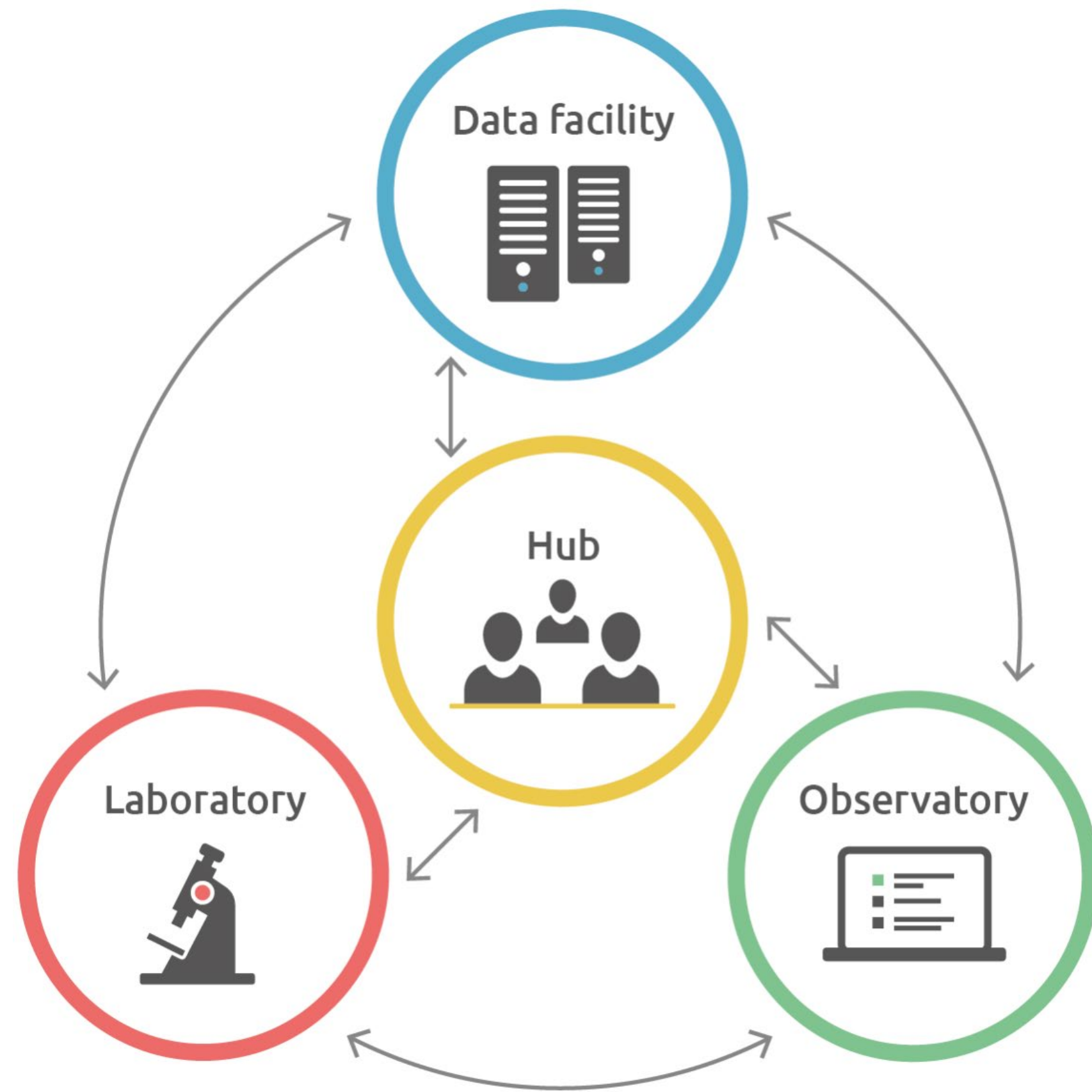
Future



**ODISSEI**

Open Data Infrastructure for Social Science and Economic Innovations

# Future



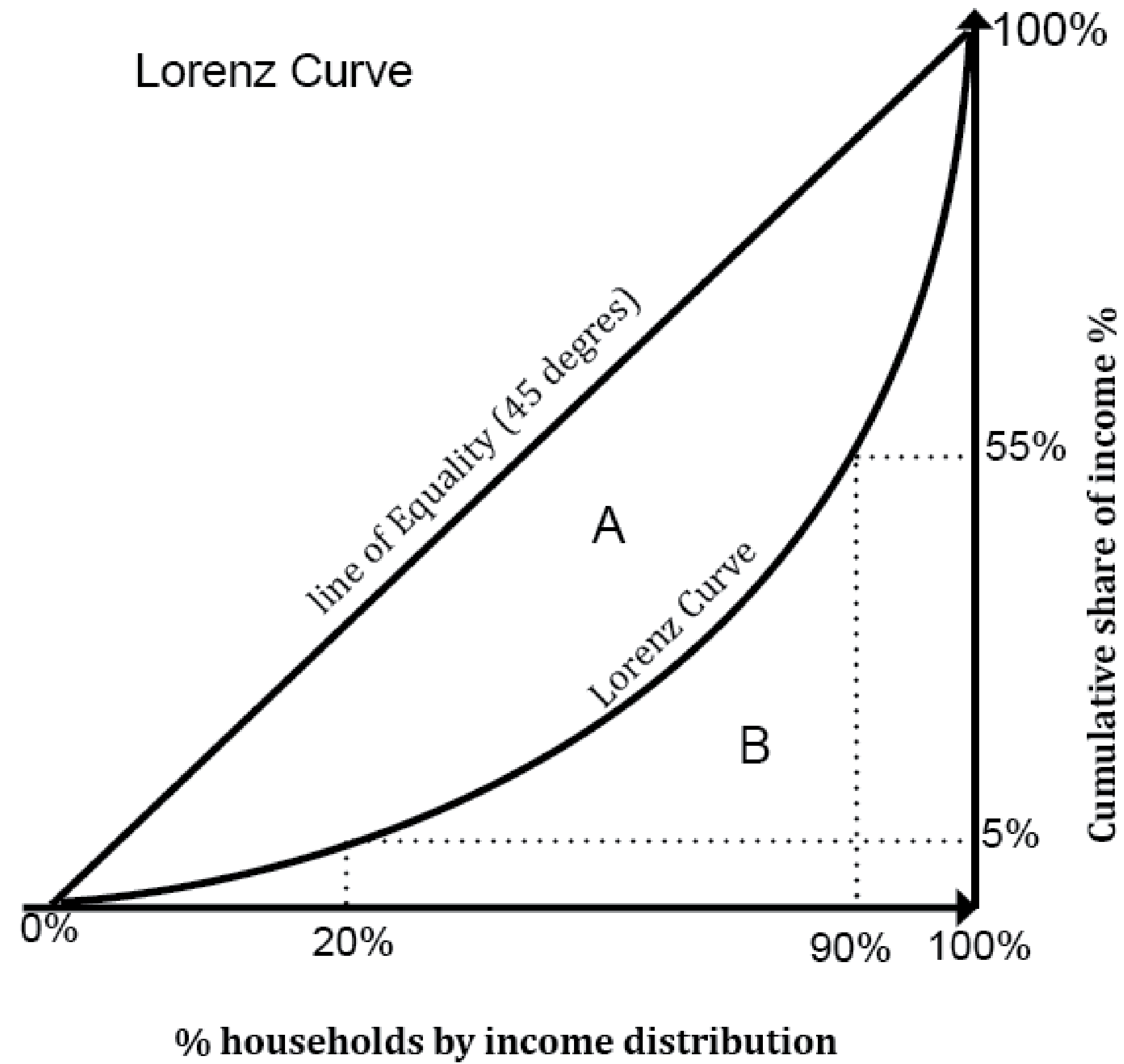
We applied for a NWO Large Infrastructure Roadmap grant. The outcome is expected in April 2020.



# Looking Ahead



# Looking Ahead





## Looking Ahead

- There is a unique opportunity for Dutch Social Science to forge ahead
- Important for analysis to build on top of each other. Data needs to be stored in a FAIR and Secure way
- Challenges in catching up and making use of the resources, tools and methods available for large and complex analysis



# Thank you

info@odissei-data.nl



**ODISSEI**

Open Data Infrastructure for Social Science and Economic Innovations