

# An Overview of Research Funding Programs that Support Sustainability-related Research in Flanders

Amr Ali-Eldin [1], Hoang-Son Pham [1], Mehmet Ali Abdulhayoglu [2] and Bart Thijs [2]

[1] ECOOM-Interoperability, UHasselt
[2] ECOOM-Biblio-Leuven, KU Leuven

UHASSELT
KNOWLEDGE IN ACTION

Ecoom-beleidsdag over Duurzaam Ondernemen en Innoveren,
Brussel, October 25th, 2024

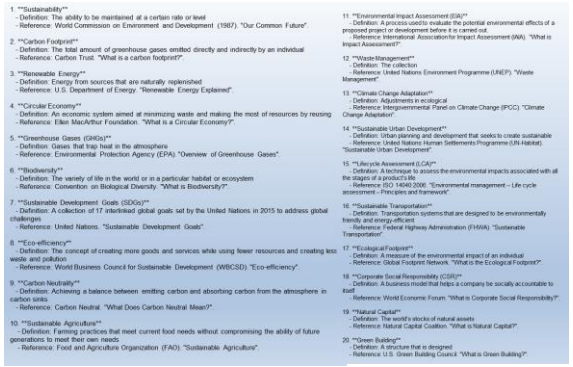
# Introduction

- Sustainability-related research addresses global challenges such as climate change, resource reduction, environmental degradation, social inequality, and economic instability.
- Identifying research projects related to sustainable development requires the ability to determine if a project aligns with sustainability principles.
- This typically involves analyzing project descriptions, abstracts, and keywords to determine if they address issues like renewable energy, circular economy, climate action, or other sustainability themes.
- Traditional methods may depend on human interpretation to categorize projects as sustainability-related, leading to subjectivity and inconsistencies, especially across interdisciplinary research areas.



# Research Approach

- Using advanced techniques like large language models (LLMs) and semantic embeddings to assess Projects' alignment and map that to related funding programs based on:
  - sustainability keywords and definitions
  - sustainability development goals (SDG)



1. **Sustainability**
   - Definition: The ability to be maintained at a certain rate or level
   - Reference: World Commission on Environment and Development (1987). "Our Common Future".

2. **Carbon Footprint**
   - Definition: The total amount of greenhouse gases emitted directly and indirectly by an individual
   - Reference: Carbon Trust. "What is a carbon footprint?".

3. **Renewable Energy**
   - Definition: Energy from sources that are naturally replenished
   - Reference: U.S. Department of Energy. "Renewable Energy Explained".

4. **Circular Economy**
   - Definition: An economic system aimed at minimizing waste and making the most of resources by reusing
   - Reference: Ellen MacArthur Foundation. "What is a Circular Economy?".

5. **Greenhouse Gases (GHGs)**
   - Definition: Gases that trap heat in the atmosphere
   - Reference: Environmental Protection Agency (EPA). "Overview of Greenhouse Gases".

6. **Biodiversity**
   - Definition: The variety of life in the world or in a particular habitat or ecosystem
   - Reference: Convention on Biological Diversity. "What is Biodiversity?".

7. **Sustainable Development Goals (SDGs)**
   - Definition: A collection of 17 interlinked global goals set by the United Nations in 2015 to address global challenges
   - Reference: United Nations. "Sustainable Development Goals".

8. **Eco-efficiency**
   - Definition: The concept of creating more goods and services while using fewer resources and creating less waste and pollution
   - Reference: World Business Council for Sustainable Development (WBCSD). "Eco-efficiency".

9. **Carbon Neutrality**
   - Definition: Achieving a balance between emitting carbon and absorbing carbon from the atmosphere in carbon sinks
   - Reference: Carbon Neutral. "What Does Carbon Neutral Mean?".

10. **Sustainable Agriculture**
    - Definition: Farming practices that meet current food needs without compromising the ability of future generations to meet their own needs
    - Reference: Food and Agriculture Organization (FAO). "Sustainable Agriculture".

11. **Environmental Impact Assessment (EIA)**
    - Definition: A process used to evaluate the potential environmental effects of a proposed project or development before it is carried out.
    - Reference: International Association for Impact Assessment (IAIA). "What is Impact Assessment?".

12. **Waste Management**
    - Definition: The collection
    - Reference: United Nations Environment Programme (UNEP). "Waste Management".

13. **Climate Change Adaptation**
    - Definition: Adjustments in ecological
    - Reference: Intergovernmental Panel on Climate Change (IPCC). "Climate Change Adaptation".

14. **Sustainable Urban Development**
    - Definition: Urban planning and development that seeks to create sustainable
    - Reference: United Nations Human Settlements Programme (UN-Habitat). "Sustainable Urban Development".

15. **Lifecycle Assessment (LCA)**
    - Definition: A technique to assess the environmental impacts associated with all the stages of a product's life
    - Reference: ISO 14040:2006. "Environmental management – Life cycle assessment – Principles and framework".

16. **Sustainable Transportation**
    - Definition: Transportation systems that are designed to be environmentally friendly and energy-efficient
    - Reference: Federal Highway Administration (FHWA). "Sustainable Transportation".

17. **Ecological Footprint**
    - Definition: A measure of the environmental impact of an individual
    - Reference: Global Footprint Network. "What is the Ecological Footprint?".

18. **Corporate Social Responsibility (CSR)**
    - Definition: A business model that helps a company be socially accountable to itself
    - Reference: World Economic Forum. "What is Corporate Social Responsibility?".

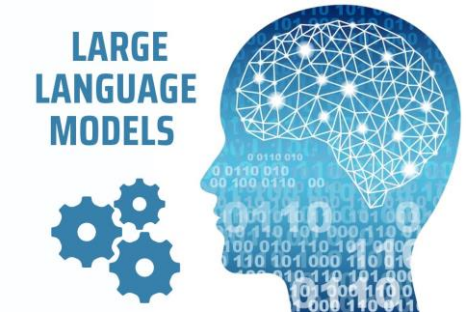
19. **Natural Capital**
    - Definition: The world's stocks of natural assets
    - Reference: Natural Capital Coalition. "What is Natural Capital?".

20. **Green Building**
    - Definition: A structure that is designed
    - Reference: U.S. Green Building Council. "What is Green Building?"



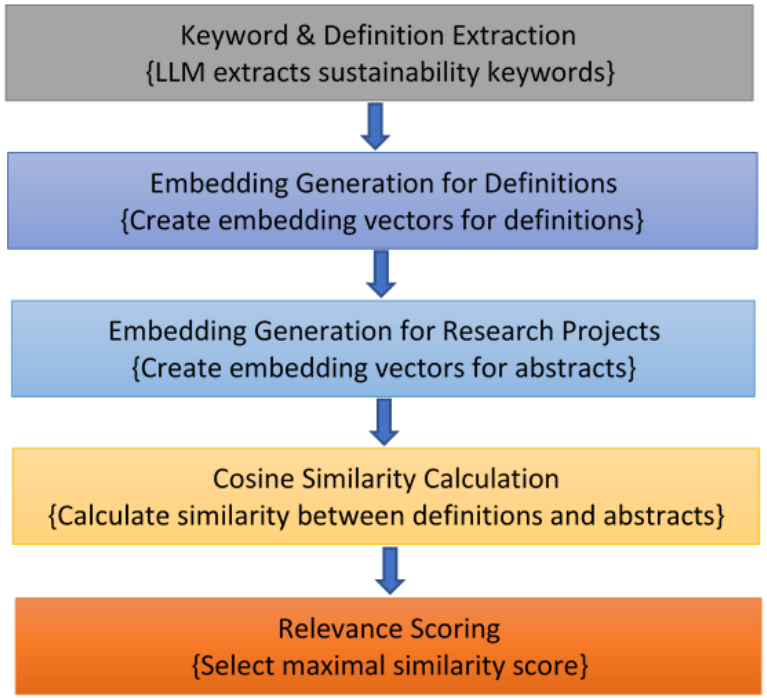
# Research Approach

- Large language models (LLMs) are advanced models trained on massive amounts of data, enabling them to understand and generate natural language and perform various tasks.
- They enable efficient analysis of large text volumes, suitable for diverse research environments.
- Semantic embedding transforms text into numerical vectors, capturing semantics for easier machine understanding and manipulation.



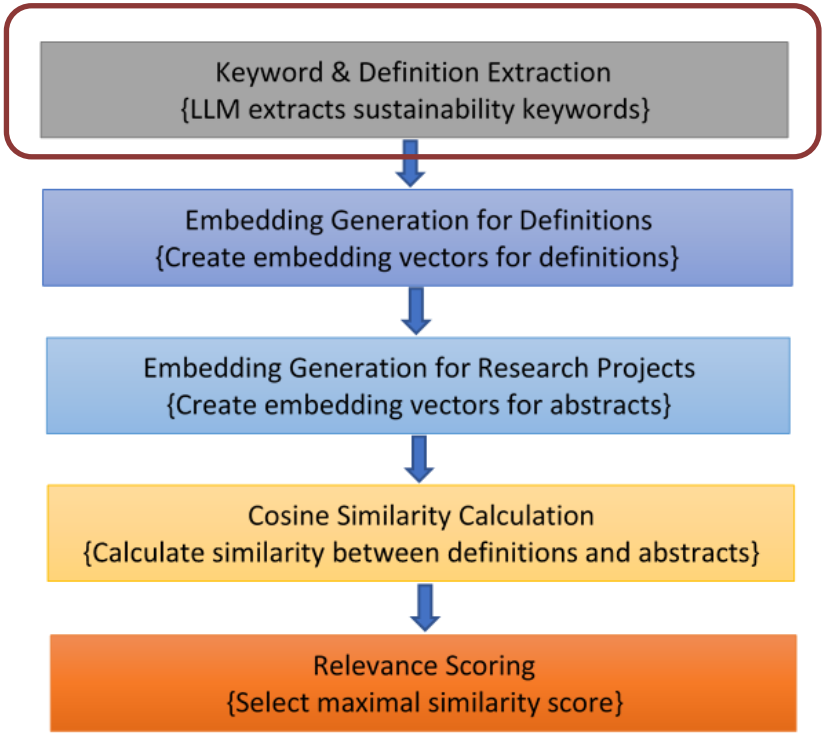
# The Sustainability Keywords and Definitions Approach

- This approach allows for a scalable and efficient method of identifying sustainability-related projects by analyzing the semantic alignment between project abstracts and sustainability keywords and their definitions.



# Keyword and Definition Extraction

- The LLM extract sustainability keywords and their definitions, creating a strong vocabulary for later research comparison.
- In this context, OpenAI was used a number of times to provide a list of sustainability keywords and their definitions.



# Keywords and definitions

1. **Sustainability**
   - Definition: The ability to be maintained at a certain rate or level
   - Reference: World Commission on Environment and Development (1987). "Our Common Future".

2. **Carbon Footprint**
   - Definition: The total amount of greenhouse gases emitted directly and indirectly by an individual
   - Reference: Carbon Trust. "What is a carbon footprint?".

3. **Renewable Energy**
   - Definition: Energy from sources that are naturally replenished
   - Reference: U.S. Department of Energy. "Renewable Energy Explained".

4. **Circular Economy**
   - Definition: An economic system aimed at minimizing waste and making the most of resources by reusing
   - Reference: Ellen MacArthur Foundation. "What is a Circular Economy?".

5. **Greenhouse Gases (GHGs)**
   - Definition: Gases that trap heat in the atmosphere
   - Reference: Environmental Protection Agency (EPA). "Overview of Greenhouse Gases".

6. **Biodiversity**
   - Definition: The variety of life in the world or in a particular habitat or ecosystem
   - Reference: Convention on Biological Diversity. "What is Biodiversity?".

7. **Sustainable Development Goals (SDGs)**
   - Definition: A collection of 17 interlinked global goals set by the United Nations in 2015 to address global challenges
   - Reference: United Nations. "Sustainable Development Goals".

8. **Eco-efficiency**
   - Definition: The concept of creating more goods and services while using fewer resources and creating less waste and pollution
   - Reference: World Business Council for Sustainable Development (WBCSD). "Eco-efficiency".

9. **Carbon Neutrality**
   - Definition: Achieving a balance between emitting carbon and absorbing carbon from the atmosphere in carbon sinks
   - Reference: Carbon Neutral. "What Does Carbon Neutral Mean?".

10. **Sustainable Agriculture**
    - Definition: Farming practices that meet current food needs without compromising the ability of future generations to meet their own needs
    - Reference: Food and Agriculture Organization (FAO). "Sustainable Agriculture".

11. **Environmental Impact Assessment (EIA)**
    - Definition: A process used to evaluate the potential environmental effects of a proposed project or development before it is carried out.
    - Reference: International Association for Impact Assessment (IAIA). "What is Impact Assessment?".

12. **Waste Management**
    - Definition: The collection
    - Reference: United Nations Environment Programme (UNEP). "Waste Management".

13. **Climate Change Adaptation**
    - Definition: Adjustments in ecological
    - Reference: Intergovernmental Panel on Climate Change (IPCC). "Climate Change Adaptation".

14. **Sustainable Urban Development**
    - Definition: Urban planning and development that seeks to create sustainable
    - Reference: United Nations Human Settlements Programme (UN-Habitat). "Sustainable Urban Development".

15. **Lifecycle Assessment (LCA)**
    - Definition: A technique to assess the environmental impacts associated with all the stages of a product's life
    - Reference: ISO 14040:2006. "Environmental management – Life cycle assessment – Principles and framework".

16. **Sustainable Transportation**
    - Definition: Transportation systems that are designed to be environmentally friendly and energy-efficient
    - Reference: Federal Highway Administration (FHWA). "Sustainable Transportation".

17. **Ecological Footprint**
    - Definition: A measure of the environmental impact of an individual
    - Reference: Global Footprint Network. "What is the Ecological Footprint?".

18. **Corporate Social Responsibility (CSR)**
    - Definition: A business model that helps a company be socially accountable to itself
    - Reference: World Economic Forum. "What is Corporate Social Responsibility?".

19. **Natural Capital**
    - Definition: The world's stocks of natural assets
    - Reference: Natural Capital Coalition. "What is Natural Capital?".

20. **Green Building**
    - Definition: A structure that is designed
    - Reference: U.S. Green Building Council. "What is Green Building?".

# Embedding Generation for Extracted Keywords & Definitions

- After extracting sustainability definitions, we convert them into numerical representations, called embedding vectors, using the OpenAI ‘’text-embedding-3-large’’ embedding model.

```
# create embedding vector stores

    # create openai embedding model

    from langchain_openai import OpenAIEmbeddings

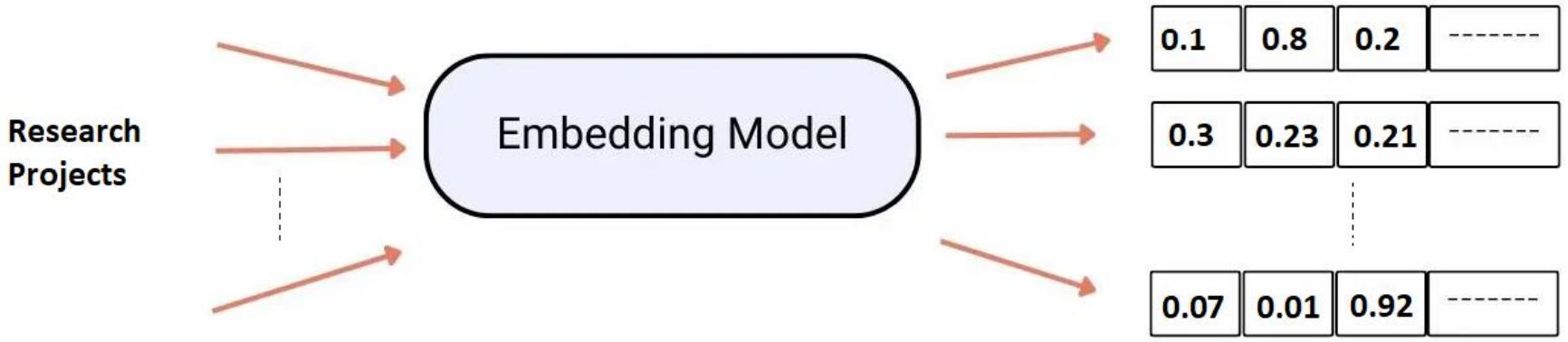
    embedding_model = OpenAIEmbeddings(model="text-embedding-3-large")
```

- Embedding vectors capture the semantic meaning of the definitions in a high-dimensional space, enabling us to capture complex relationships between words.
- This is an important step as it allows us to transform the textual definitions into a format that can be quantitatively compared with research project abstracts.



# Embedding Generation for Research Projects

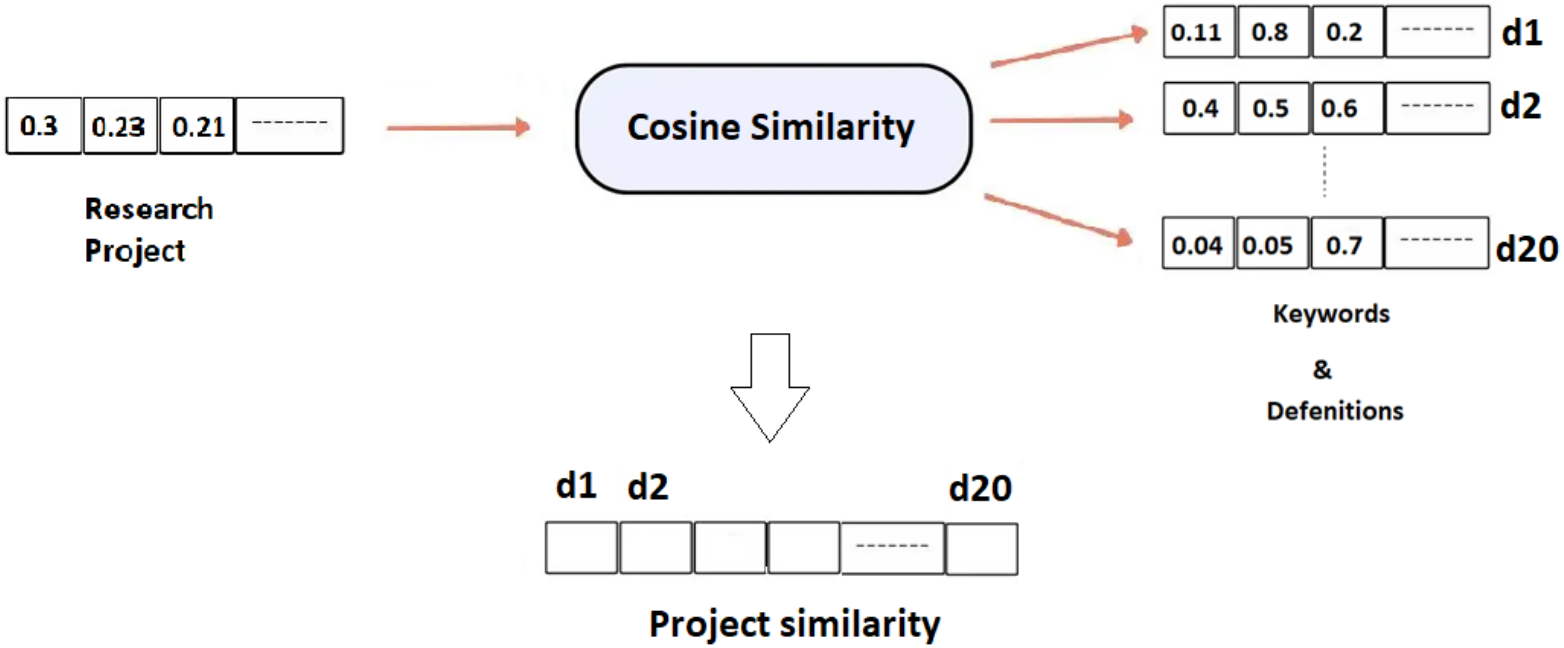
- For every research project, we gather the most important details: the title, keywords, and abstract.
- The abstract is especially crucial as it provides a summary of the project's scope, objectives, and methodology.
- We then use the same embedding model to generate vectors for these abstracts.
- These resulting vectors capture the essential meaning of each research project in a format that can be compared to sustainability definitions.



# Cosine Similarity Calculation

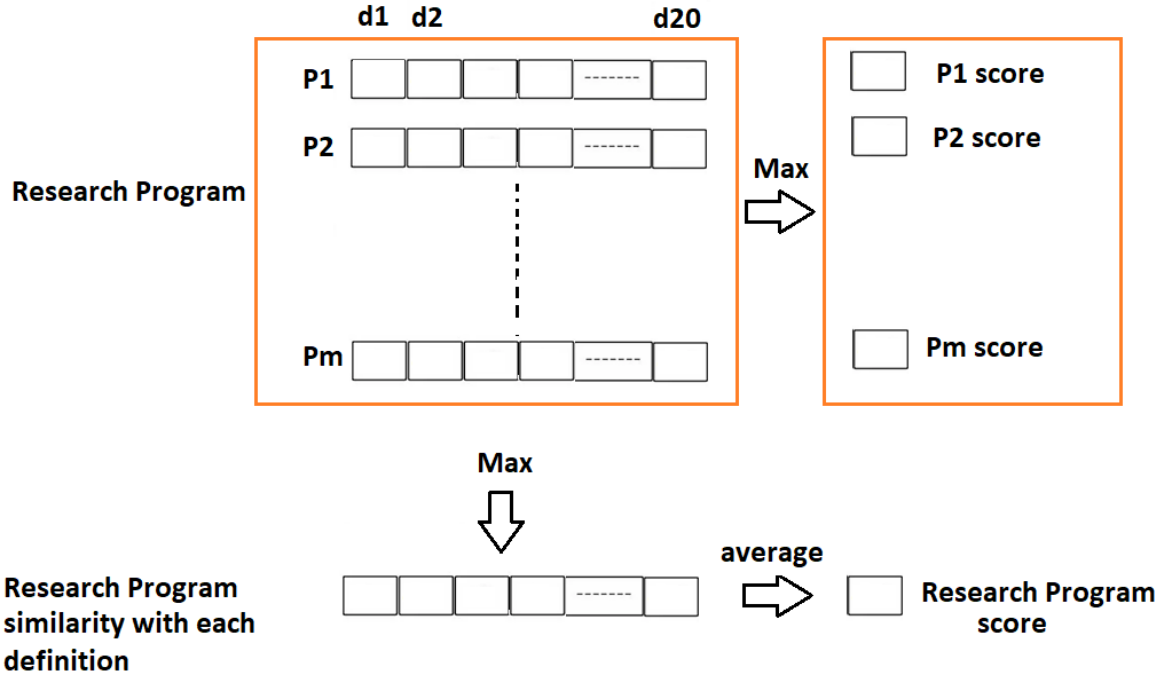
- For each project, we compare the abstract's embedding vector to each of the sustainability definition vectors.
- Given two n-dimensional embedding vectors A and B, the cosine similarity, cos(θ), is as follows:

$$\text{cosine similarity} = S_C(A,B) := \cos(\theta) = \frac{\mathbf{A}\cdot\mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\cdot\sqrt{\sum_{i=1}^{n} B_i^2}},$$

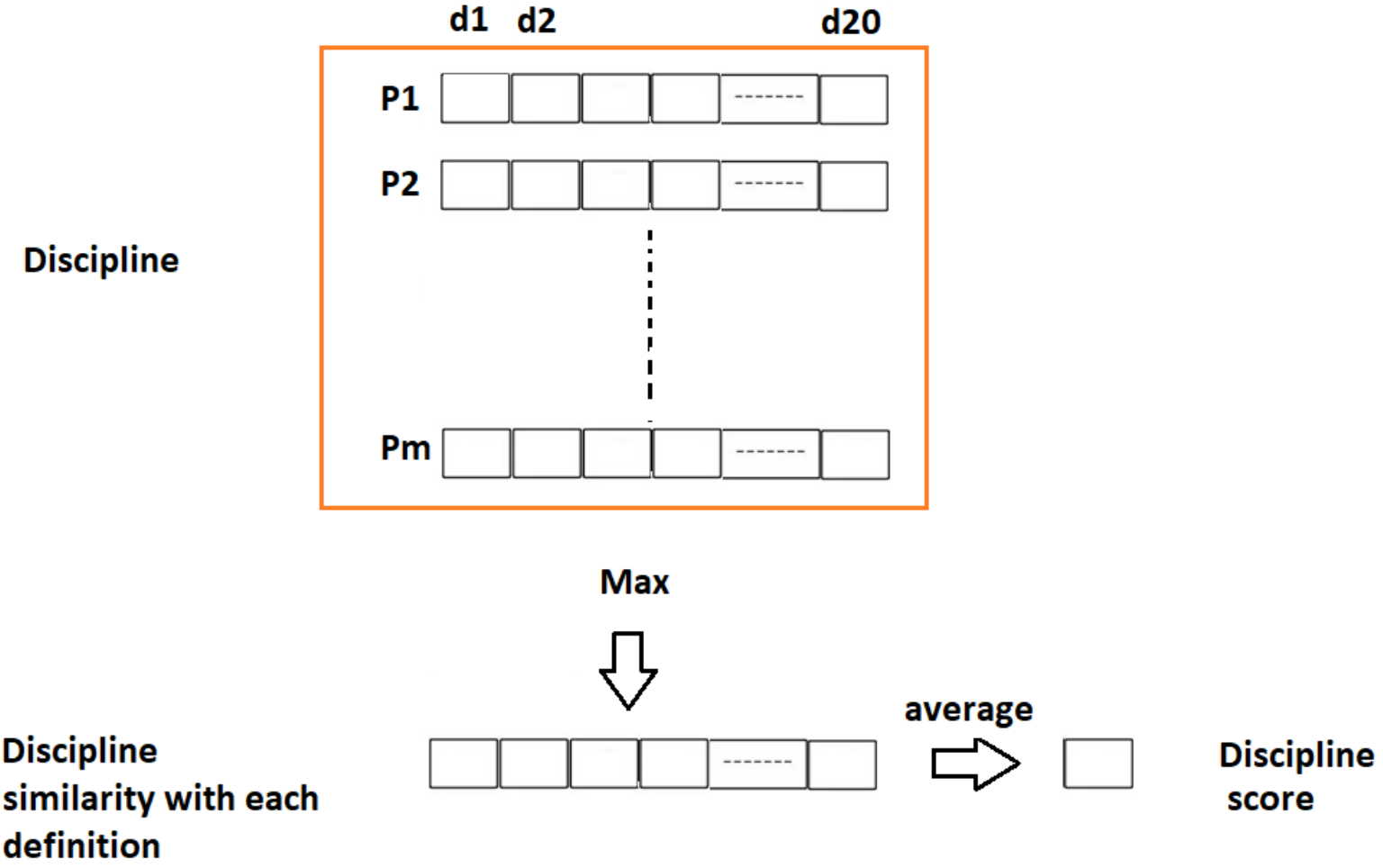


# Relevance Scoring

- In order to evaluate the relevance of each project to sustainability, we calculate the highest cosine similarity score between the project's abstract and a set of sustainability definitions.
- This score indicates how closely the project aligns with sustainability concepts.
- Projects with higher scores are considered to be more relevant to sustainability.



# Relevance Scoring



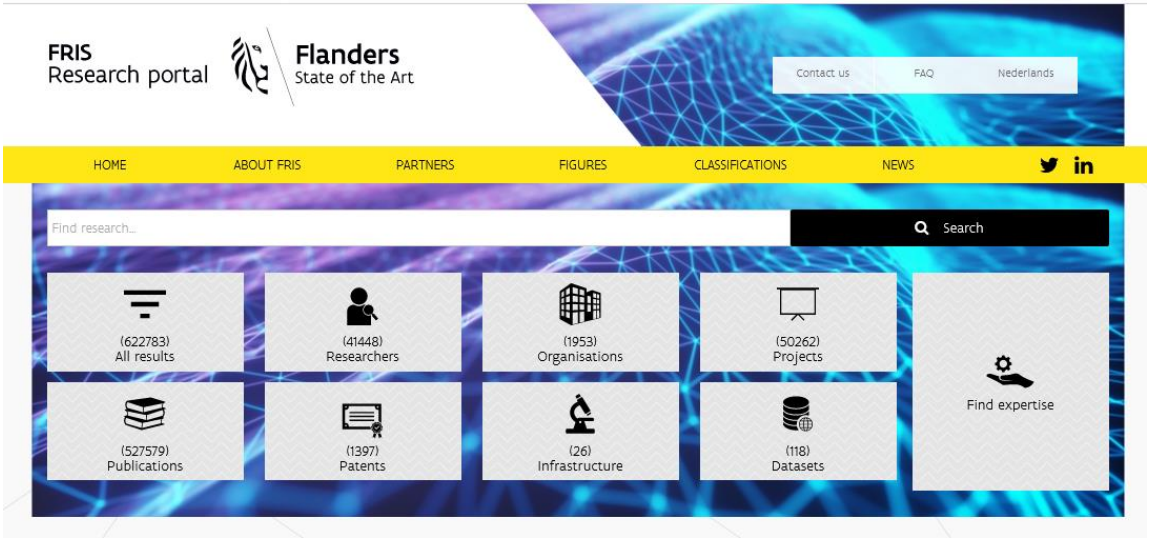
# Experimental Work

- **Project dataset selection:**

— 16942 projects with abstracts and funding information were selected from FRIS for this analysis (projects from 2019 – 2024).

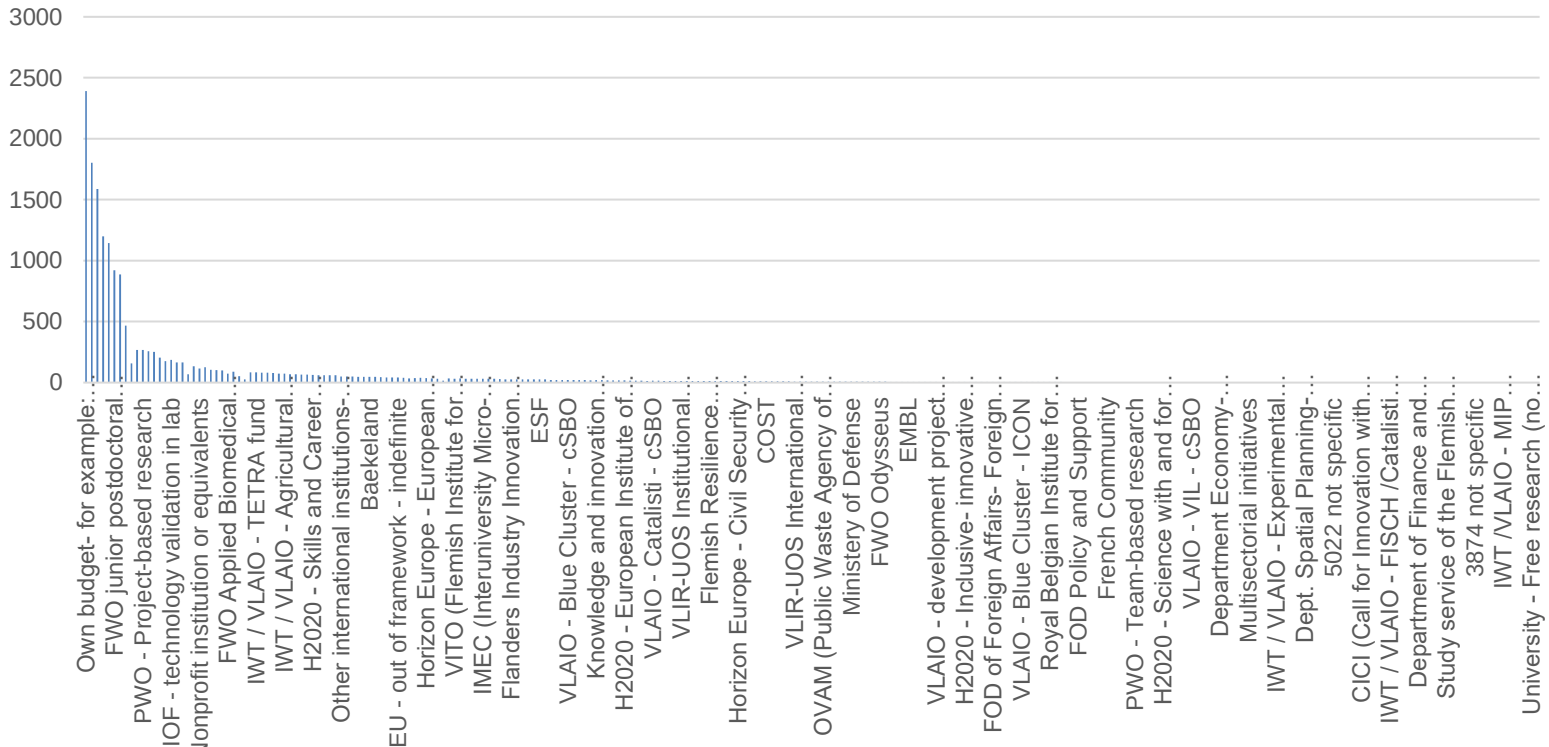
— Each project's metadata includes: id, title, keywords, abstract, principle funding codes, complementary funding codes, disciplines

- For each project, we calculate the cosine similarity between the embedding vector of abstract and embedding vectors of sustainability definitions.



# Research Projects Distribution Over Research Programs

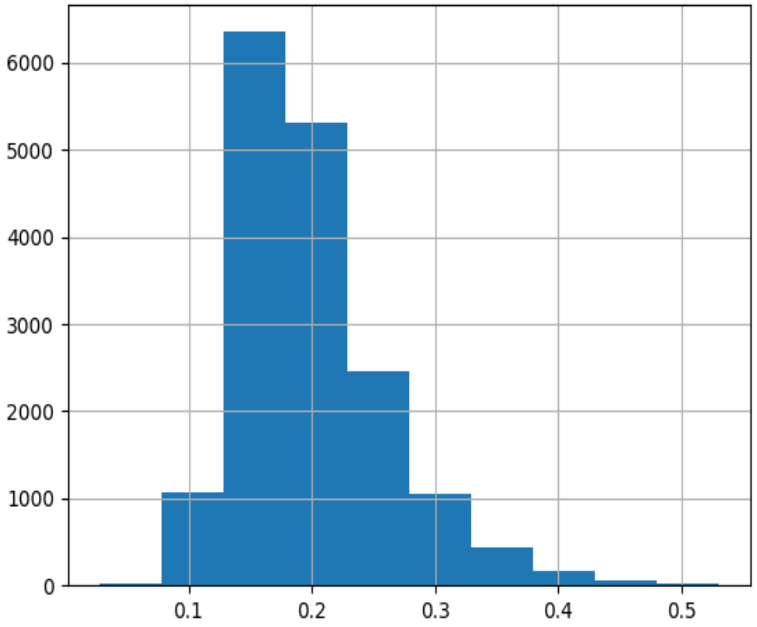
- Funding programs with higher scores are considered to have a greater focus on sustainable research compared to those with lower scores.
- In total, 257 funding programs were associated with the projects.
- The figure shows the number of projects by funding program. For readability, we selected funding programs associated with more than 100 projects.

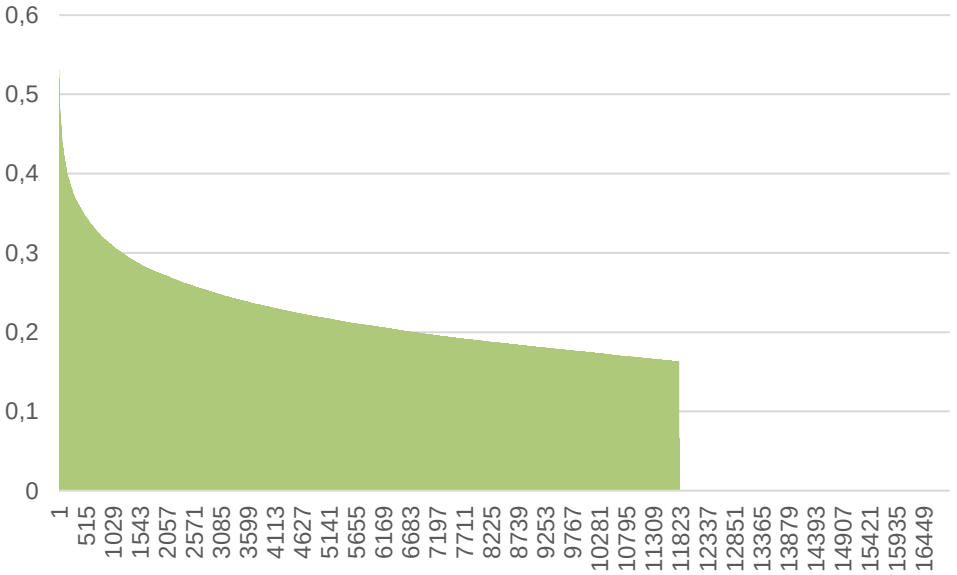



number of projects by funding programs

# Projects Similarity Distribution

- Average similarity scores for all projects was 0.2
- The maximal similarity score was 0.530, and the minimal similarity value was 0.03.



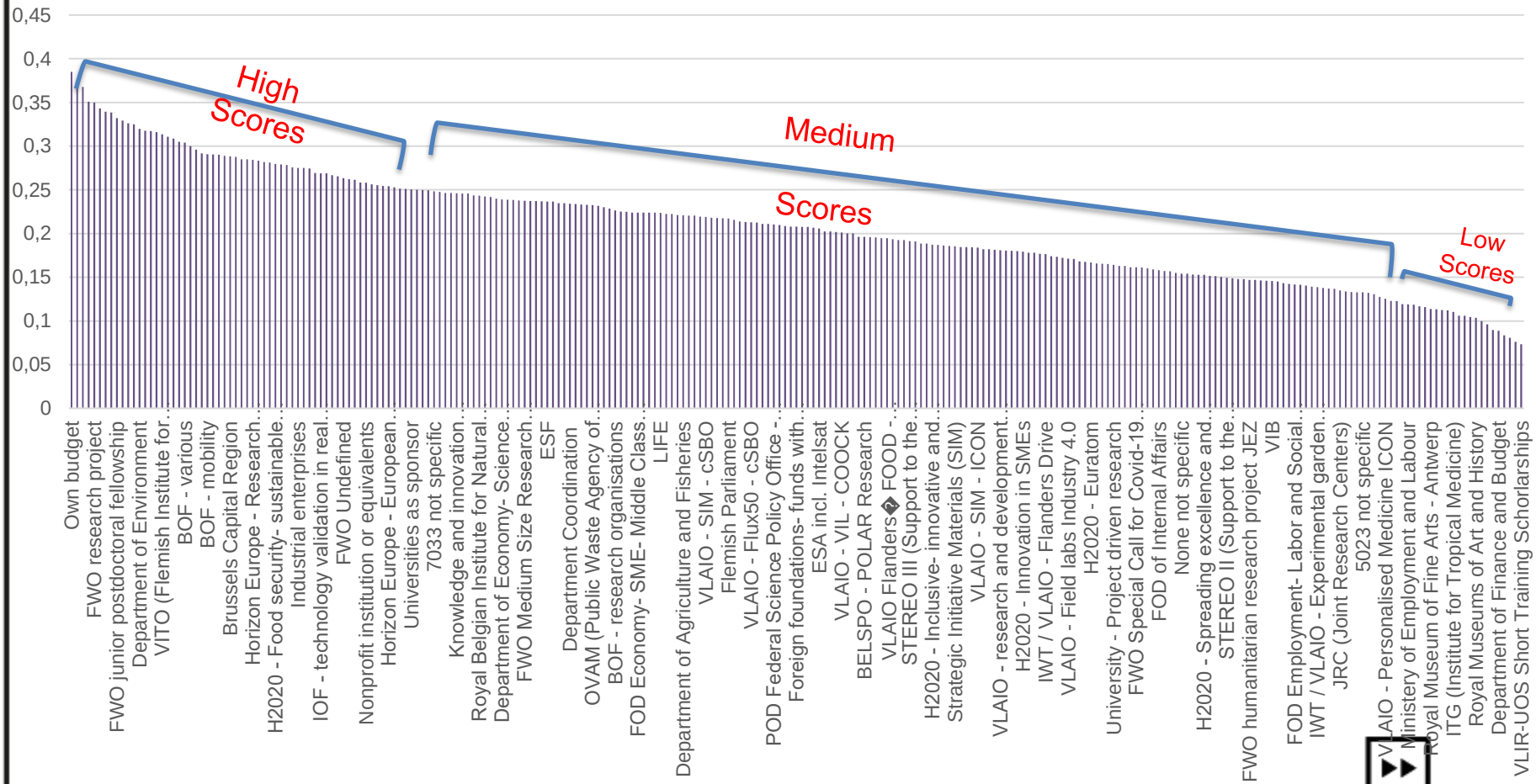

Similarity values distribution

ecoom

UHASSELT
KNOWLEDGE IN ACTION

# Sustainability research programs by funding program

- Funding programs such as ‘Own budget’, ‘BOF – Projects’, ‘BOF – Doctoral Mandates’, ‘FWO Strategic Basic Research Grant’ and ‘FWO research project’, achieved high scores compared to other programs.

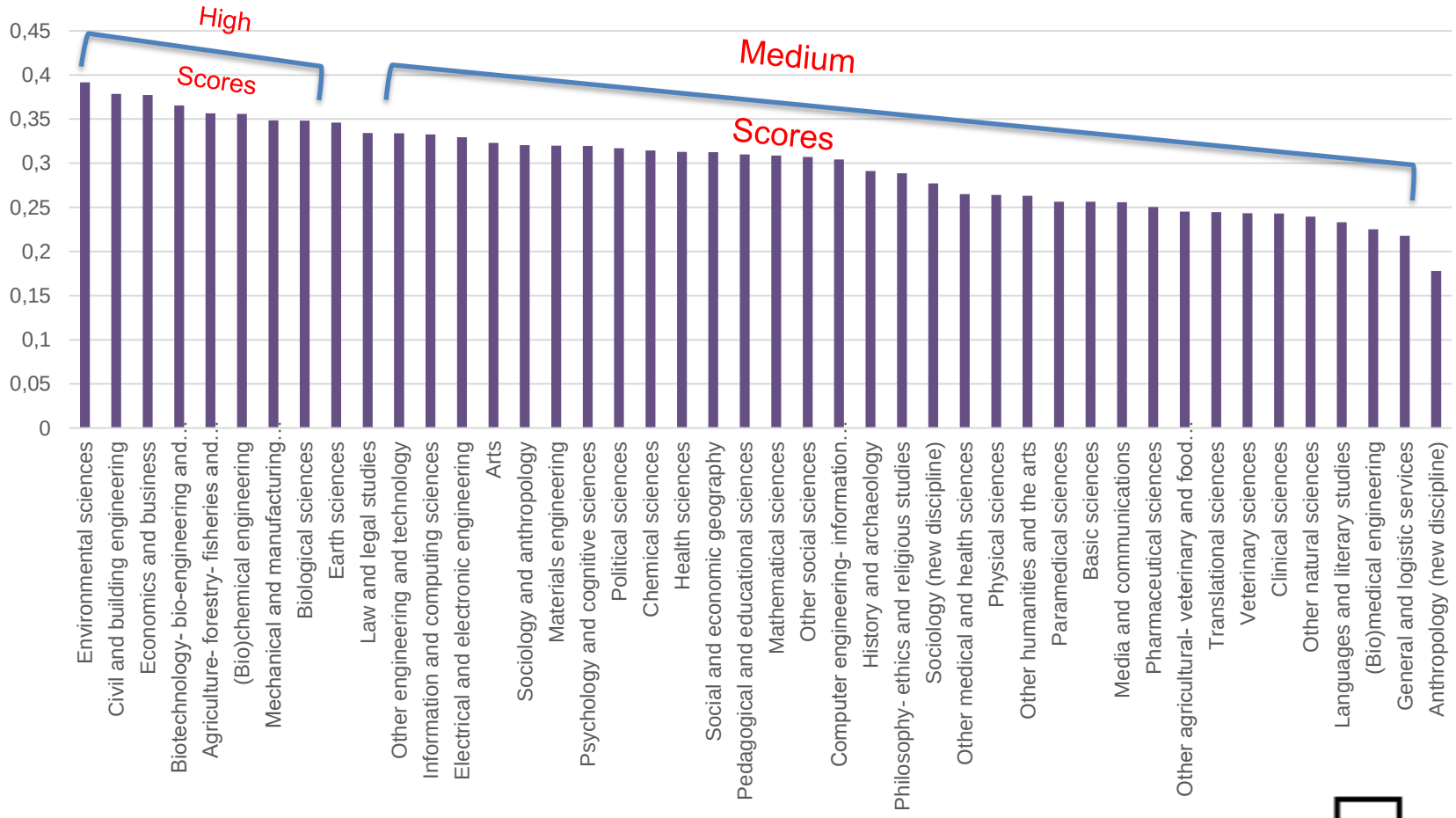



ecoom

UHASSELT
KNOWLEDGE IN ACTION

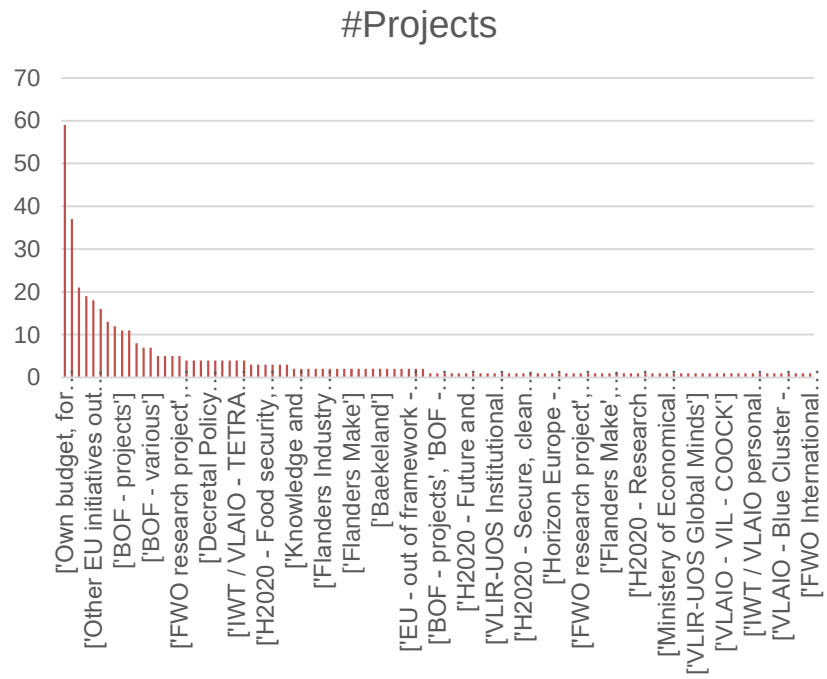
# Sustainability by disciplines

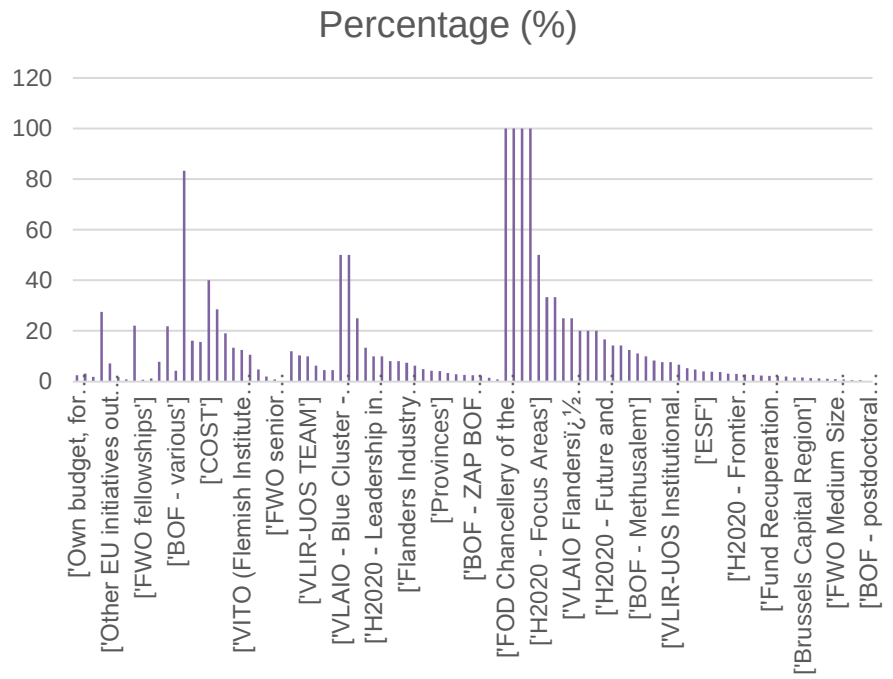
- Projects associated with Environmental sciences achieved the highest average similarity score, followed by civil and building engineering, and economics and business.




Average similarity values by disciplines

# Research Funding programs with most relevant projects




ecoom

UHASSELT
KNOWLEDGE IN ACTION

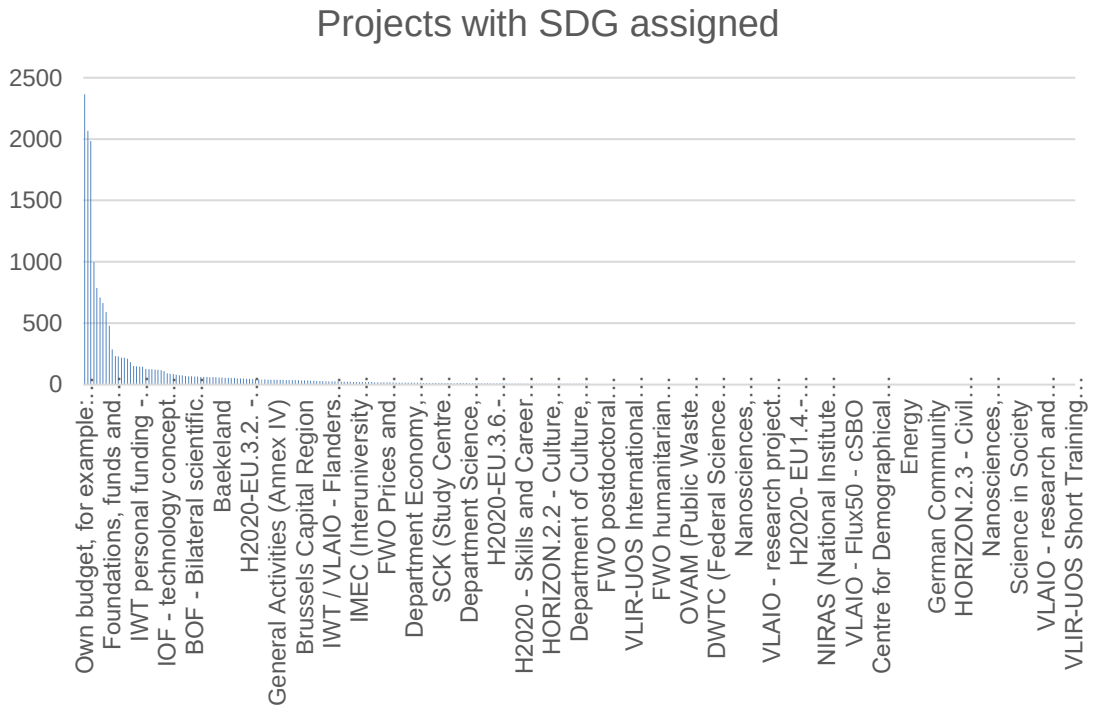
# The Sustainable Development Goals (SDGs)

- a set of 17 global objectives established by the United Nations in 2015.
- to address global challenges and promote a sustainable future for all
- Promote well-being, safeguard the planet and ensure peace and prosperity by 2030.



# Funding Programs Evaluation based on SDGs

- In this work, sustainable development goals were used to identify sustainability research projects.
- the results are based on several runs of two LLM Meta's Llama 3.1 and Mistral-Nemo.
- Own budget, FWO fellowships, FWO research project, and FWO Strategic Basic Research Grant involve a large number of projects with SDGs assigned.



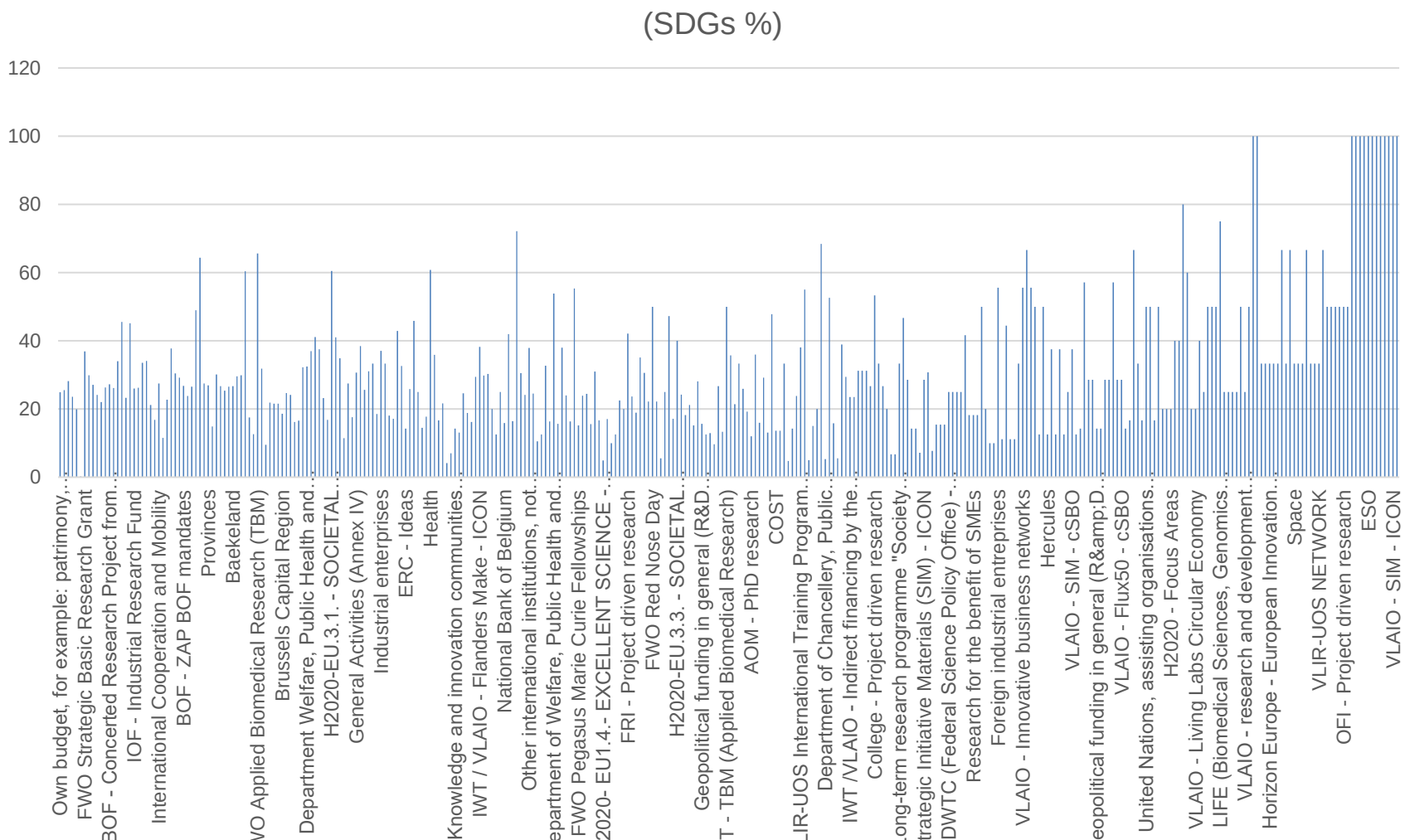

Dataset

| | |
|---|---|
| Projects with abstracts | 65149 |
| Projects with Funding info | 74850 |
| Projects with Funding And Abstract | 62924 |
| Projects with Funding / Abstract and SDG Assigned | 17025 |

ecoom

UHASSELT
KNOWLEDGE IN ACTION

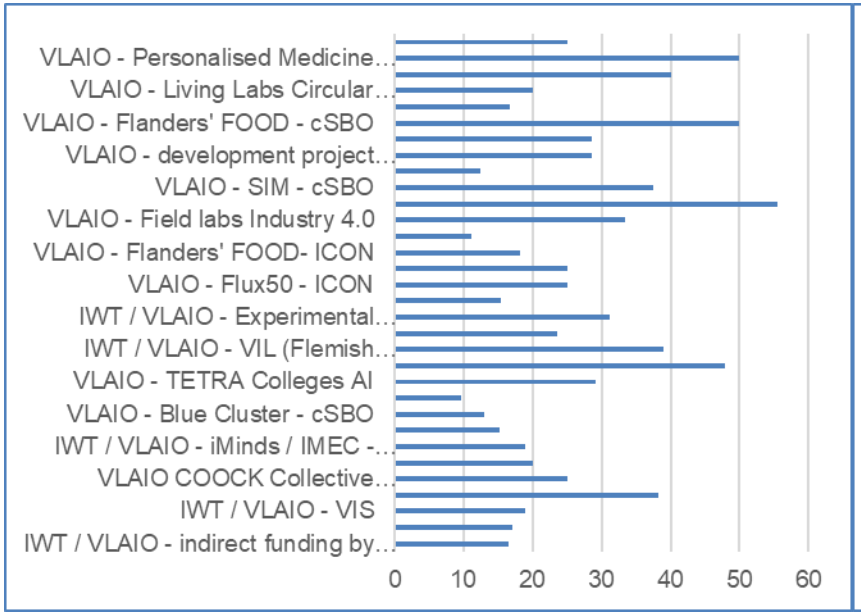
# Funding Programs Project Shares with SDGs Assigned



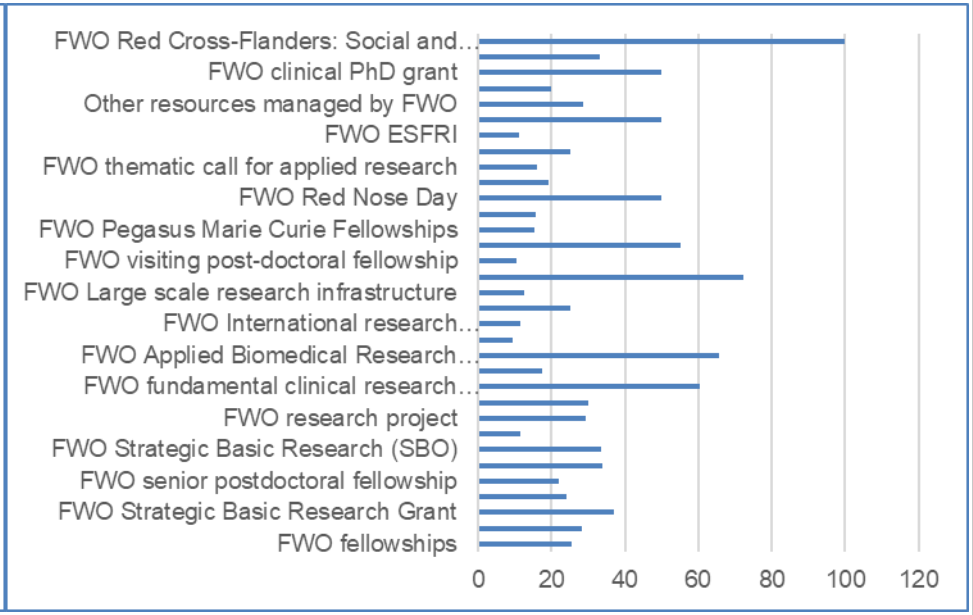

ecoom

UHASSELT
KNOWLEDGE IN ACTION

# SDGs Shares assigned within Programs




343 assigned SDGs / 1442 Projects
23,8%

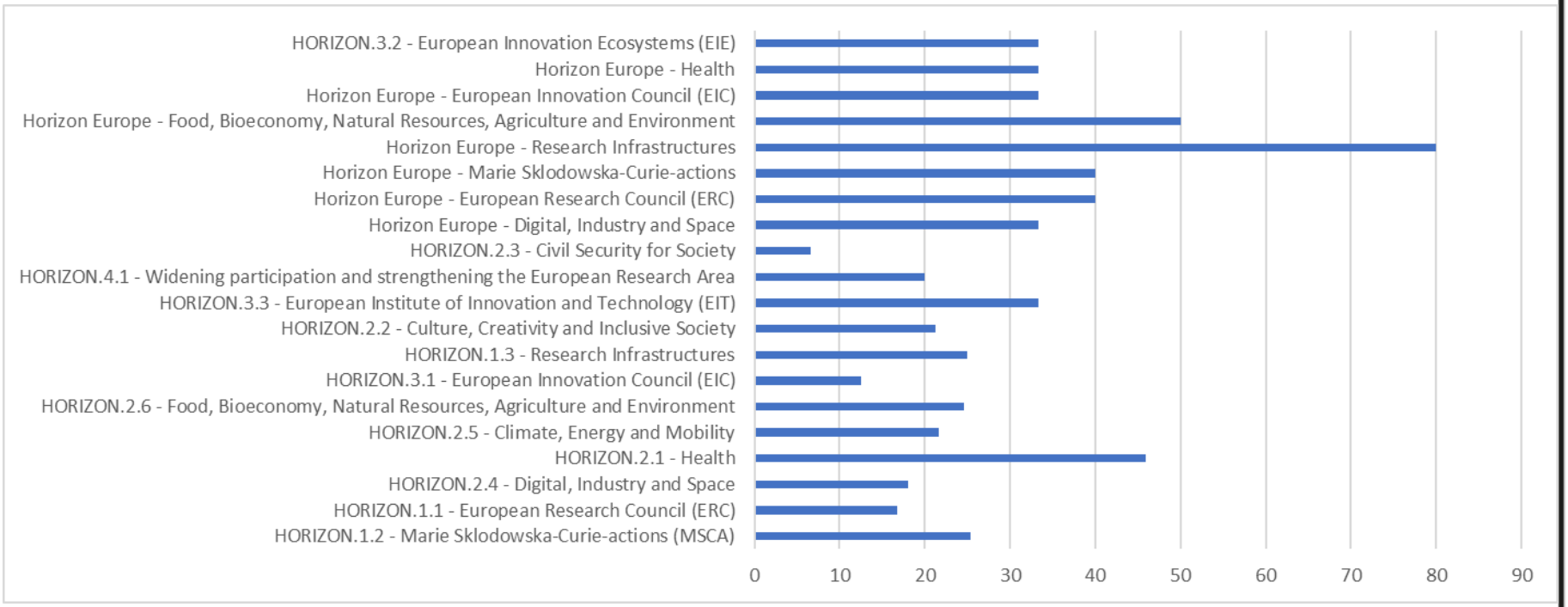



6778 assigned SDGs / 24169 Projects
28%

ecoom

UHASSELT
KNOWLEDGE IN ACTION

# SDGs Shares assigned within Programs




219 assigned SDGs / 892 Projects
24,56%

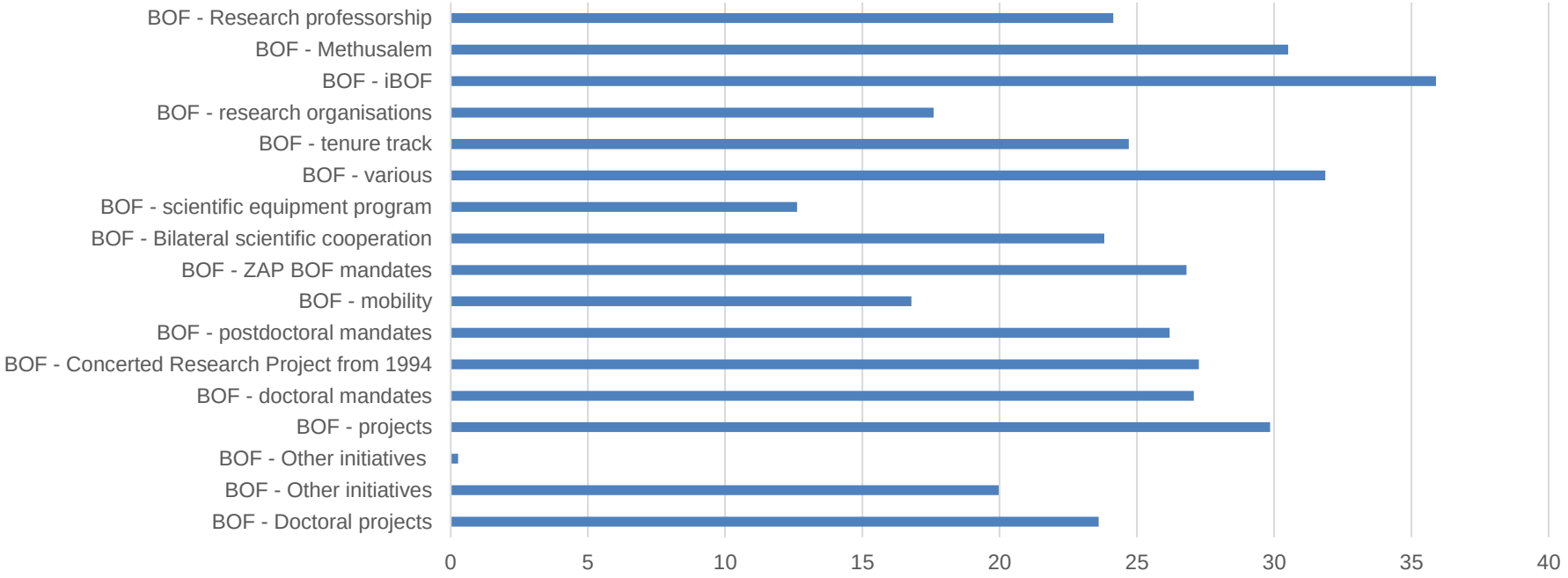
**Own Budget**

2363 assigned SDGs / 9487 Projects
24,9 %

ecoom

UHASSELT
KNOWLEDGE IN ACTION

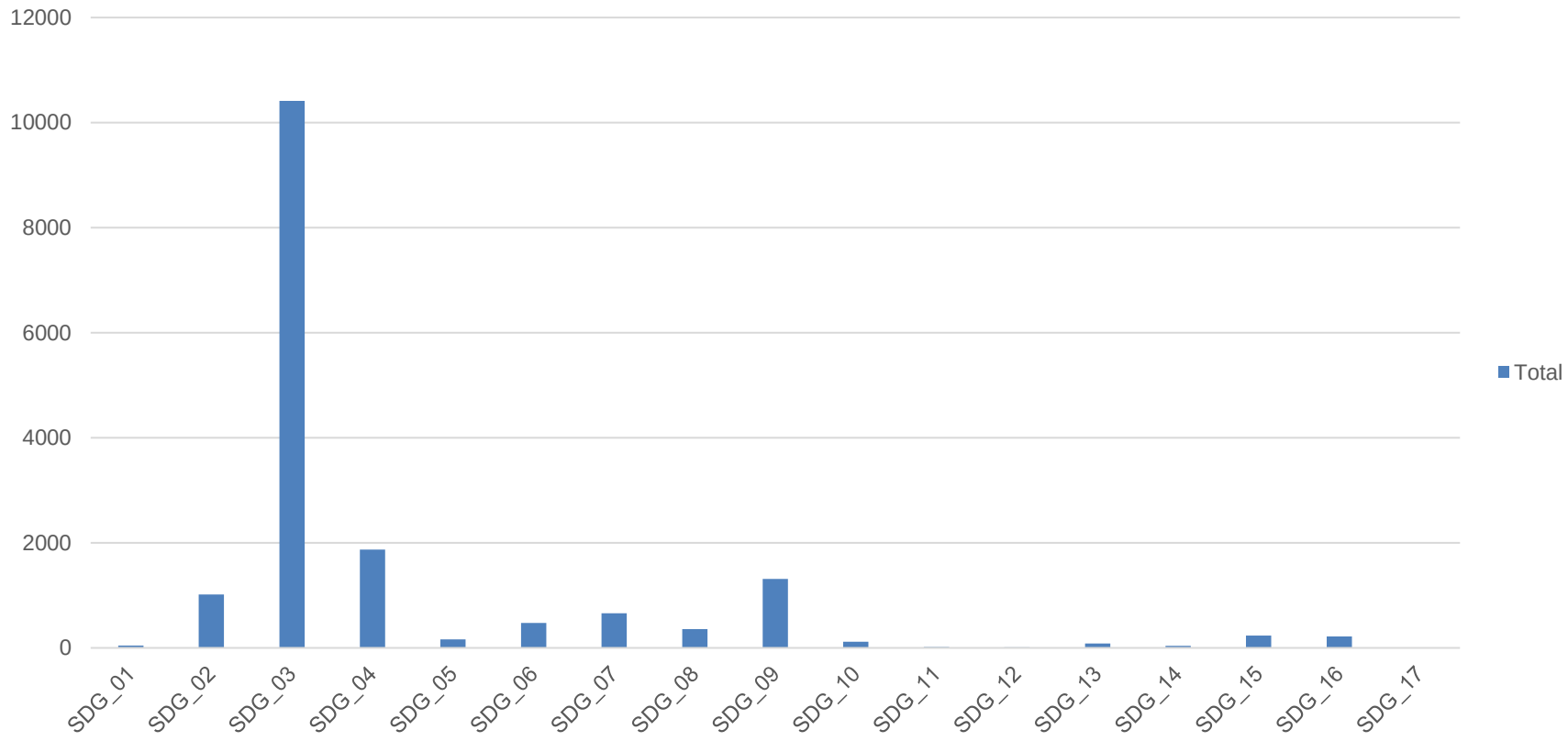
# SDGs Shares assigned within Programs




3564 assigned SDGs / 17242 Projects
20,07 %

# SDGs distribution over projects

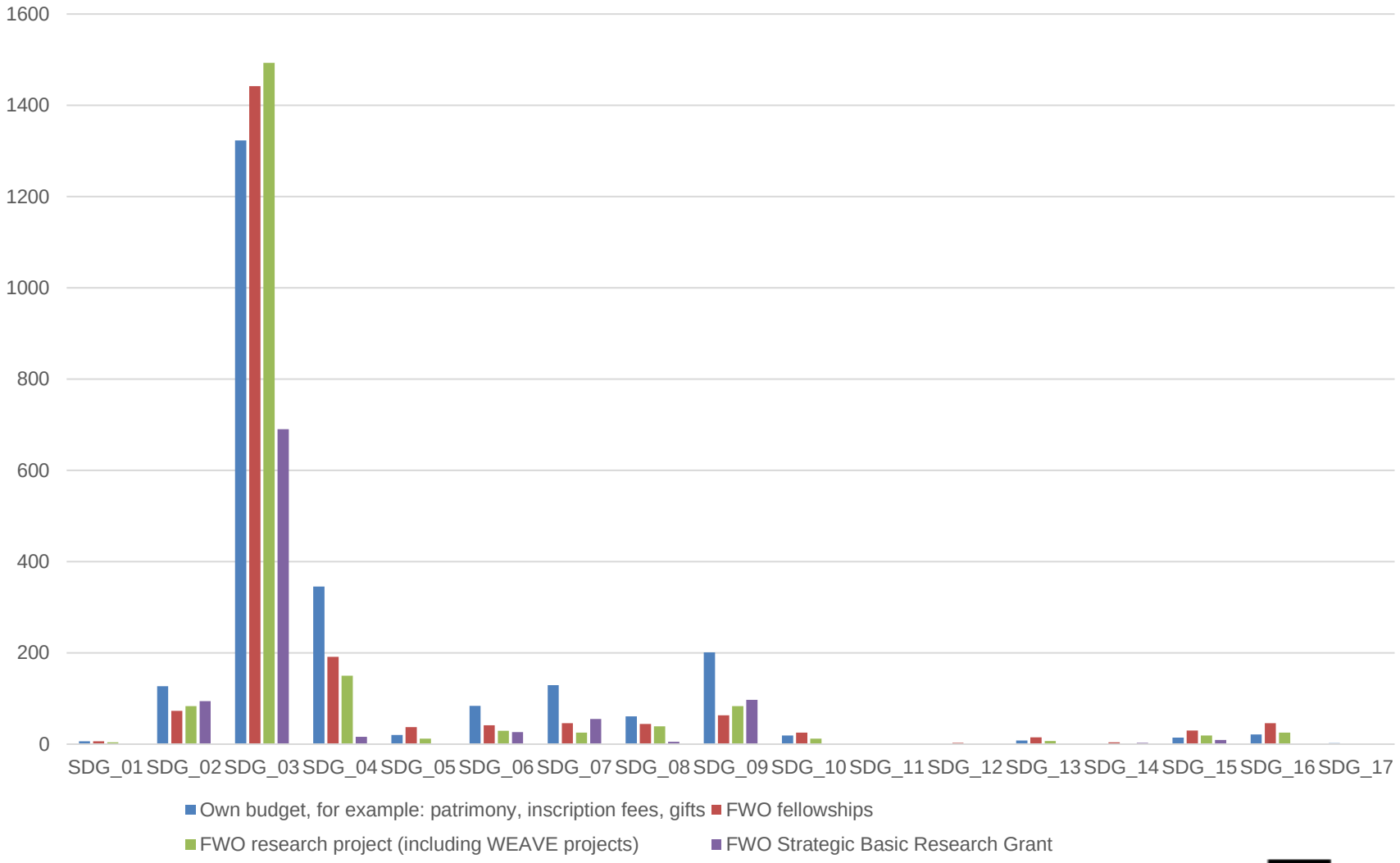
- Most projects have been assigned to SDG_03: Health.




ecoom

UHASSELT
KNOWLEDGE IN ACTION

# SDGs distribution in selected programs

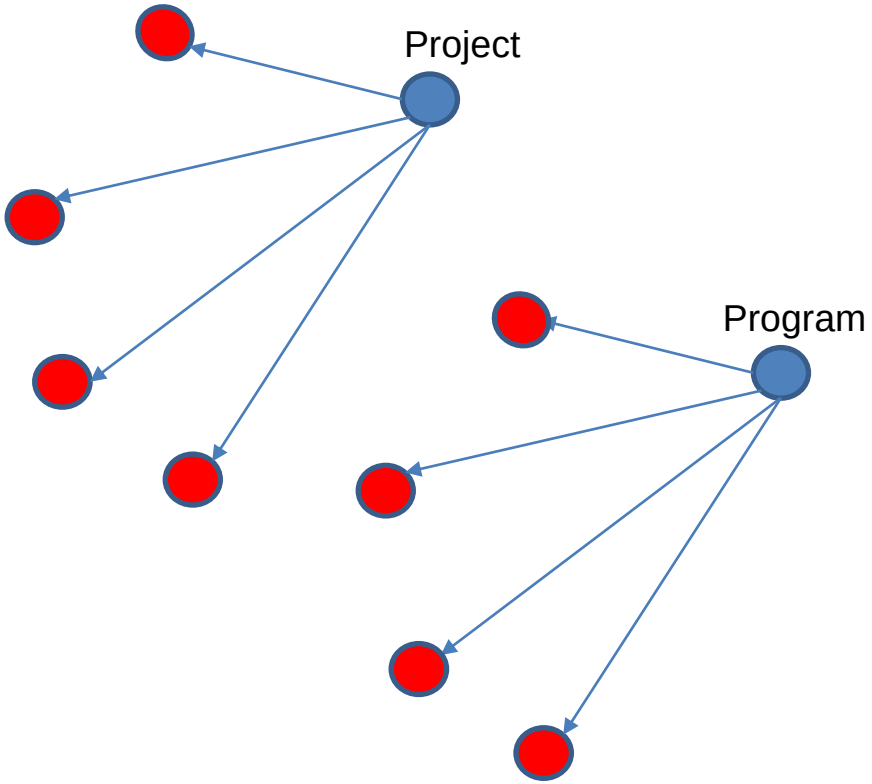


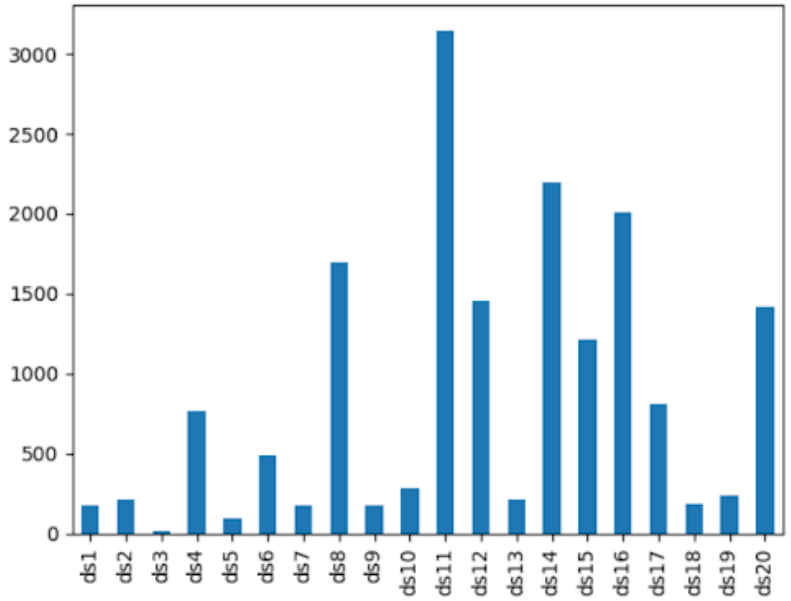

ecoom

UHASSELT
KNOWLEDGE IN ACTION

# Using Euclidean Distances

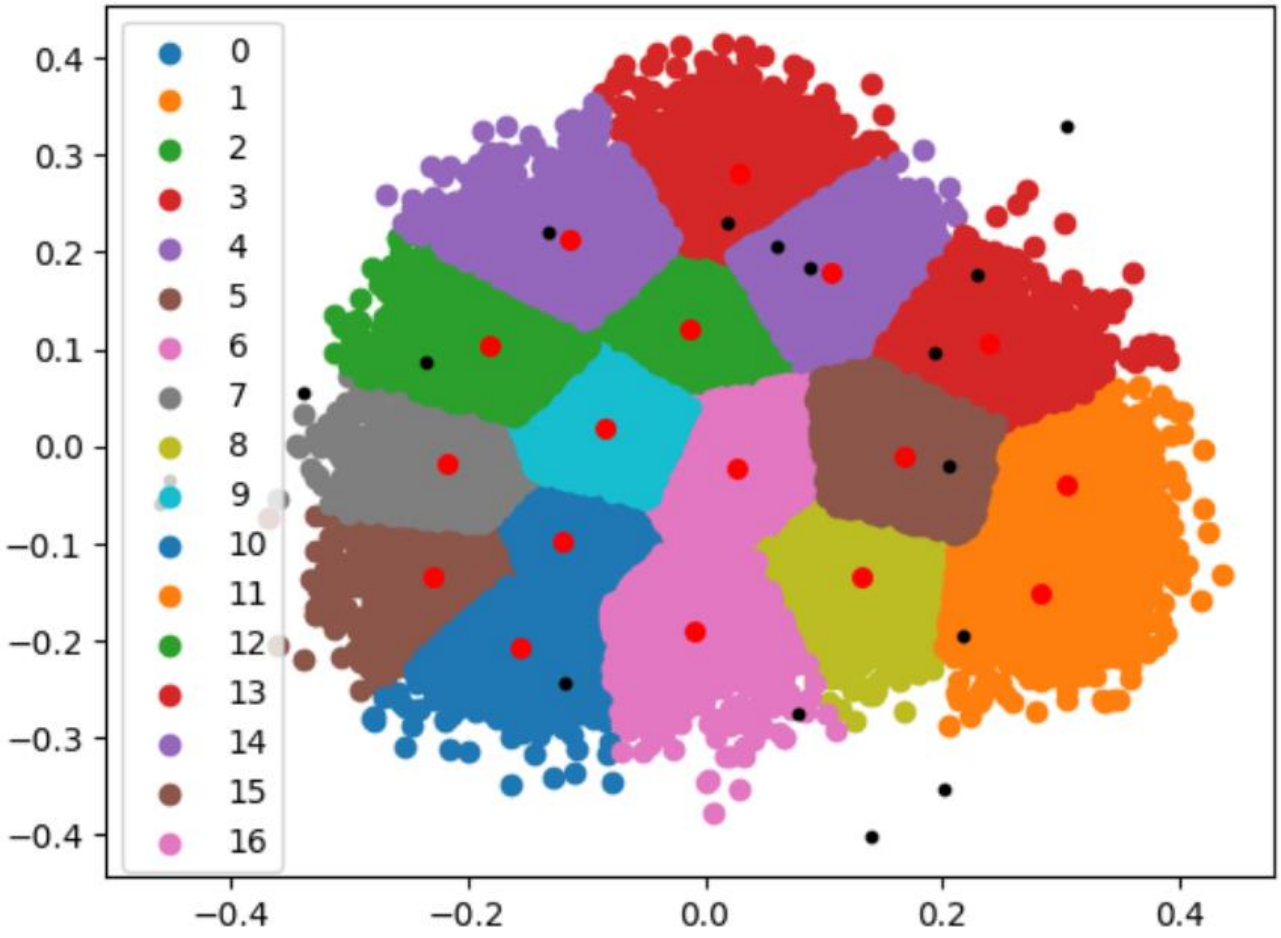
- We used a different method based on Euclidean distance calculation between each project and the 20 keywords and their definitions
- We selected the minimum distance to represent a score
- We achieved similar results to that by cosine similarity




Projects Distribution over Definitions

# Ongoing: A clustering approach to project alignment with the Sustainable Development Goals (SDGs)




Centroids

SDGs

ecoom

UHASSELT
KNOWLEDGE IN ACTION

# Conclusions

- We presented two approaches based on Large Language Models for identifying sustainability-related research projects and funding programs applied to the Flemish region.
- Most funding programs contain projects that are aligned with sustainability to some degree.
- Funding programs with a large number of projects can have different support levels for sustainability which can impact their overall alignment to SDGs.
- Projects associated with disciplines like Environmental Sciences, Civil and Building Engineering, and economics and business could be more aligned with sustainability.
- Most projects have been assigned to SDG_03: Health.

# Future Work

- Enrich FRIS project dataset with metadata from external sources to help improve the accuracy of the results.
- to investigate more advanced techniques to further improve the accuracy of the model.
- to build a classifier to predict new projects' alignment to SDGs based on the improved dataset.



# Thank you for your attention



Dr. Amr Ali-Eldin
amr.alieldin@Uhasselt.be