Data Science Project

Digital Ethics concepts mapping

Nikolaenko Yuliia
M2 DIgital Sciences track

Research question:

How the concept space of the Digital Ethics research looks like?

Aim of the research:

To identify the range of the topics around Digital Ethics studies

Research process

Data Collection	Data Preparation and Cleaning	Data Frame Creation	Network Creation	Network Analysis
Articles from open sources	Transforming pdf to text files, merging texts, lower case, identify and remove stopwords, noise and punctuation	Data frame of the words co-appearance in the texts	Setting the nodes and edges sizes	Community detection and clustering

Data

Context:

10 articles for each topic connected to Digital Ethics studies:

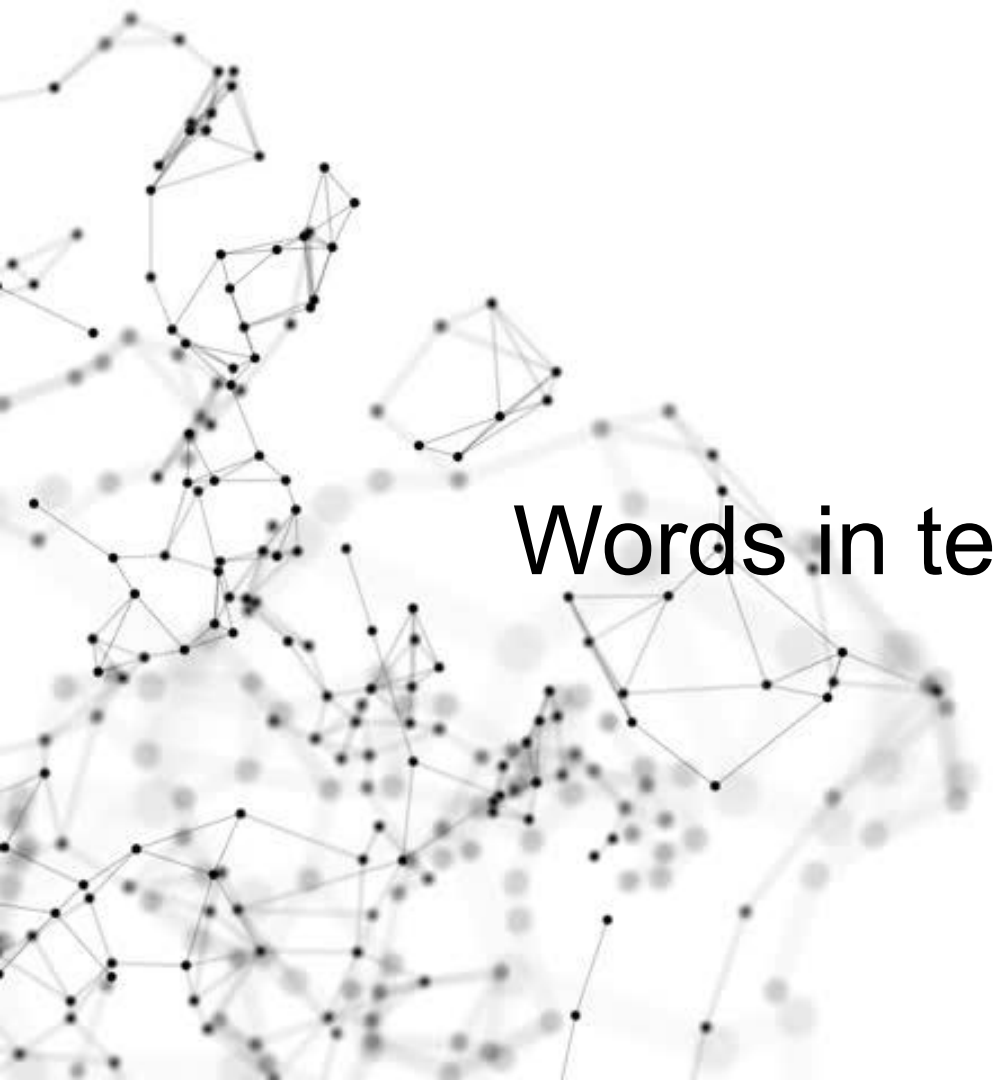
1. Artificial intelligence
2. Robotics
3. Open Government, Trust, Ethics in e-democracy
4. Open Source
5. Digital Accessibility

Overall:

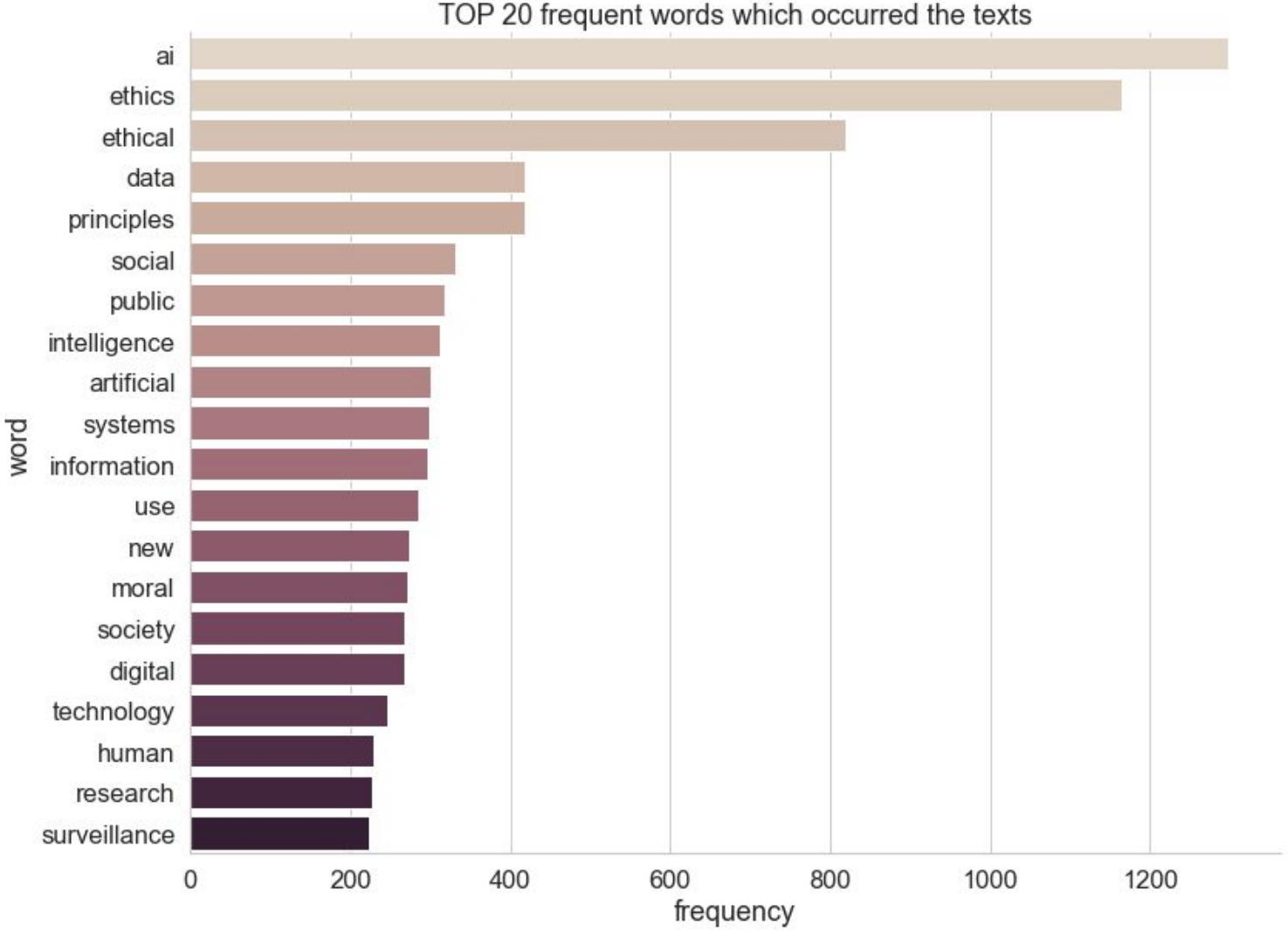
50 texts

76318 tokens after cleaning

1000 top bigrams for the network analysis

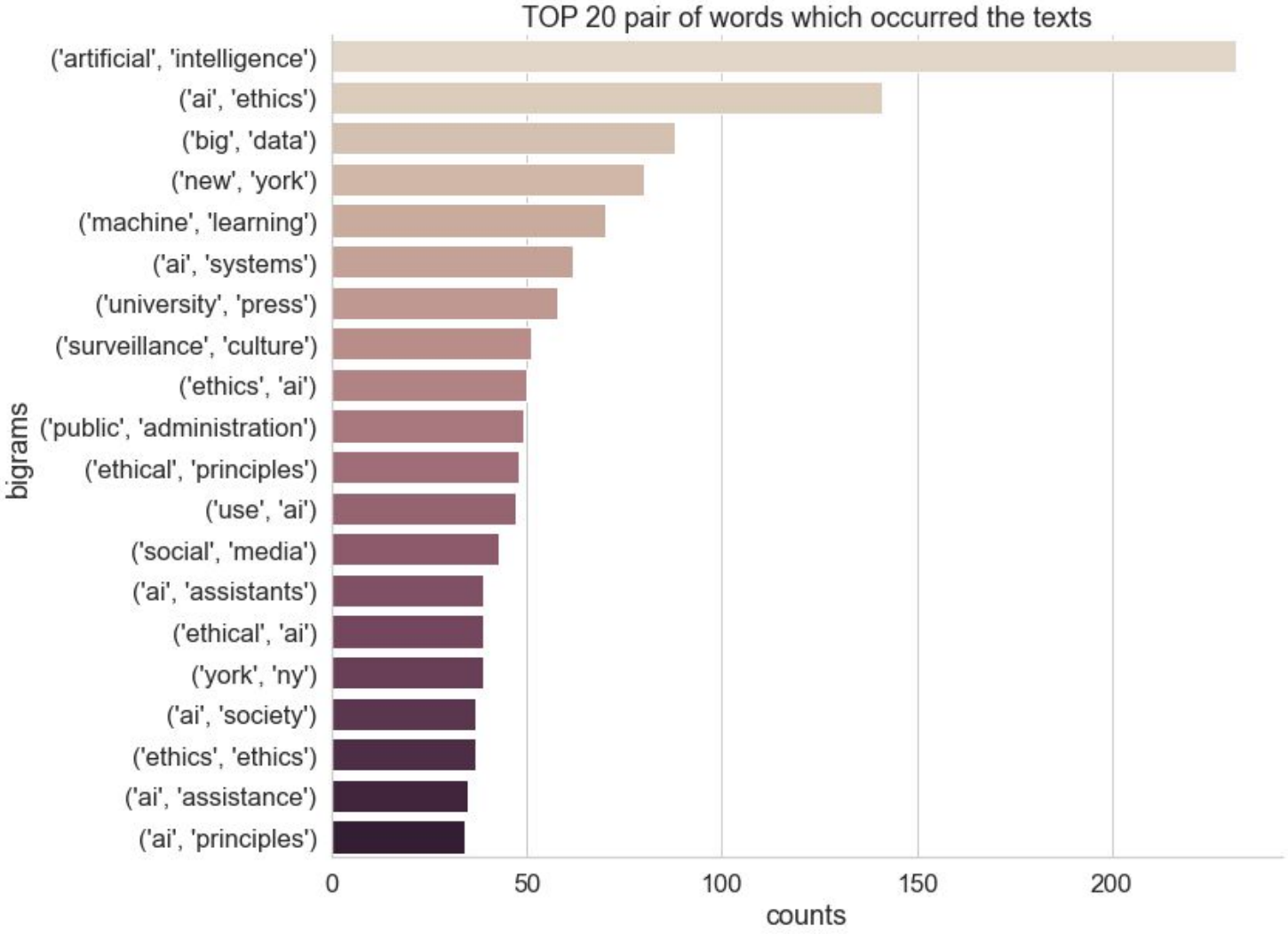


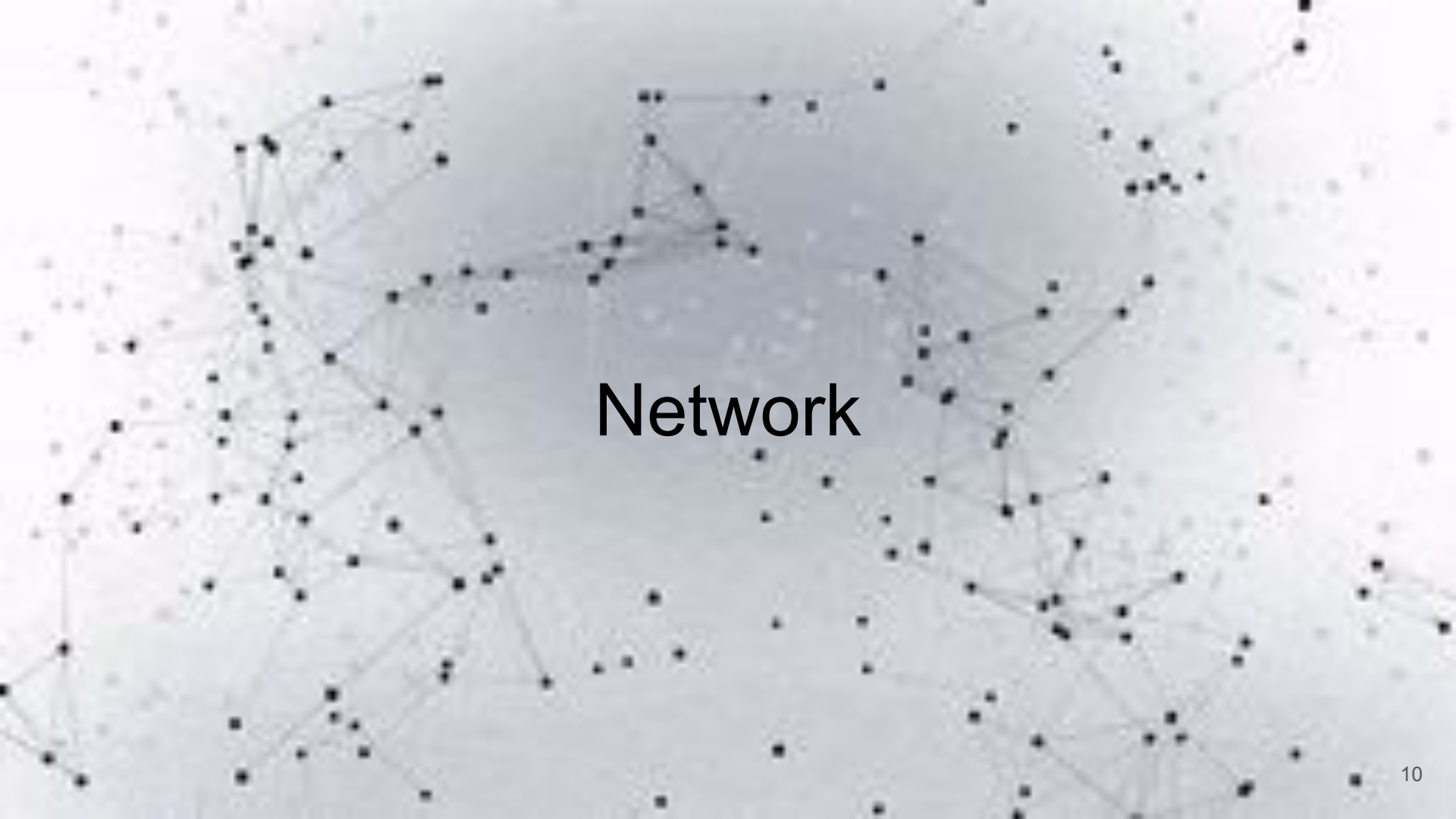
Words in text context


TOP 20 frequent words which occurred the texts
ai
ethics
ethical
data
principles
social
public
intelligence
artificial
systems
information
use
new
moral
society
digital
technology
human
research
surveillance
word
0
200
400
600
800
1000
1200
frequency


Wordcloud of TOP 100 frequent words which occurred the texts
ethical
social
ethic
artificial
intelligence
ai
principle
use
technology
value
human
individual
people
data
privacy
society
moral
information
digital
ai ethics
guérin
guideline
big data
nature
state
problem
possible
framework
future
power
issue
systems
context
citizen
design
internet
impact
political
group
example
new
case
robot
health
concern
risk
action
order
set
democracy
machine
science
analysis
surveillance
specific
change
new york
approach
governance
autonomy
term
government
challenge
process
provide
work
right
result
law
different
world
based
benefit
application
code
article
professional
need
form
effect
public
mean
important
general
time
paper
way
research
agent
question
decision
development
point
algorithm
culture
choice
situation
policy

Bigrams in text context


TOP 20 pair of words which occurred the texts
('artificial', 'intelligence')
('ai', 'ethics')
('big', 'data')
('new', 'york')
('machine', 'learning')
('ai', 'systems')
('university', 'press')
('surveillance', 'culture')
('ethics', 'ai')
('public', 'administration')
('ethical', 'principles')
('use', 'ai')
('social', 'media')
('ai', 'assistants')
('ethical', 'ai')
('york', 'ny')
('ai', 'society')
('ethics', 'ethics')
('ai', 'assistance')
('ai', 'principles')
bigrams
0
50
100
150
200
counts


Network

Network

Context:

Undirected Graph

Nodes:782

Edges: 998

Weight: co-appearance frequency

Filters:

Degree Range 3-71

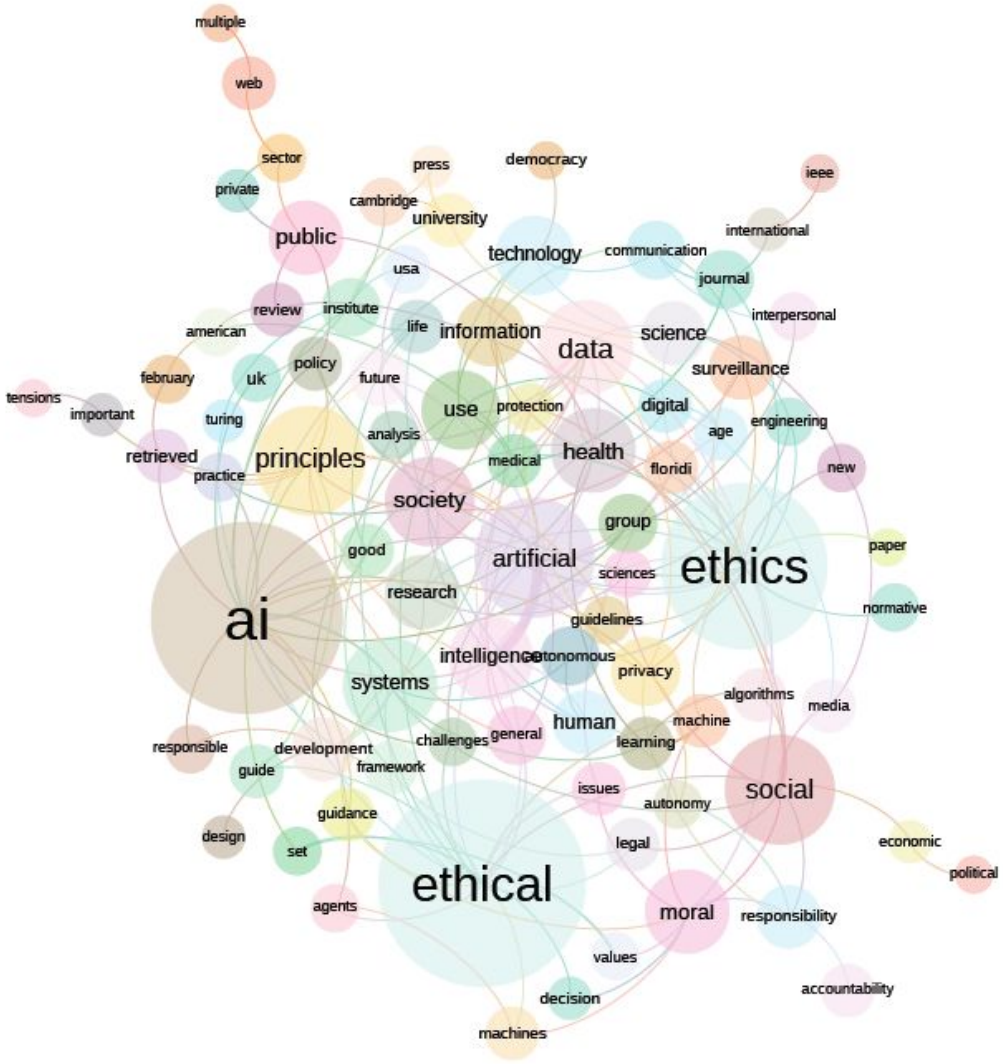
Giant Components

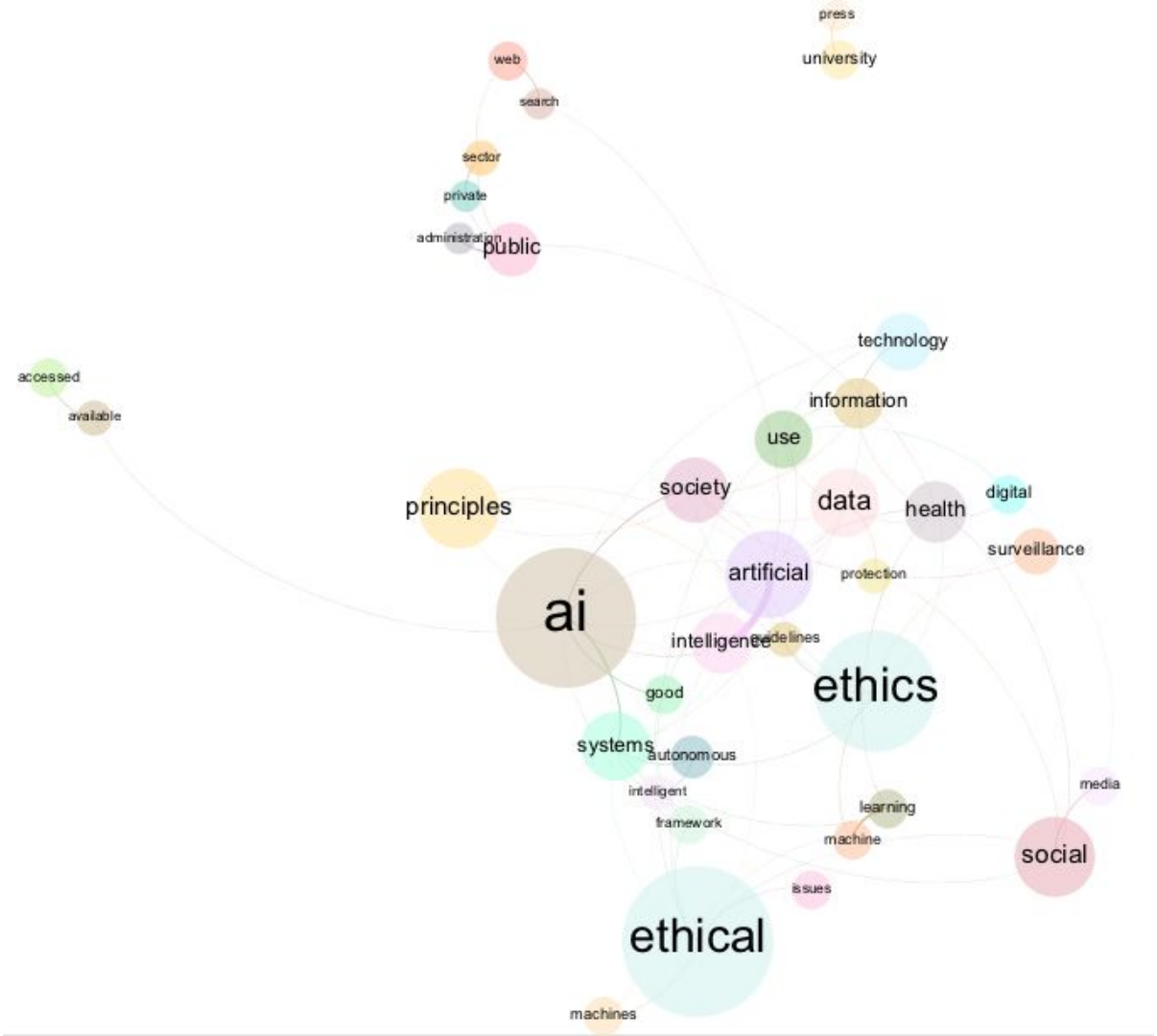
Edge Weight 4-232

Statistics:

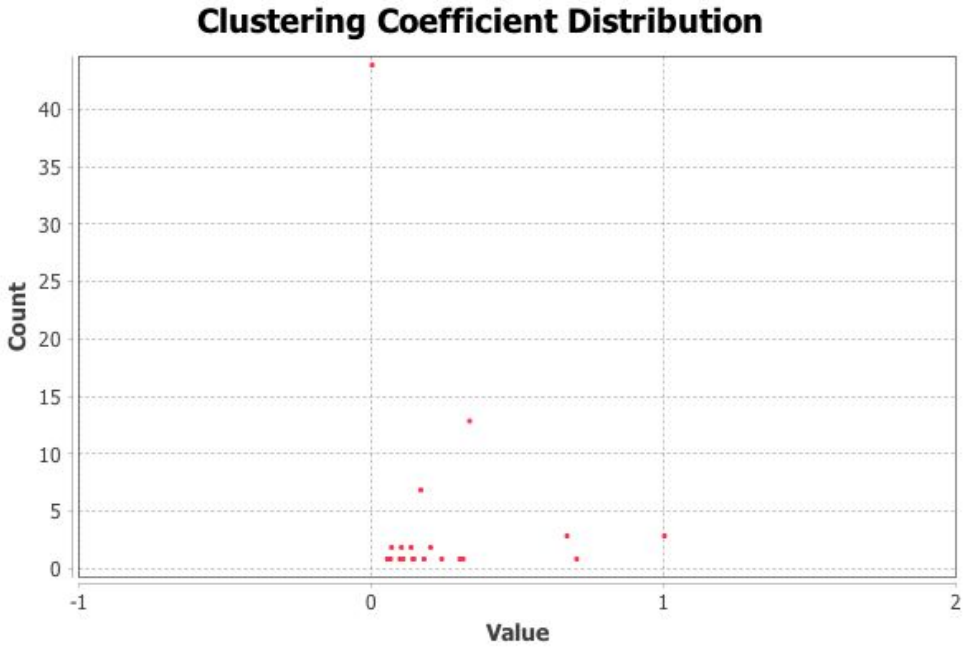
Density: 0.048

Connected components: 6


multiple
web
sector
private
press
democracy
cambridge
university
ieee
international
public
technology
communication
journal
usa
review
institute
life
information
data
science
interpersonal
american
surveillance
february
uk
policy
future
tensions
important
use
protection
digital
engineering
turing
analysis
age
retrieved
principles
medical
health
floridi
new
practice
society
group
good
artificial
ethics
paper
sciences
research
normative
ai
guidelines
intelligence
autonomous
privacy
systems
algorithms
media
human
machine
responsible
development
challenges
general
learning
framework
guide
social
issues
guidance
autonomy
design
set
legal
economic
political
ethical
agents
moral
responsibility
values
accountability
decision
machines


press
university
web
search
sector
private
administration
public
technology
accessed
available
information
use
society
principles
data
health
digital
surveillance
artificial
protection
ai
intelligence
guidelines
ethics
good
systems
autonomous
intelligent
media
learning
framework
machine
social
issues
ethical
machines

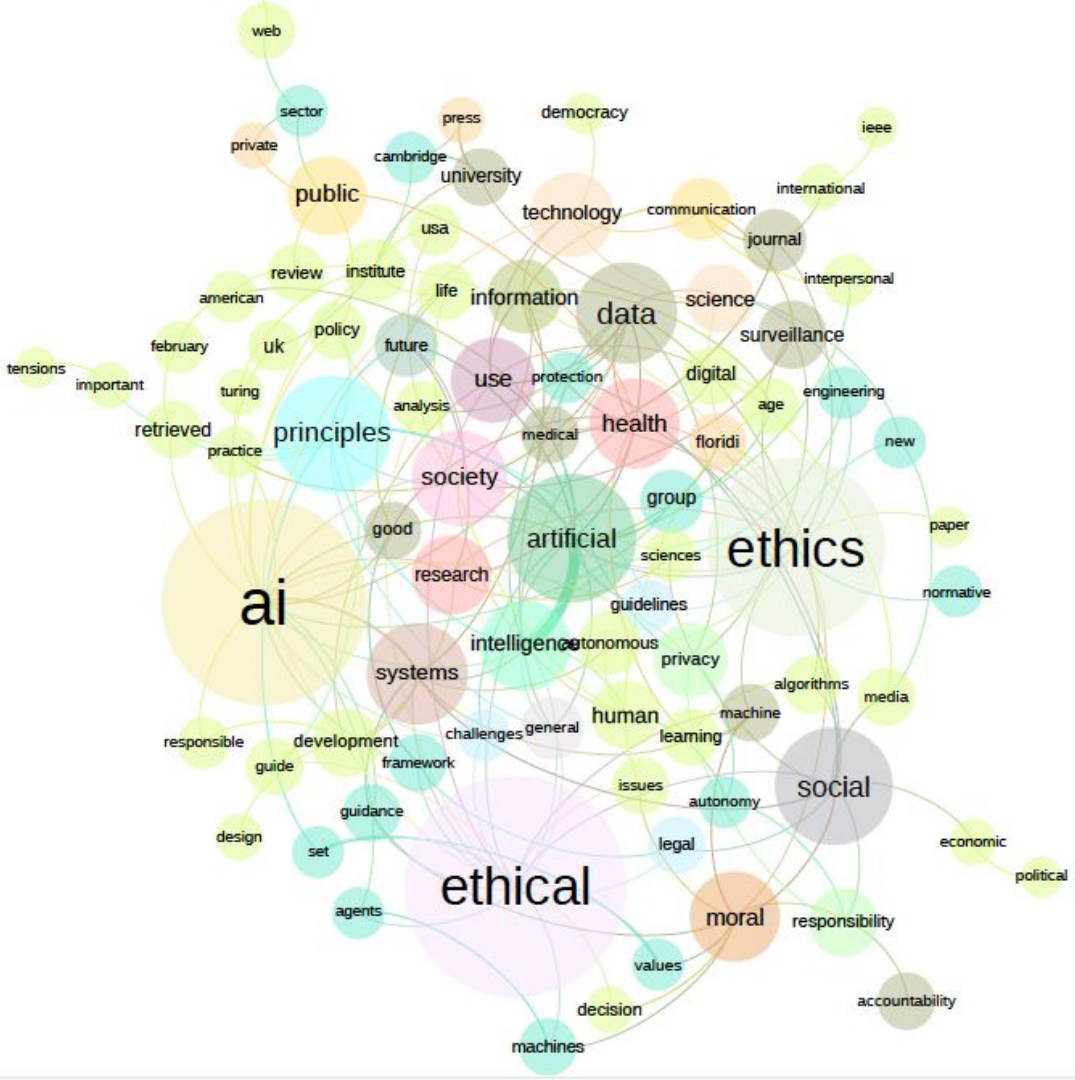
Additional analysis: detecting the clusters



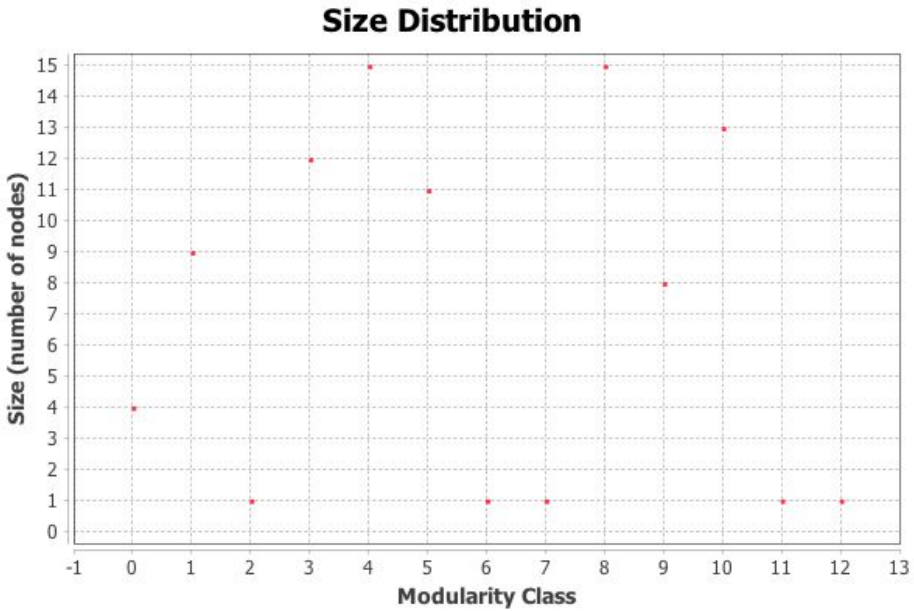
Average Clustering Coefficient: 0.179

Total triangles: 57

The Average Clustering Coefficient is the mean value of individual coefficients.


web
sector
press
democracy
ieee
private
cambridge
university
public
international
technology
communication
usa
journal
review
institute
interpersonal
life
information
science
american
data
policy
surveillance
february
uk
future
tensions
use
protection
digital
important
turing
engineering
analysis
age
retrieved
principles
health
medical
floridi
new
practice
society
group
good
artificial
ethics
paper
sciences
research
normative
ai
guidelines
intelligence
autonomous
privacy
systems
algorithms
media
human
machine
responsible
development
challenges
general
learning
guide
framework
issues
social
autonomy
guidance
design
legal
economic
set
ethical
political
agents
moral
responsibility
values
accountability
decision
machines

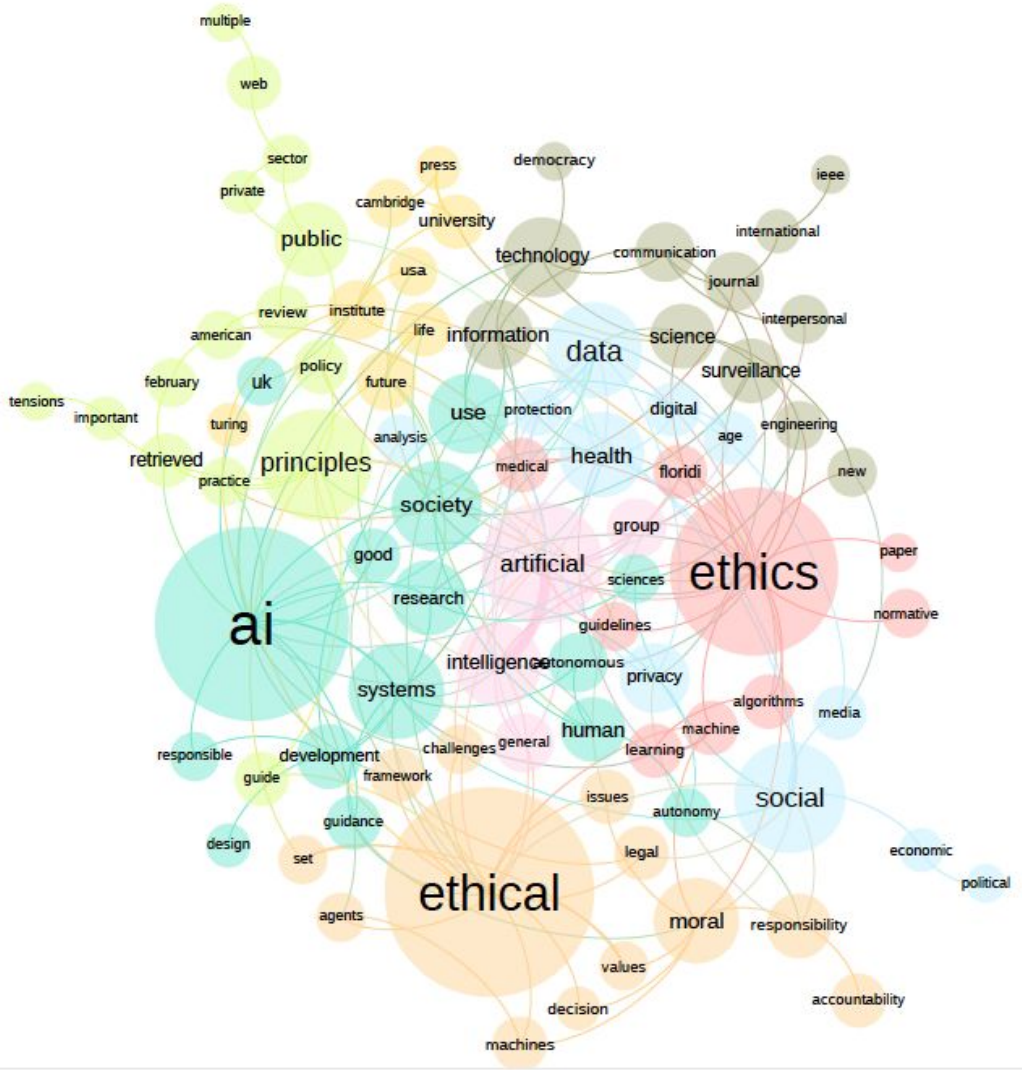
Additional analysis: detecting the communities

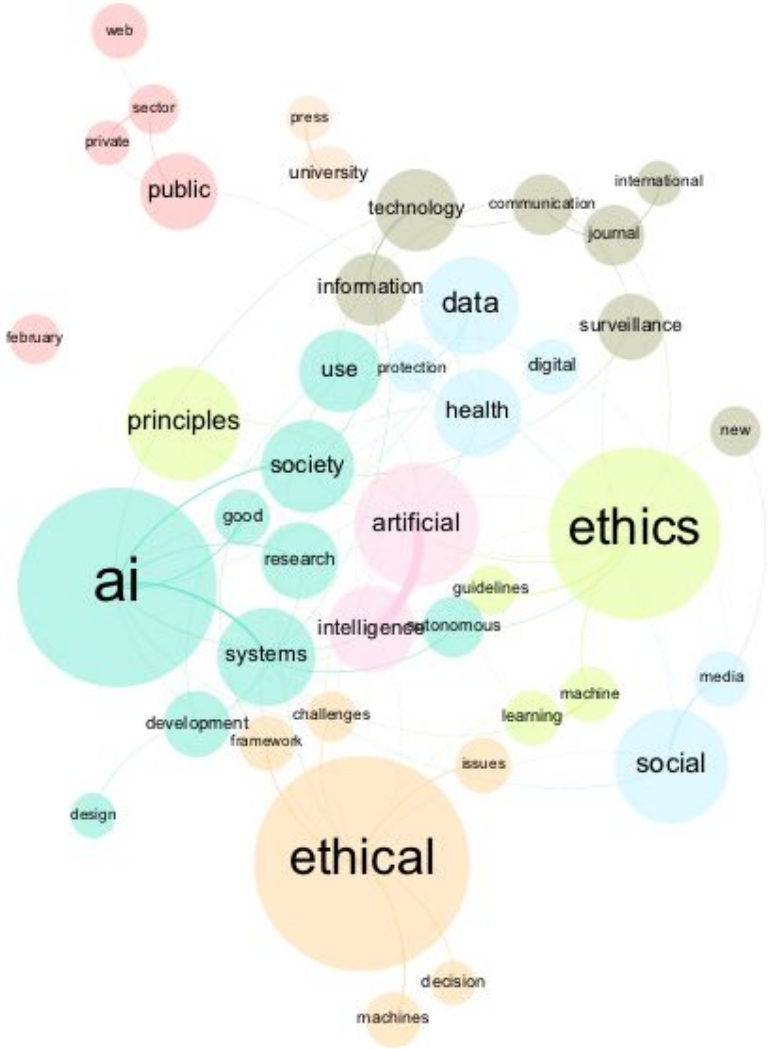


Modularity: 0.558

Modularity with resolution: 0.558

Number of Communities: 13


multiple
web
sector
private
public
press
cambridge
university
democracy
ieee
international
technology
communication
journal
usa
review
institute
interpersonal
american
life
information
data
science
surveillance
policy
february
uk
future
tensions
important
turing
use
protection
digital
engineering
analysis
age
retrieved
principles
practice
medical
health
floridi
new
society
group
good
artificial
ethics
paper
sciences
research
normative
guidelines
ai
intelligence
autonomous
privacy
systems
algorithms
media
machine
human
responsible
development
challenges
general
learning
framework
guide
issues
social
autonomy
guidance
design
set
legal
economic
political
ethical
agents
moral
responsibility
values
accountability
decision
machines


web
sector
private
public
press
university
technology
communication
international
journal
information
data
surveillance
february
use
protection
digital
principles
health
new
society
good
artificial
ethics
ai
research
guidelines
intelligence
autonomous
systems
media
machine
challenges
learning
development
framework
issues
social
design
ethical
decision
machines

Thank you for your attention!