

IC
InfoCamere

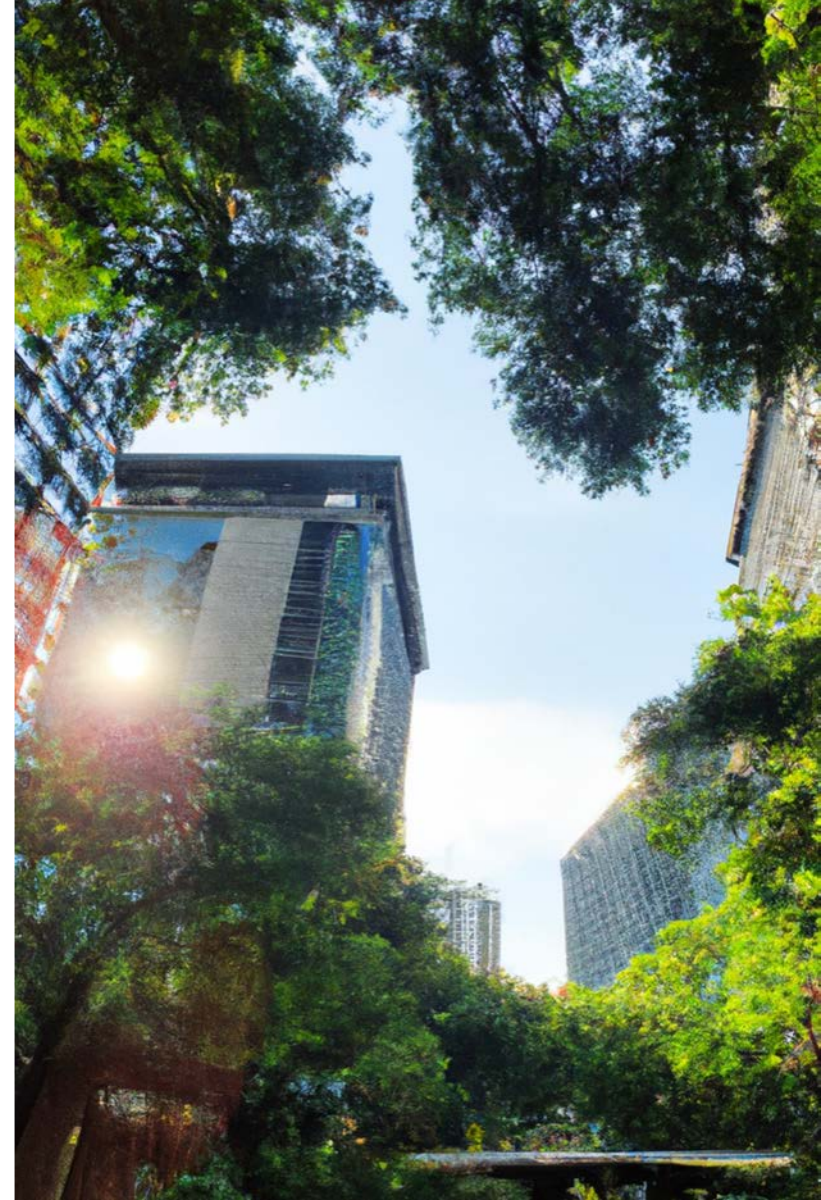
EBRA Conference, Paris, 31st May 2023

Using machine learning to map Sustainable Development Goals (SDGs)

A textual exploration of SDGs-alignment of
innovative firms and not-for-profit entities

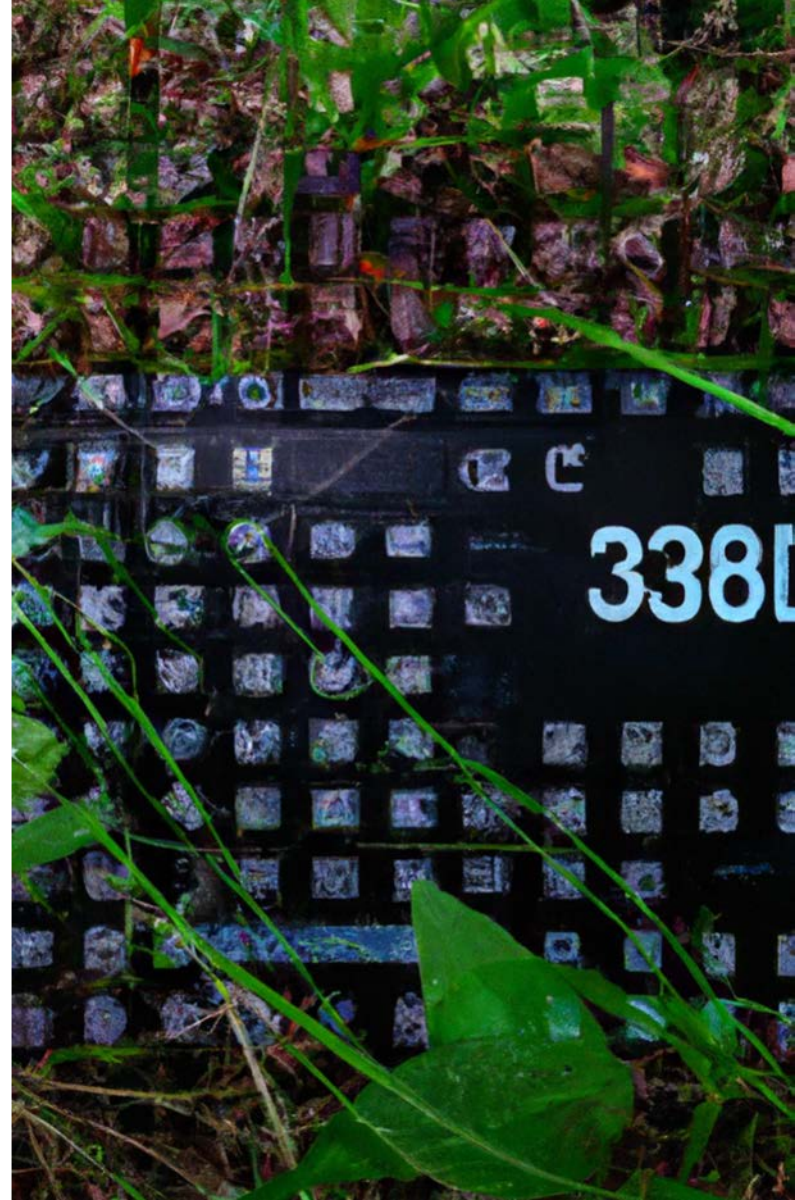
Sustainability: why it matters

1. Strong interest on sustainability by policy makers and investors;
1. Several policy efforts for its definition: ESG movement, 2030 agenda etc.;
1. Need for harmonization in definitions and measurements.



Sustainability: a quick outlook

1. EU's Corporate Sustainability Directive is setting the rules for reporting (taxonomy);
1. Information likely to be organized in semi-structured format, with reports starting in 2025;
1. Focus on large organizations;
1. Descriptive texts to be expected (as in the notes to the financial statements)?



Language-based mapping for SDGs

Joint effort with *Values 20.* **√20**

1. Can **textual analysis** help us explore **Sustainable Development Goals**?

1. Can **small and innovative** companies play a role in the sustainability challenge?



NICE STORY, BUT HOW DO WE MEASURE IT?



Language-based mapping for SDGs - explorative approach

Goal: to measure firms' business purpose alignment with SDGs.

Experimental method:

1. Provided a short descriptive text & keywords per SDG (jointly with V20);
1. Computed a *similarity algorithm* between companies' business purpose and SDGs descriptions;
1. Run descriptive analysis & classification.


Perimeter: innovative firms and not-for-profit entities, approx **53K** companies ~ 70% non-profit, 4th Q 2022.

A visual example

Business case

The association is a **non-profit organization** dedicated to conducting **socially beneficial activities** for its members and others, while upholding their **freedom and dignity**. Its main objective is to promote **widespread awareness** and understanding of the **value** and **history** of women...

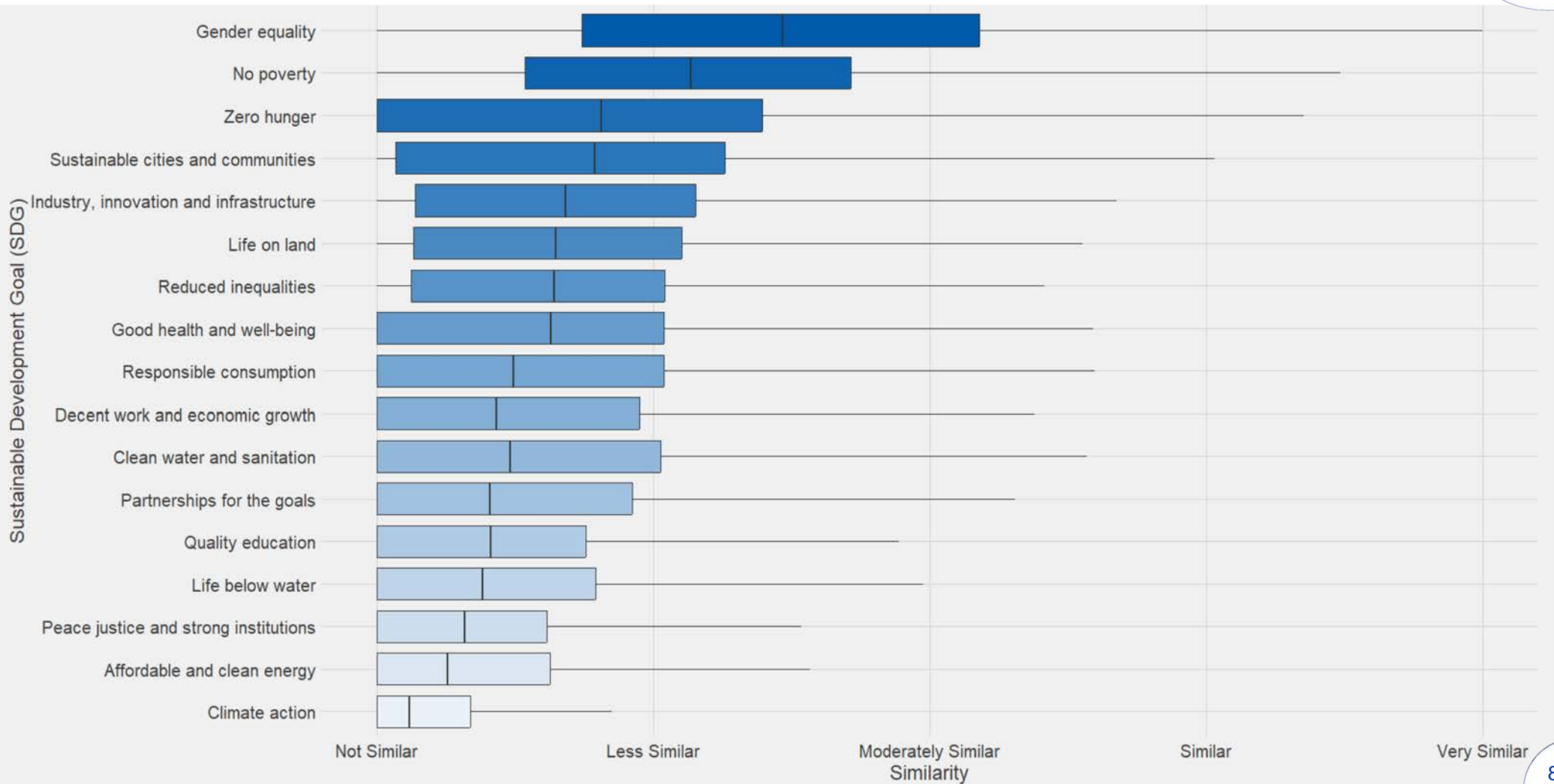
High similarity
Score ~ 0.54



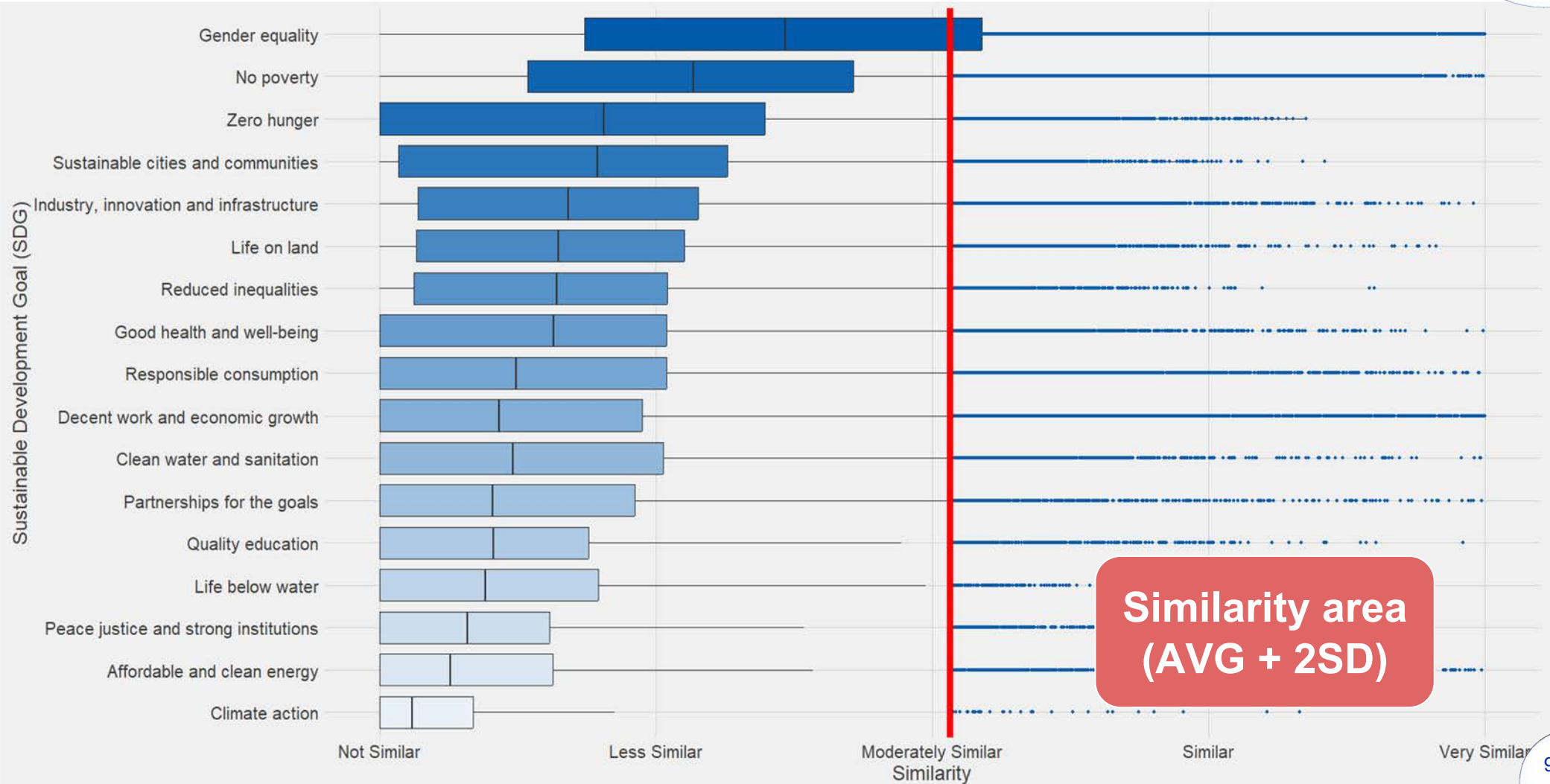
SDG #5: Gender equality

...to **promote gender equality** and the **empowerment of women and girls**, eliminating all forms of gender discrimination and violence. This involves promoting the **rights of women and girls**, their active participation in political, economic, and **social life**, **access to education and healthcare services**, and **promoting equal pay**.

HOW SIMILAR ARE FIRMS' BUSINESS PURPOSES TO SDGs?

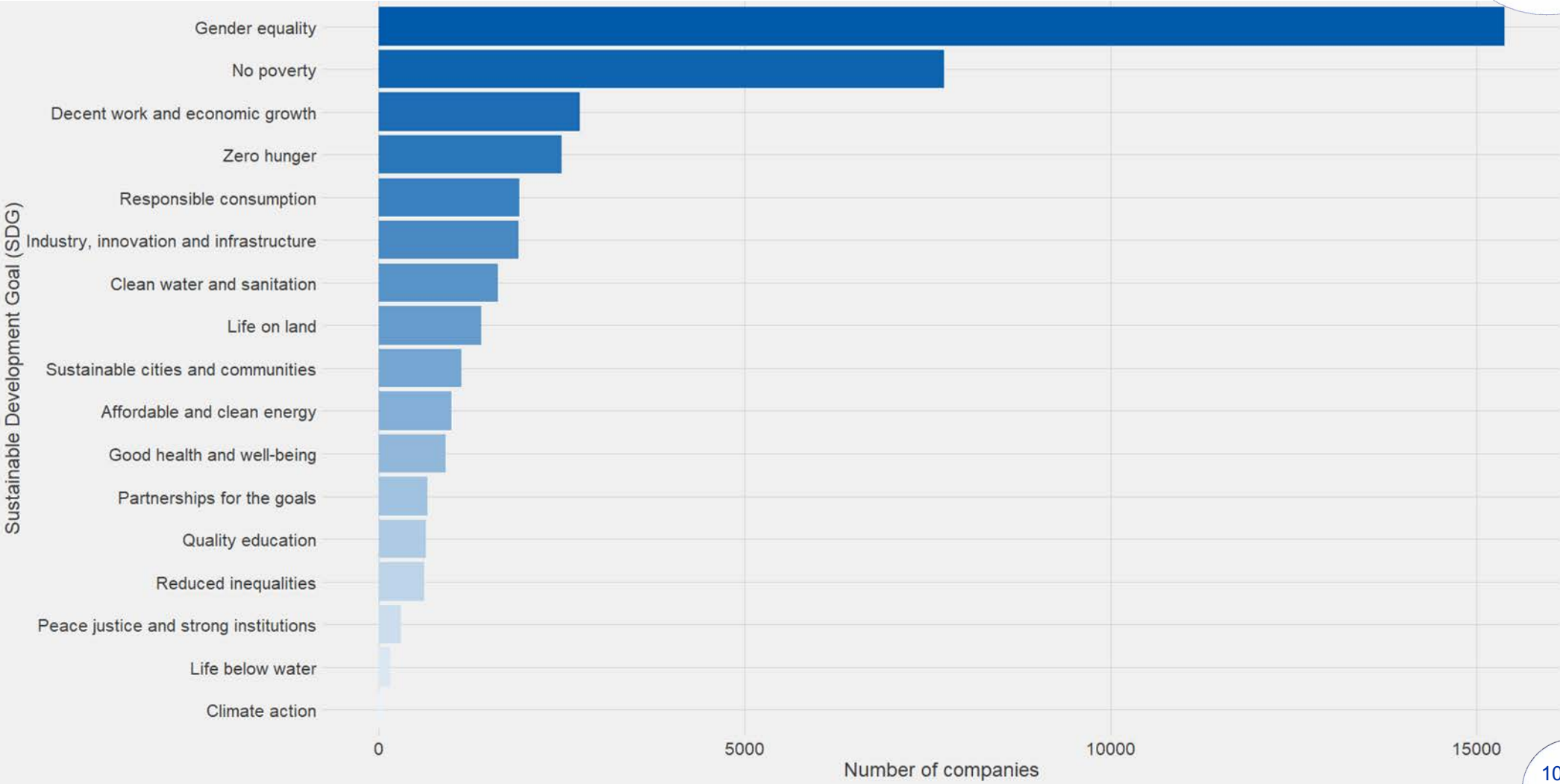


TEXT SIMILARITY FOR COMPANY CLASSIFICATION?

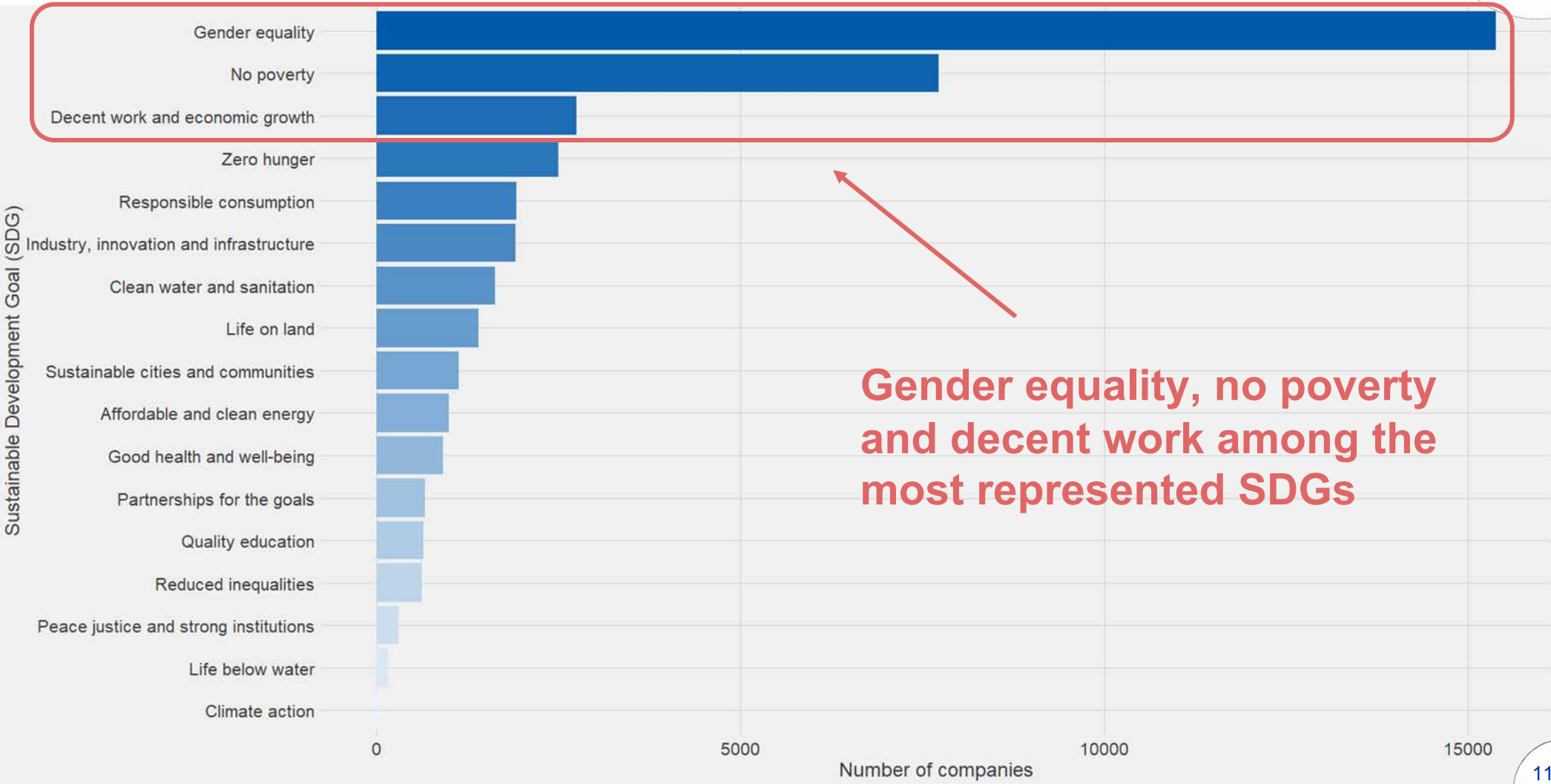


Similarity area
(AVG + 2SD)

HOW MANY COMPANIES DID WE FIND PER SDG?

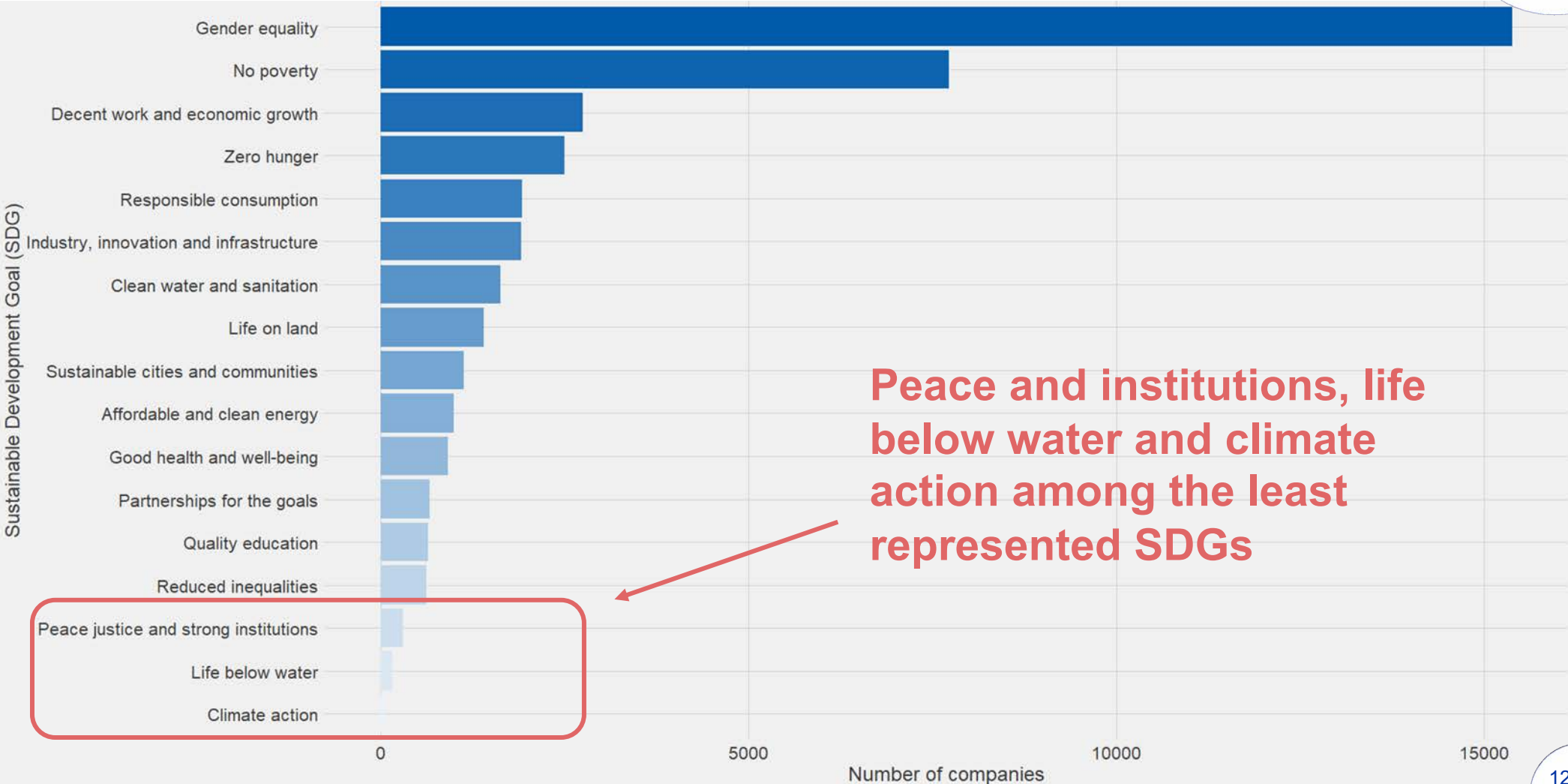


HOW MANY COMPANIES DID WE FIND PER SDG?



Gender equality, no poverty and decent work among the most represented SDGs

HOW MANY COMPANIES DID WE FIND PER SDG?



Peace and institutions, life below water and climate action among the least represented SDGs



OK, BUT ARE THESE RESULTS ROBUST?

We tested our results by sampling *company websites*:

1. We manually *classified* companies to individual **SDGs** based on the their ***website content*** (n = 150);
1. The similarity algorithm and the manual labeling were aligned in approx **70%** of the cases;
1. *An higher similarity score* was associated with *correct classification*.

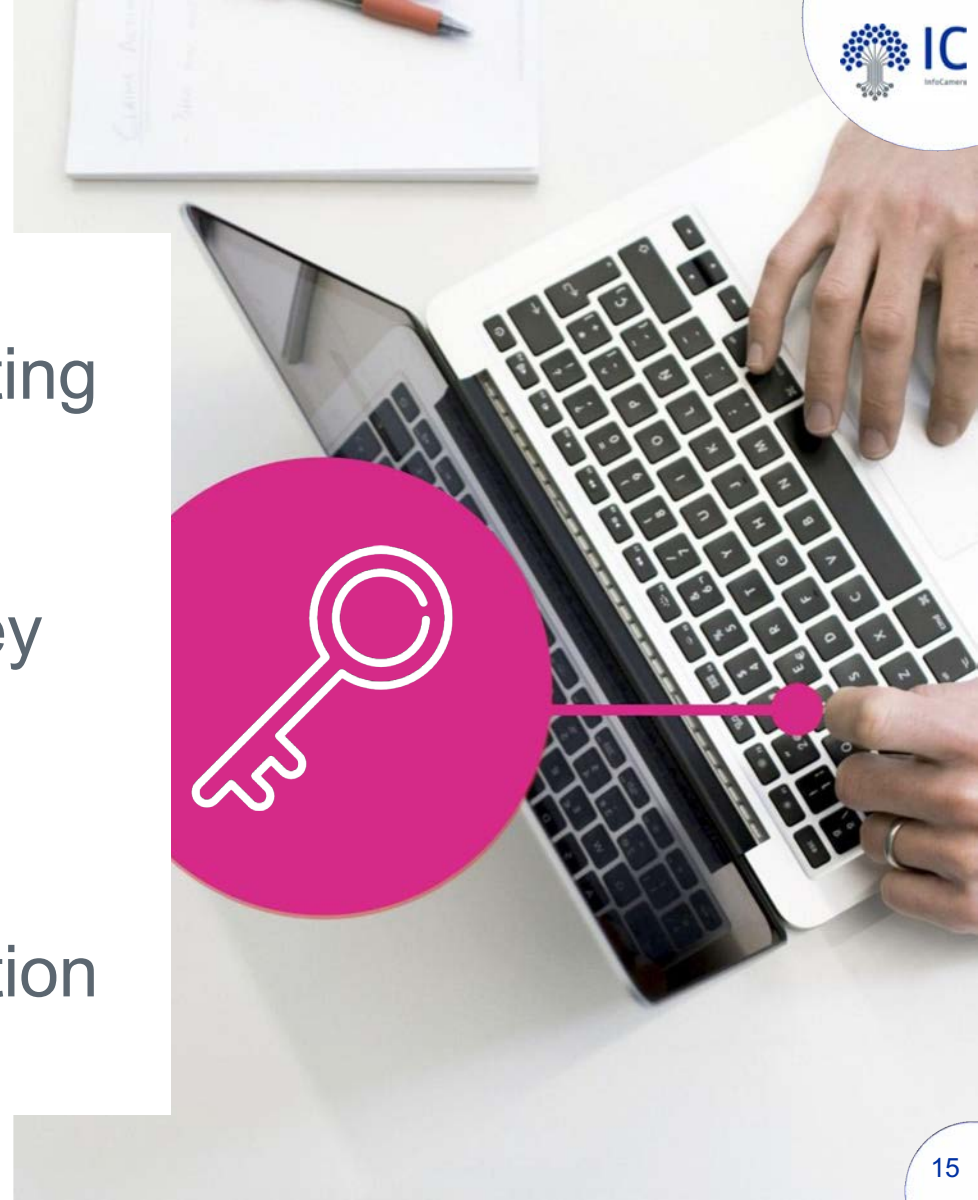
Limitations: jargon, purpose not actions, parameters.



Opportunities for Business Registries

Opportunities for Business Registries (BR)

1. Dedicated sustainability reporting is on the way (for big orgs);
1. Textual analysis might be a key approach to extract insights;
1. BRs might leverage this opportunity through collaboration and an ecosystem mindset.



Thank you

Contacts

niccolo.stamboglis@infocamere.it



IC

InfoCamere
Innovare è Crescere