



# **What does language say about human rights?**

## **Webinar**

### **Applying Data Science Techniques in the Evaluation of Human Rights-Based Approach (HRBA) in Finland's Development Policy and Cooperation**

**28th April 2023**

Development Evaluation Unit, Ministry for Foreign Affairs of Finland

In partnership with Particip GmbH-Niras Finland; and DavConsulting

# Purpose of the assignment



Part of an ongoing evaluation on human rights and human rights-based approach in Finland's development policy and cooperation.

- Serves as one stream of documentary evidence to the overall evaluation assignment.
- Other evaluation methods will also be used, particularly qualitative methods, to answer each of the evaluation questions.
- The results of the evaluation published in Oct 2023

Systematically examine level of HRBA application as evidenced by selected documents across different cooperation instruments.

Analysis based on [HRBA Guidance Note](#).

Piloting the use of methods.

1. Feeding into inception phase, identifying interesting cases in order to inform sampling
2. Feeding into evidence on EQ1 as one line of evidence on how human rights and HRBA feature in selected plans and reports for interventions

## Evaluation questions:

**EQ1: How and to what extent has the Human Rights-Based Approach been applied in the planning, implementation, monitoring and reporting of development cooperation funded by the Ministry?**

**EQ2: What have been the specific effects and value in actual terms of using the Human Rights-Based Approach for the effectiveness of various interventions, more transformative changes and ultimately for the realization of human rights and development policy objectives?**

**EQ3: How is the HRBA interacting with risk management of development cooperation interventions?**

# How to read the results from this analysis?



These methods yield **estimates**, not exact calculations. They give a broad **overview** of the topic.

**No direct conclusions about HRBA implementation** in actual terms or answers to evaluation questions can be drawn from the data science results alone.

The results **do not explore the reasons** why a given result arises.

The results of the three **different methods are not comparable** with each other. They each provide a different perspective to the topic.

The **cooperation instruments are not comparable** with each other. Look at each instrument in itself.

- What do the two types of methods reveal about the cooperation instrument or intervention relevant to you?
- To what extent are the analyses on par with the Ministry's self-assessed HRBA markers that assign the level of ambition to the interventions?
- What do the results suggest as any potential needs for capacity development; points for discussion; or needs for further development of the instruments or interventions in relation to the information in their documentation etc.?
- What do the results suggest about the interpretation of HR, HRBA and the Guidance Note?



# Programme

- 14:00 Opening words  
Presentation of results  
Questions and comments from  
the audience
- 15:30 End of webinar



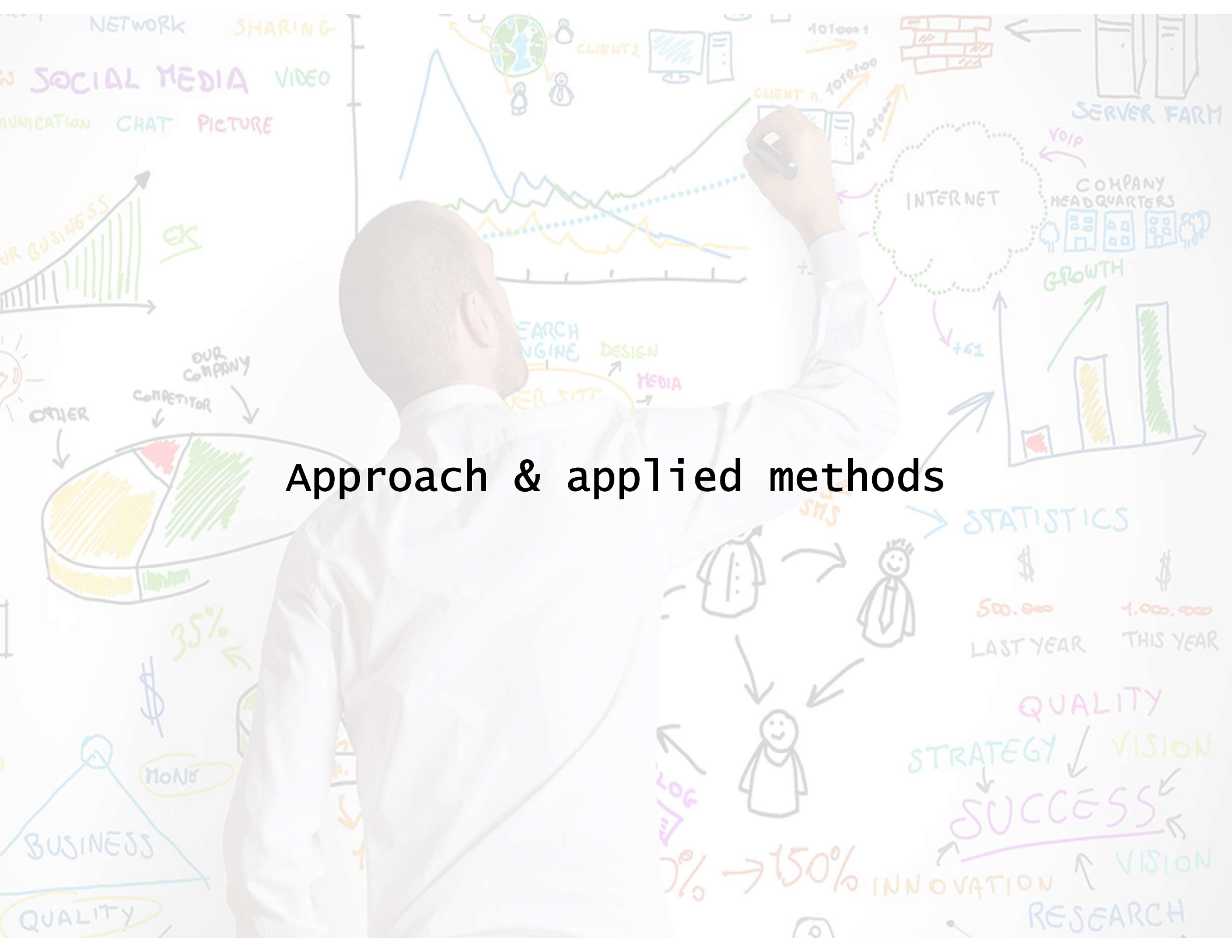
**dav|consulting**

# Process & Results from a Mixed-Methods NLP Approach



# Talking points

1. Approach and applied methods
2. Results and estimates from applied methods
3. Questions and answers



# Approach & applied methods

NETWORK SHARING  
SOCIAL MEDIA VIDEO  
COMMUNICATION CHAT PICTURE

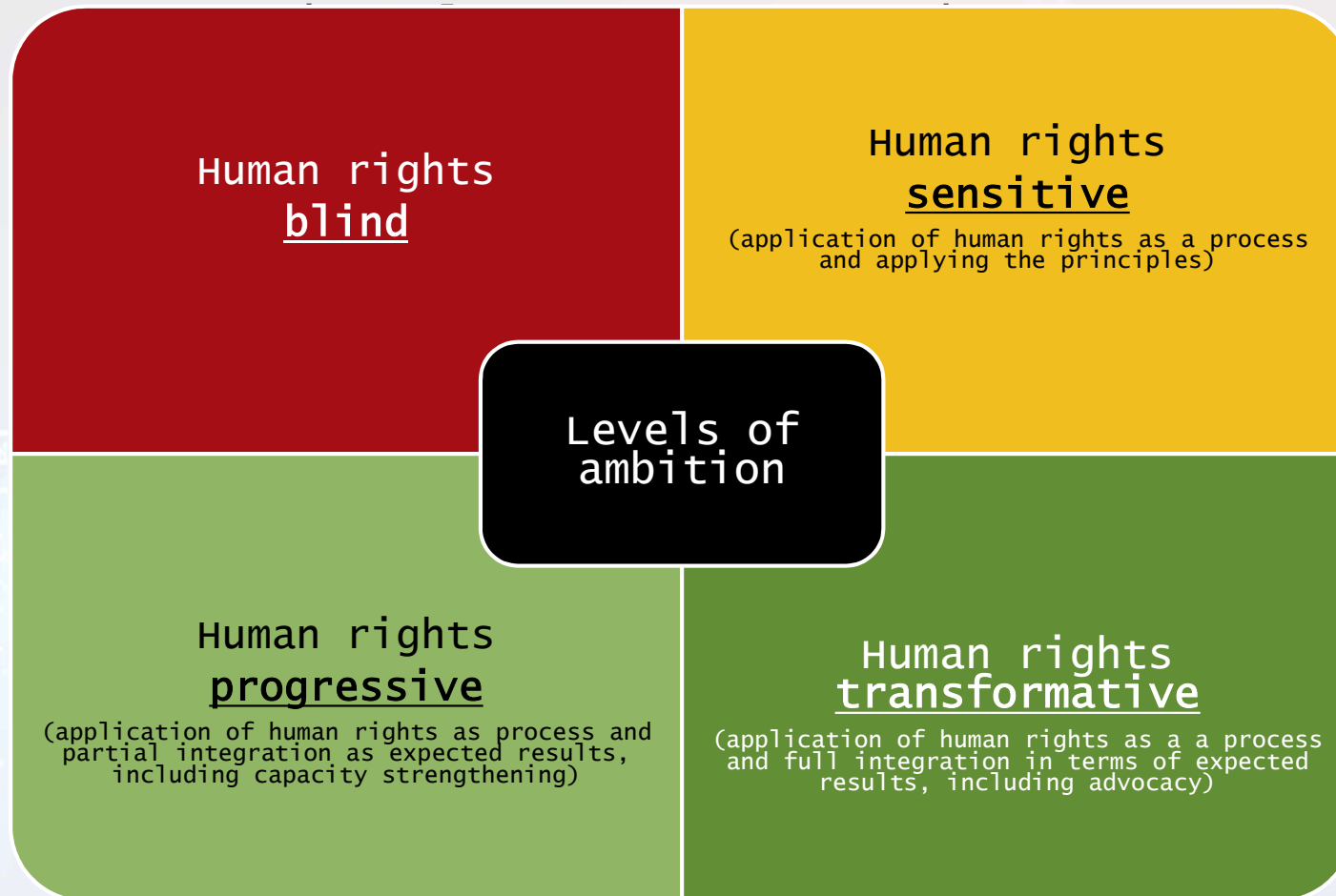


SEARCH ENGINE  
DESIGN  
MEDIA  
WEB SITE

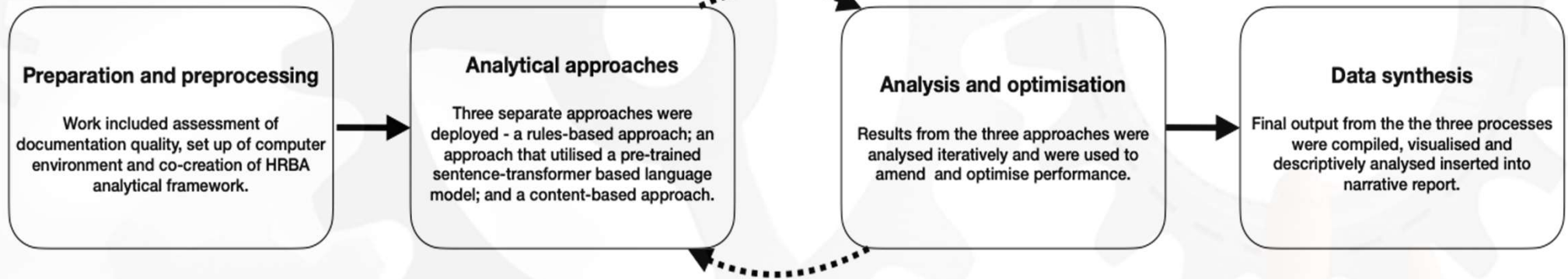




# Task in a nutshell



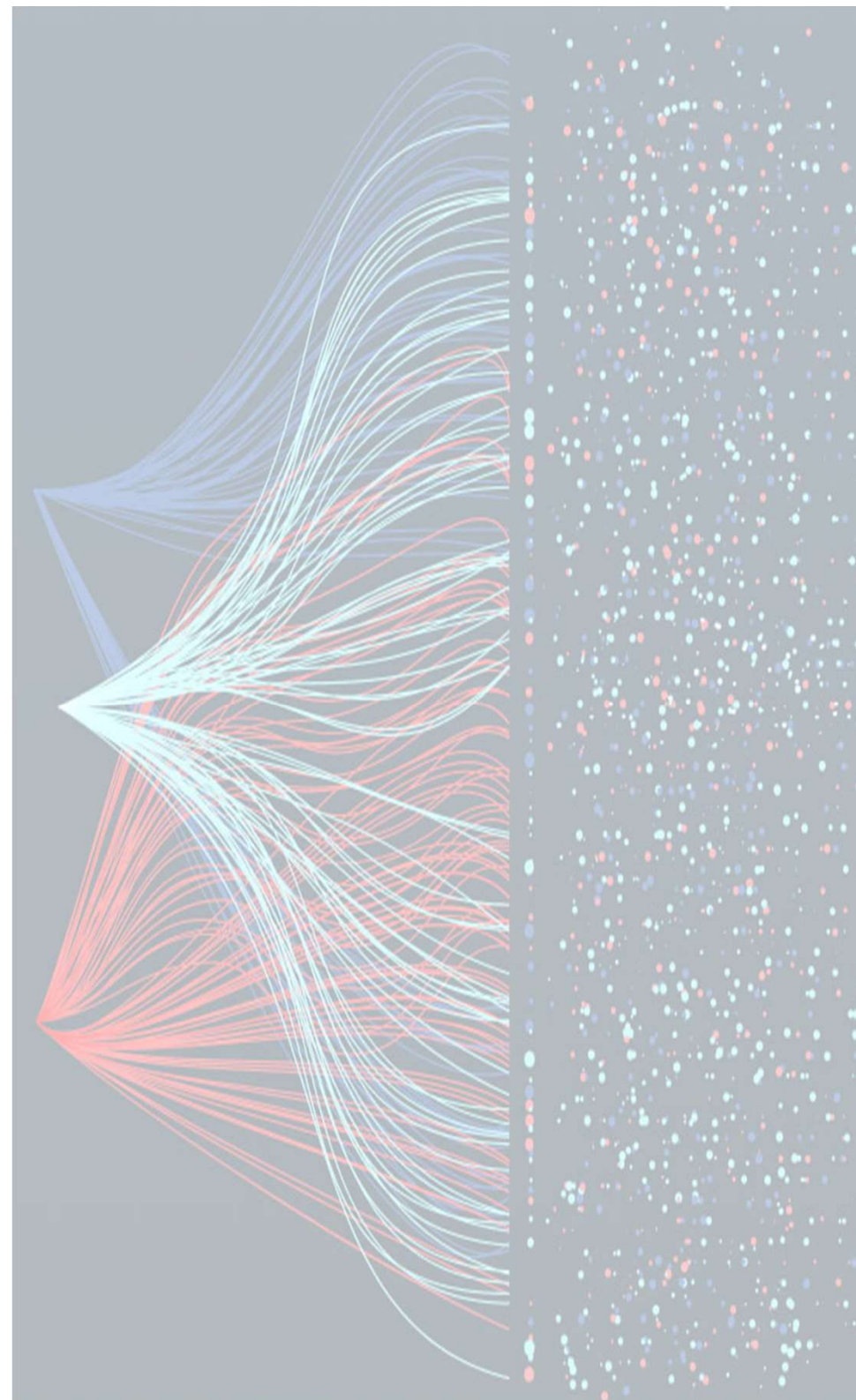
# Work process & NLP pipeline





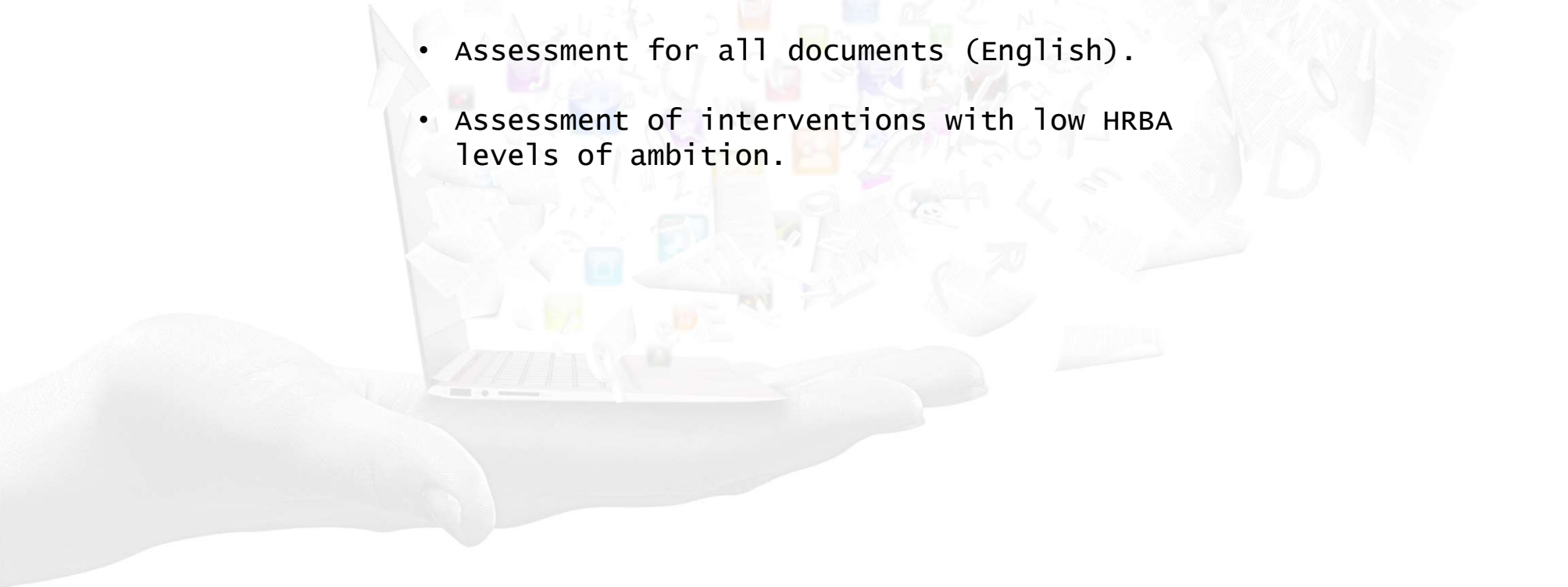
# Pre-trained language model

- Machine learning approach to classification.
- Transforming text in all documents into vector representations (i.e. text embeddings).
- Pre-processing and cleaning exercise (i.e. removal of content without bearing on HRBA).
- Identification/draft of text paragraphs typical for each of the HRBA levels of ambition.
- Similarity comparison of the typical sections and intervention documents



# Content analysis

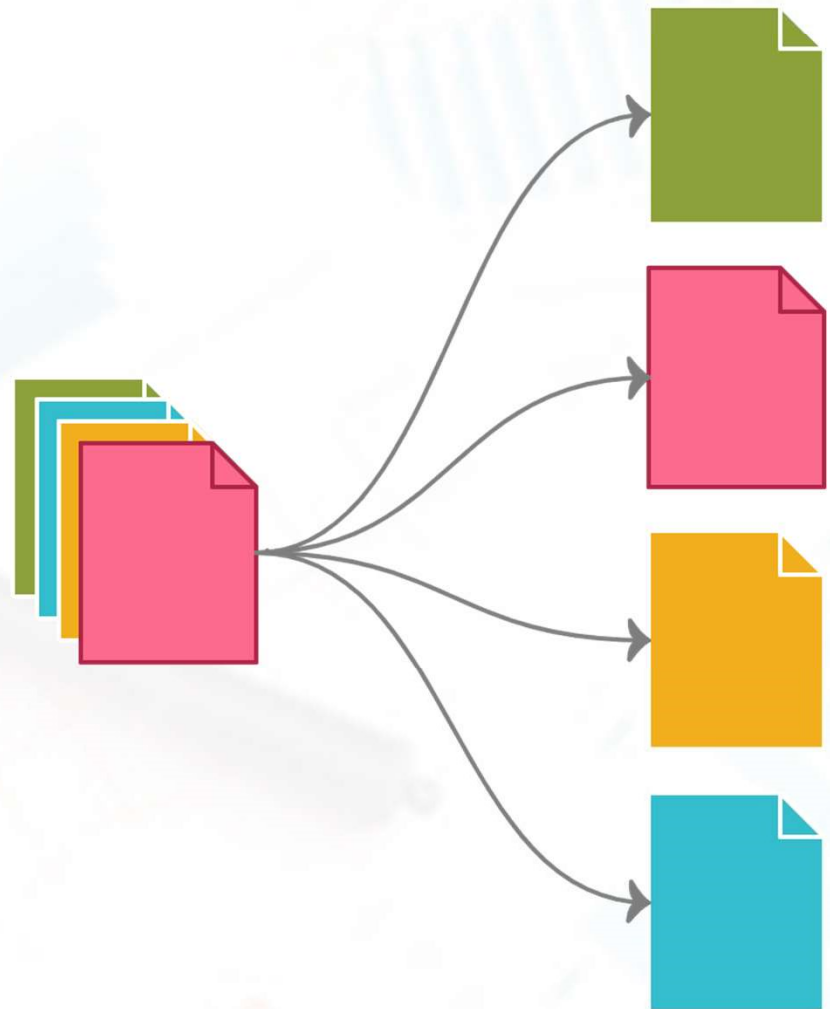
- Extracting content of relevance.
- Utilising tokenised documentation to extract nouns.
- Data cleaning and compilation of frequency of found nouns.
- Assessment for all documents (English).
- Assessment of interventions with low HRBA levels of ambition.





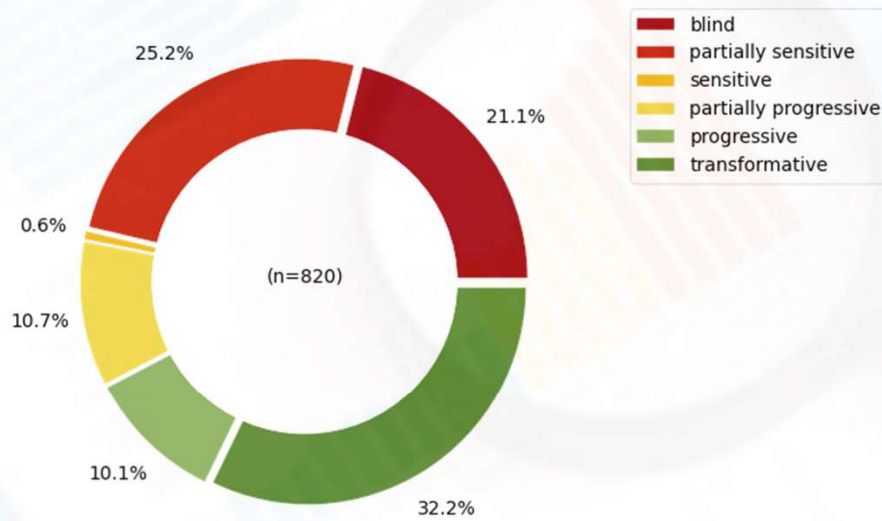
R E S U L T S

# Classification process

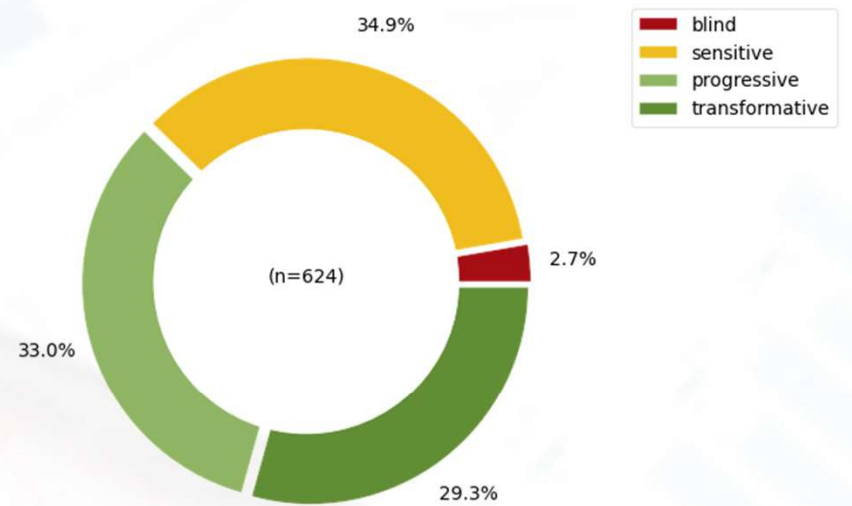


# HRBA Levels of ambition by approach

## Rules-based



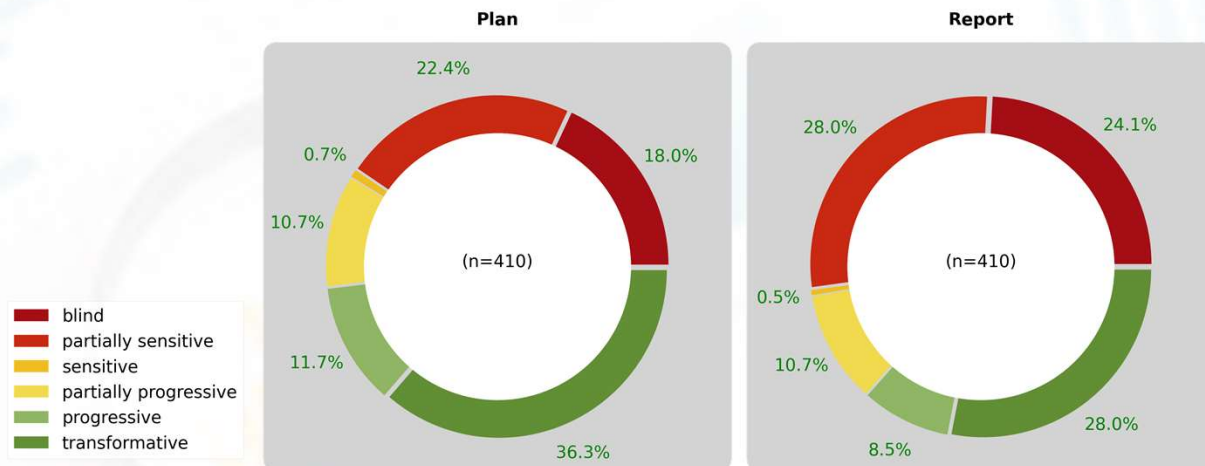
## Machine Learning



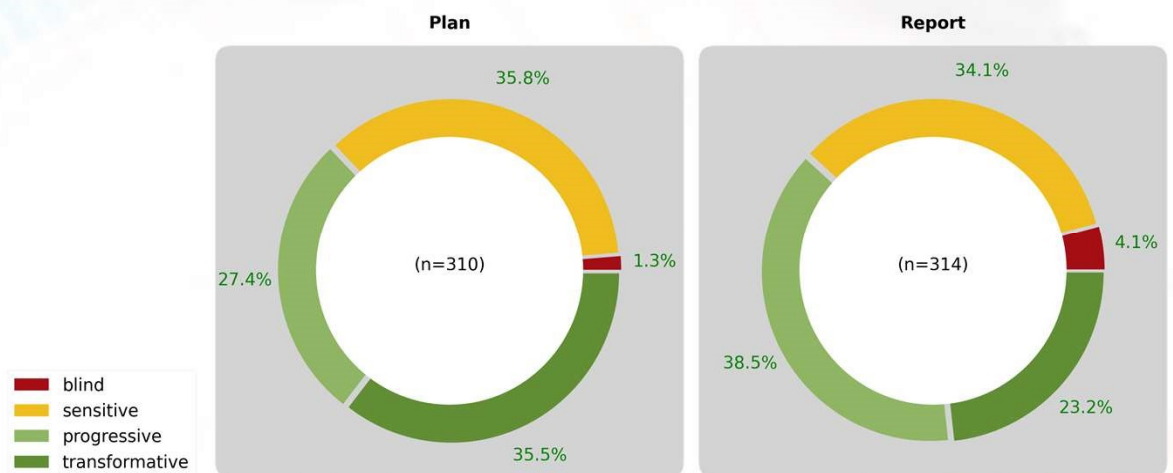


# HRBA levels of ambition by document type

Rules-based

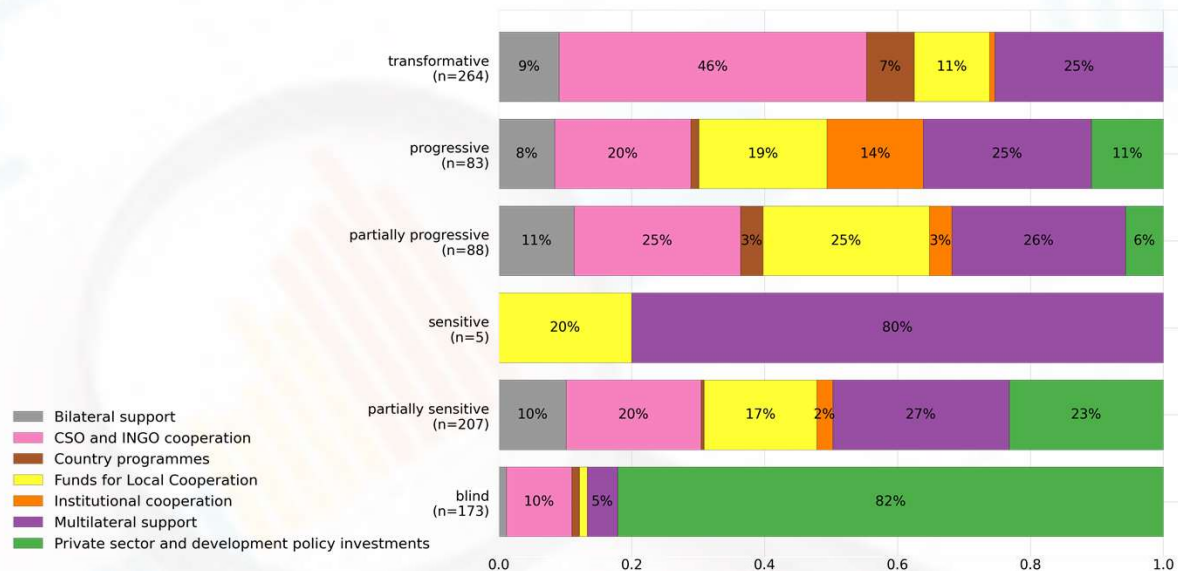


Machine Learning

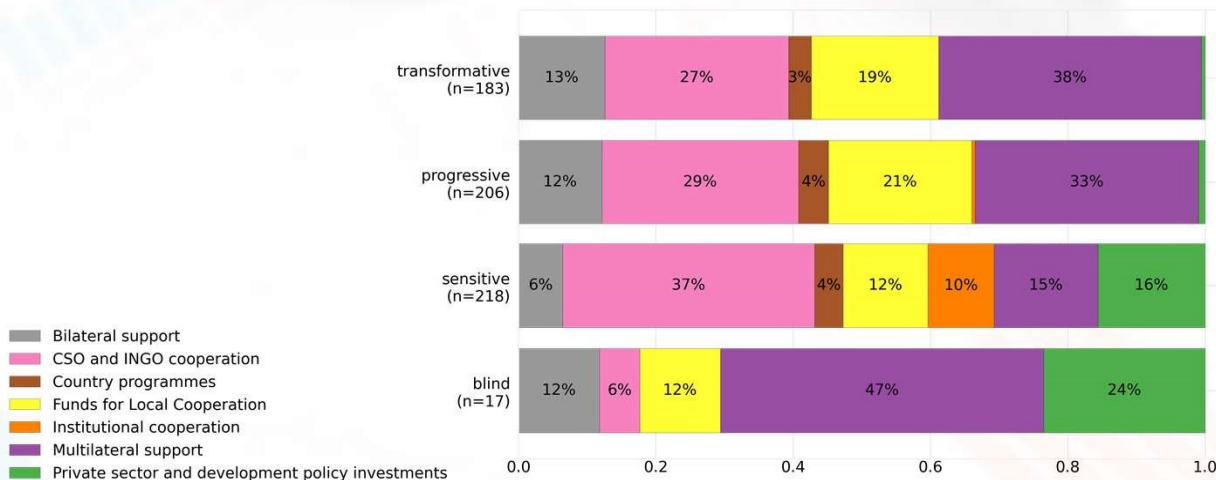


# HRBA levels of ambition by cooperation instrument

Rules-based

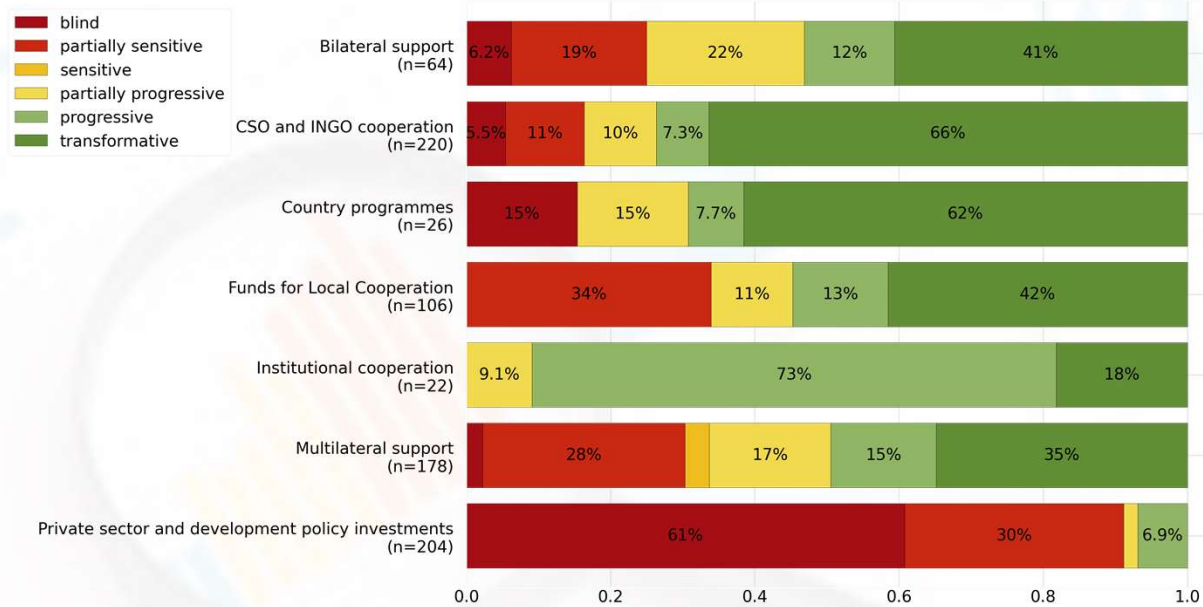


Machine Learning

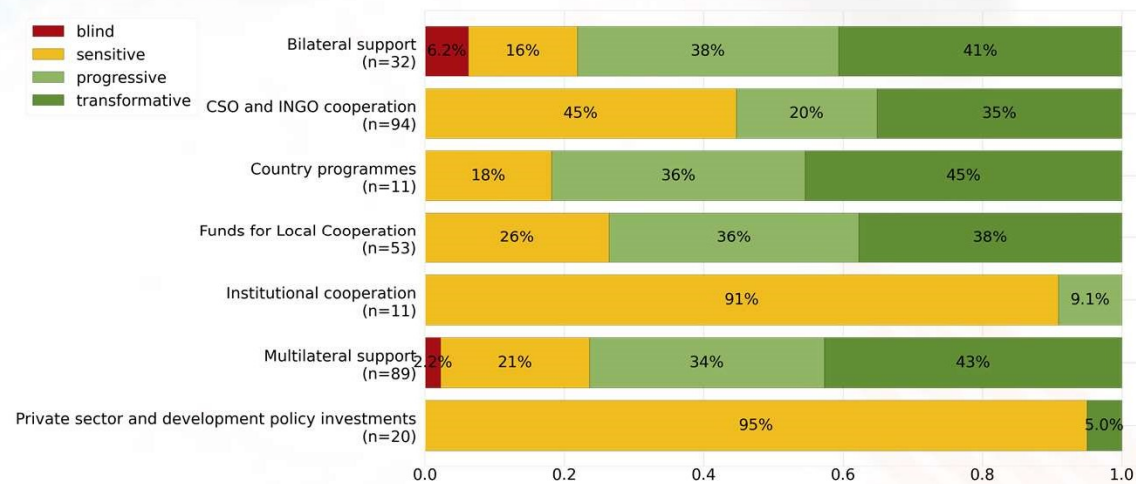


# HRBA levels of ambition in plans by cooperation instrument

RuLes-based

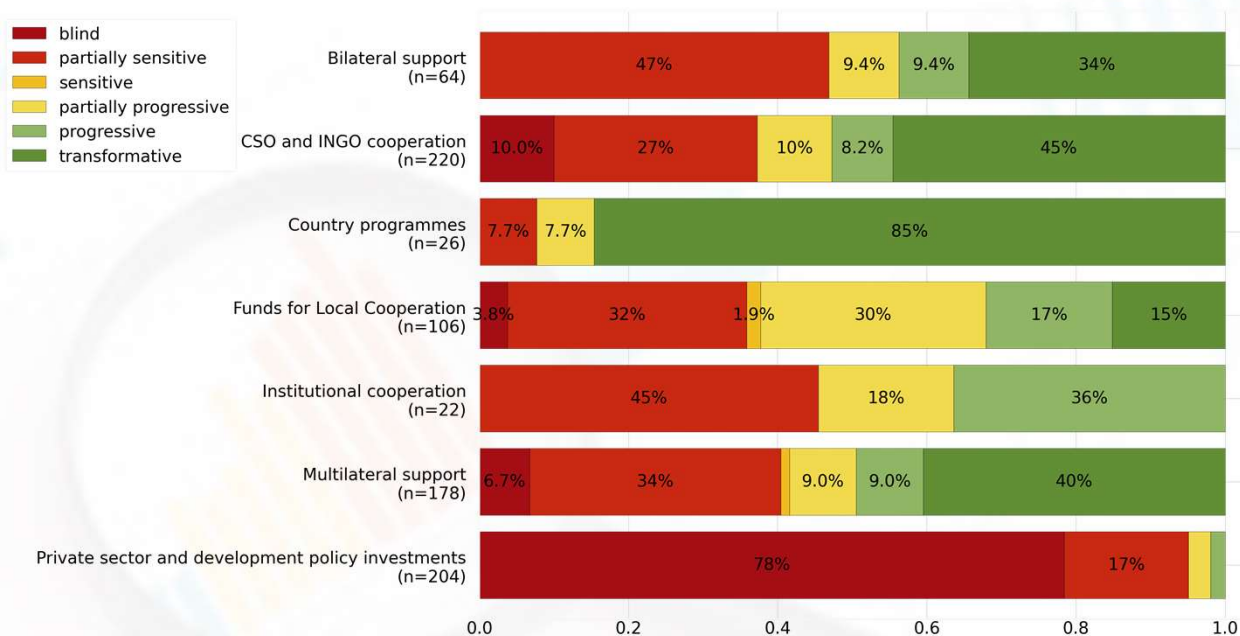


Machine Learning

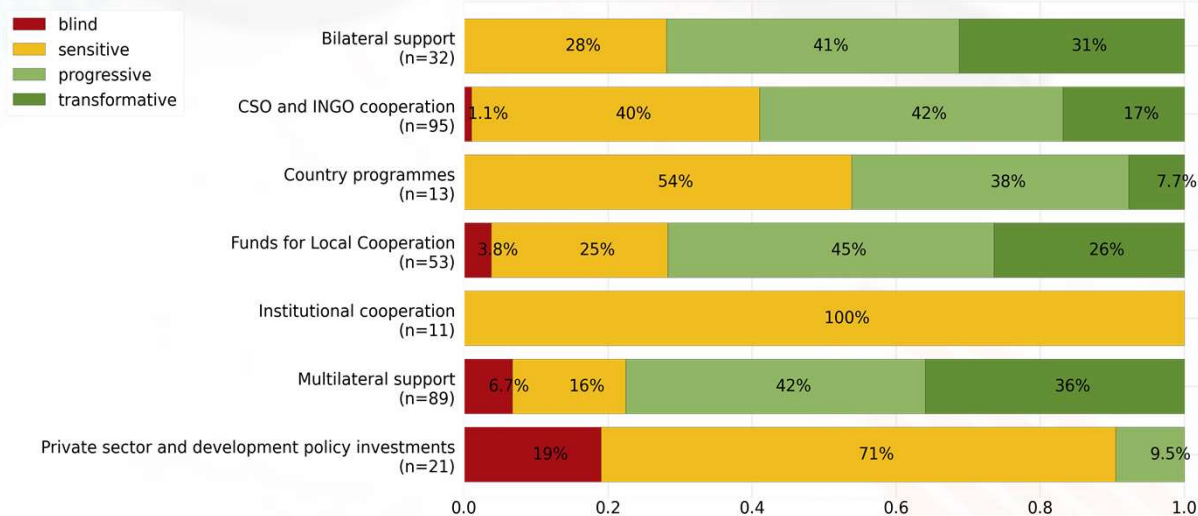


# HRBA levels of ambition in reports by cooperation instrument

RuTes-based

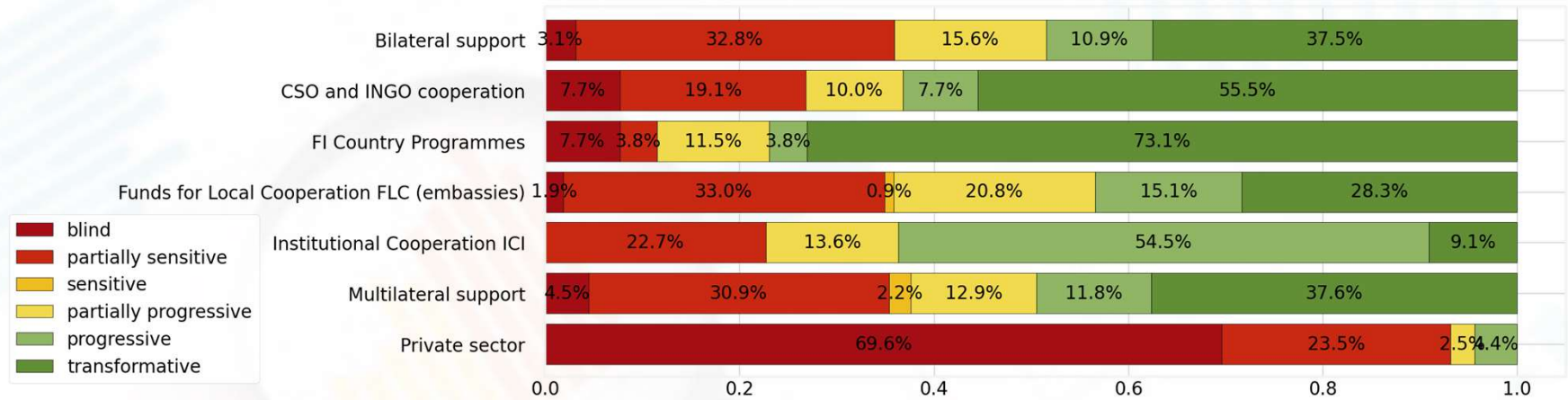


Machine Learning

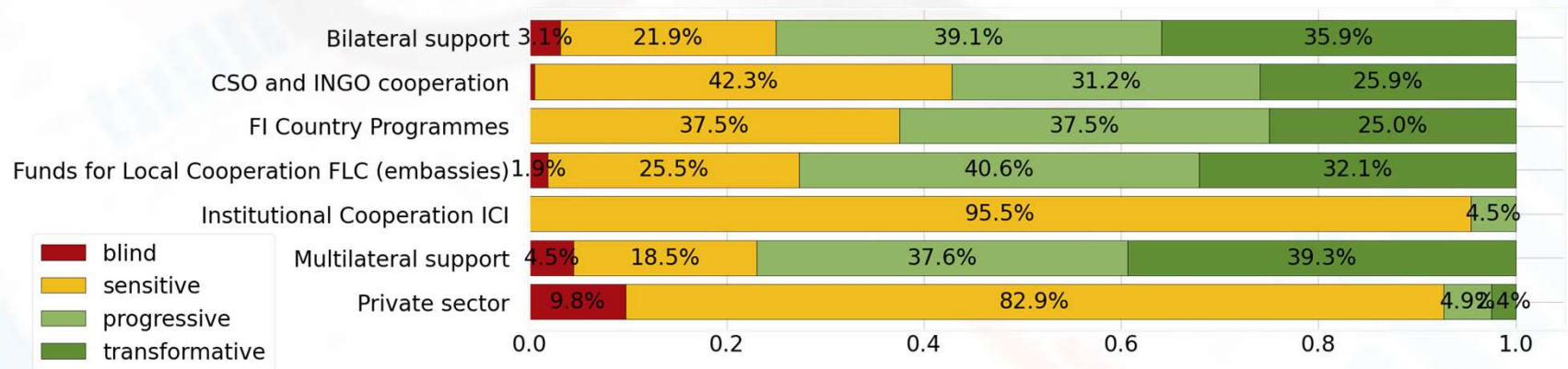


# HRBA levels of ambition by cooperation sub-instrument (compiled)

Rules-based

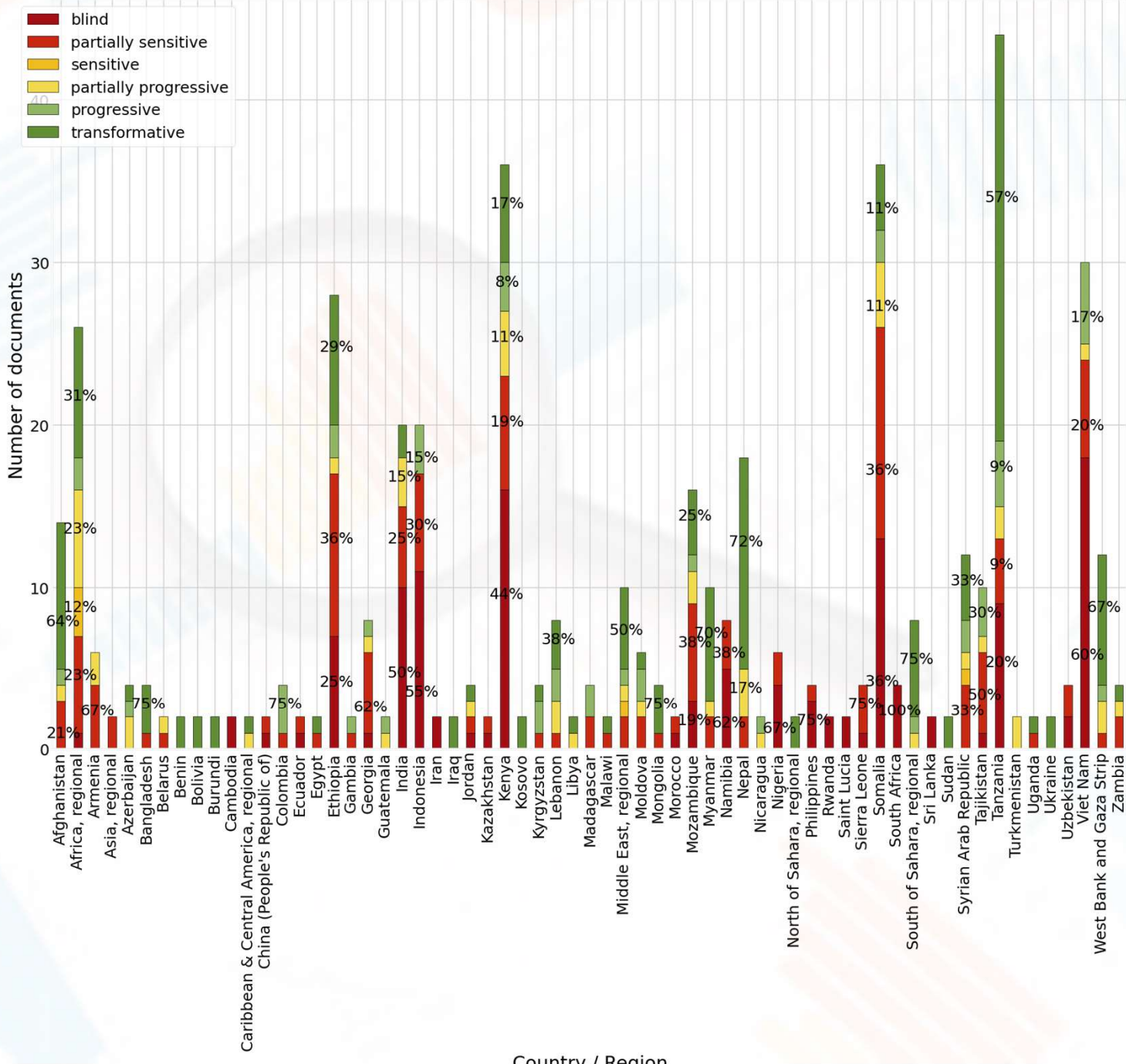


Machine Learning



# HRBA levels of ambition by country

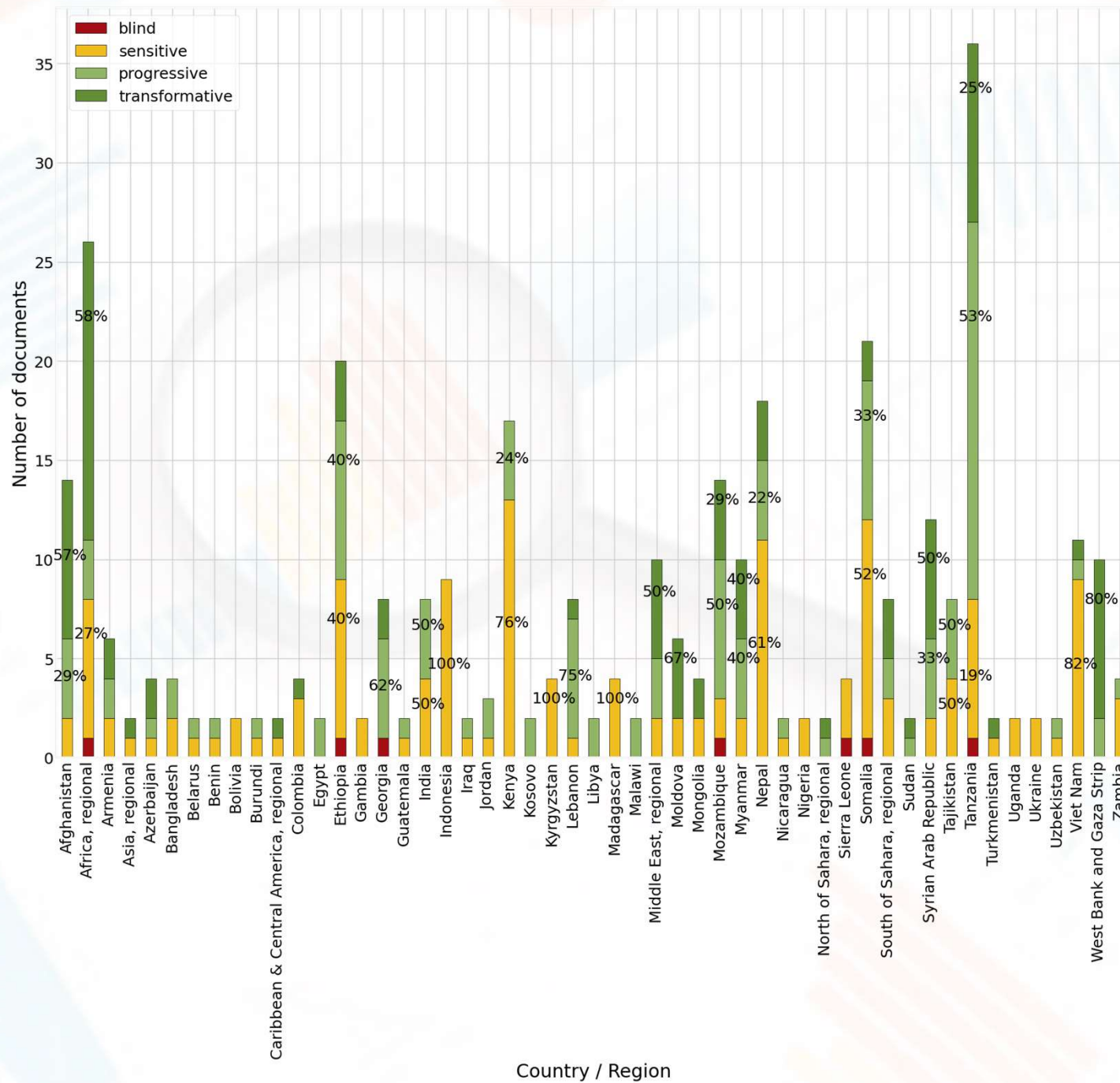
RuT es - based



Country / Region

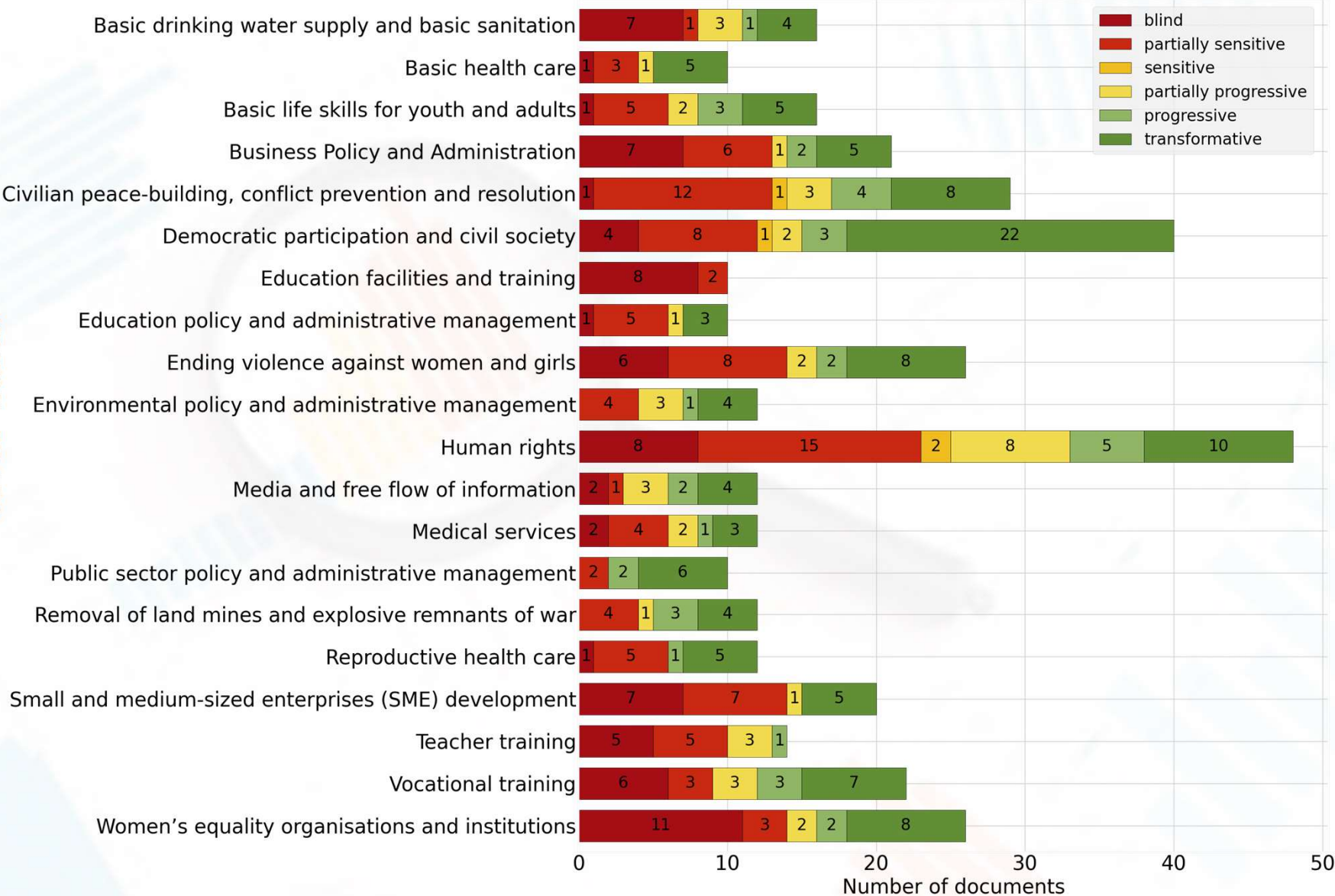
# HRBA Levels of ambition by country

Machine Learning



# HRBA Levels of ambition by sector

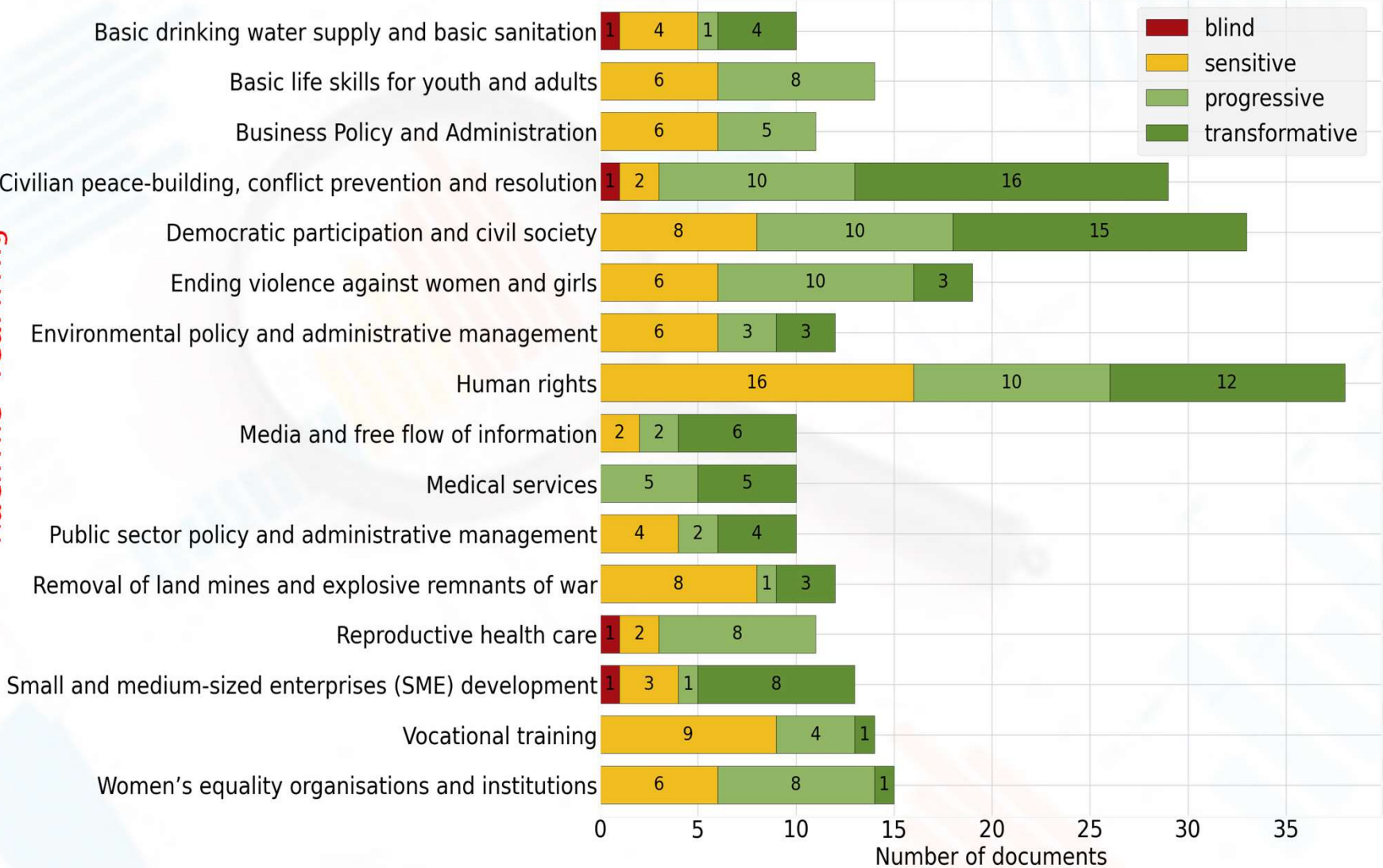
RuTes-based





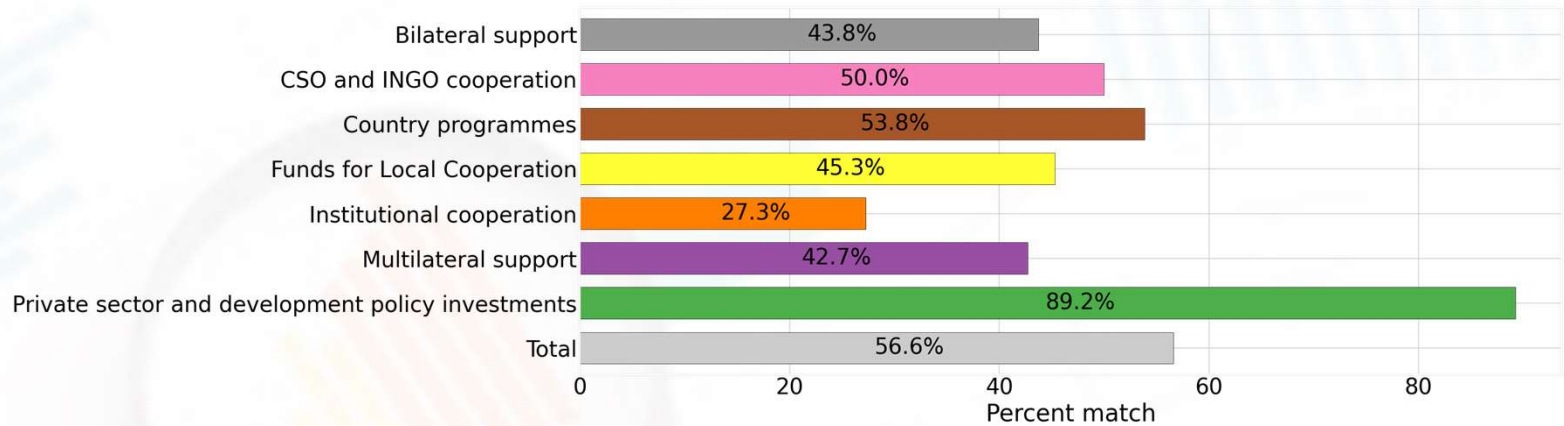
# HRBA levels of ambition by sector

Machine Learning

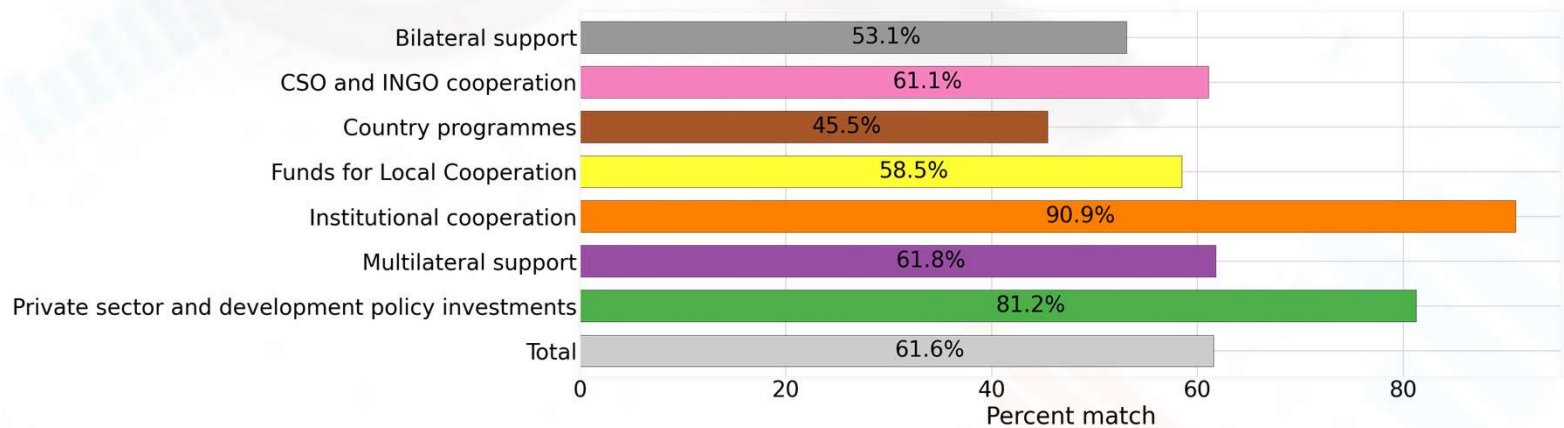


# Consistency in HRBA levels between plans and reports

Rules-based

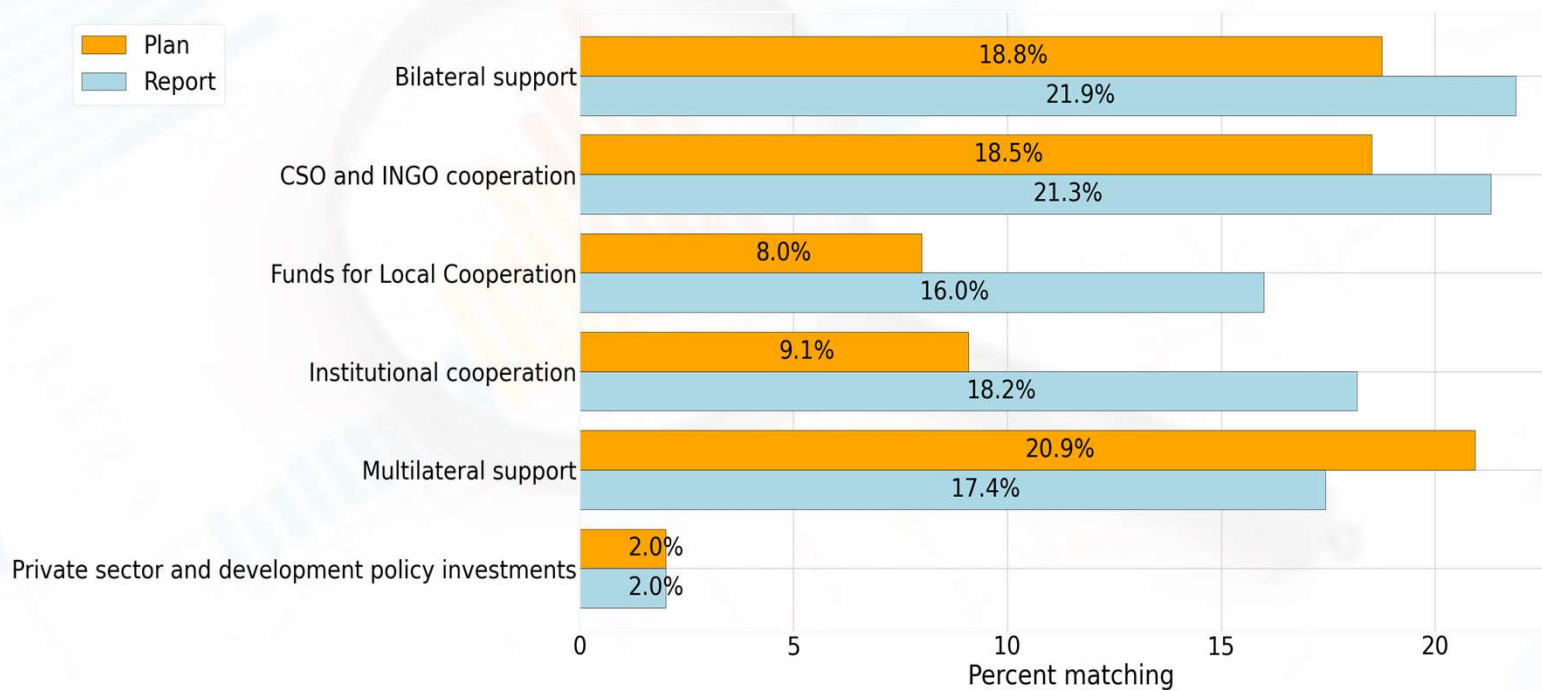


Machine Learning



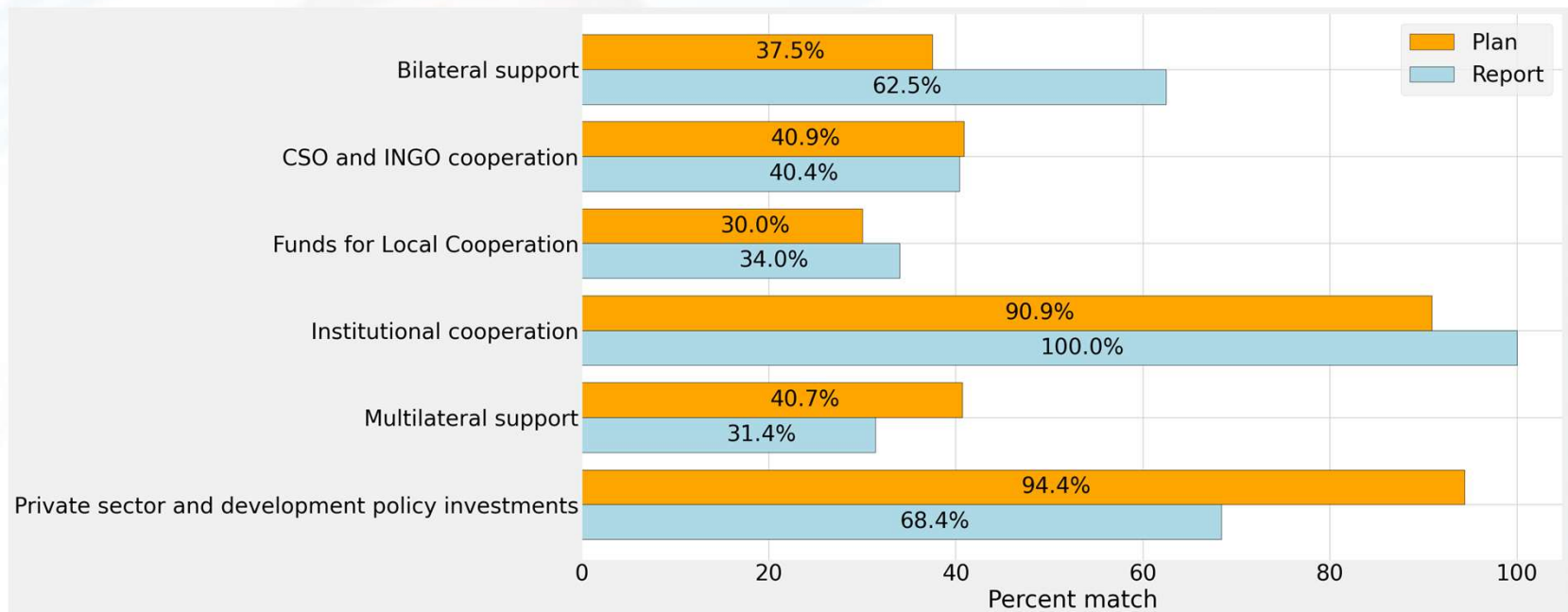
# Share of HRBA levels matching self-assessment by cooperation instrument and document type

## Rules-based



# Share of HRBA levels matching self-assessment by cooperation instrument and document type

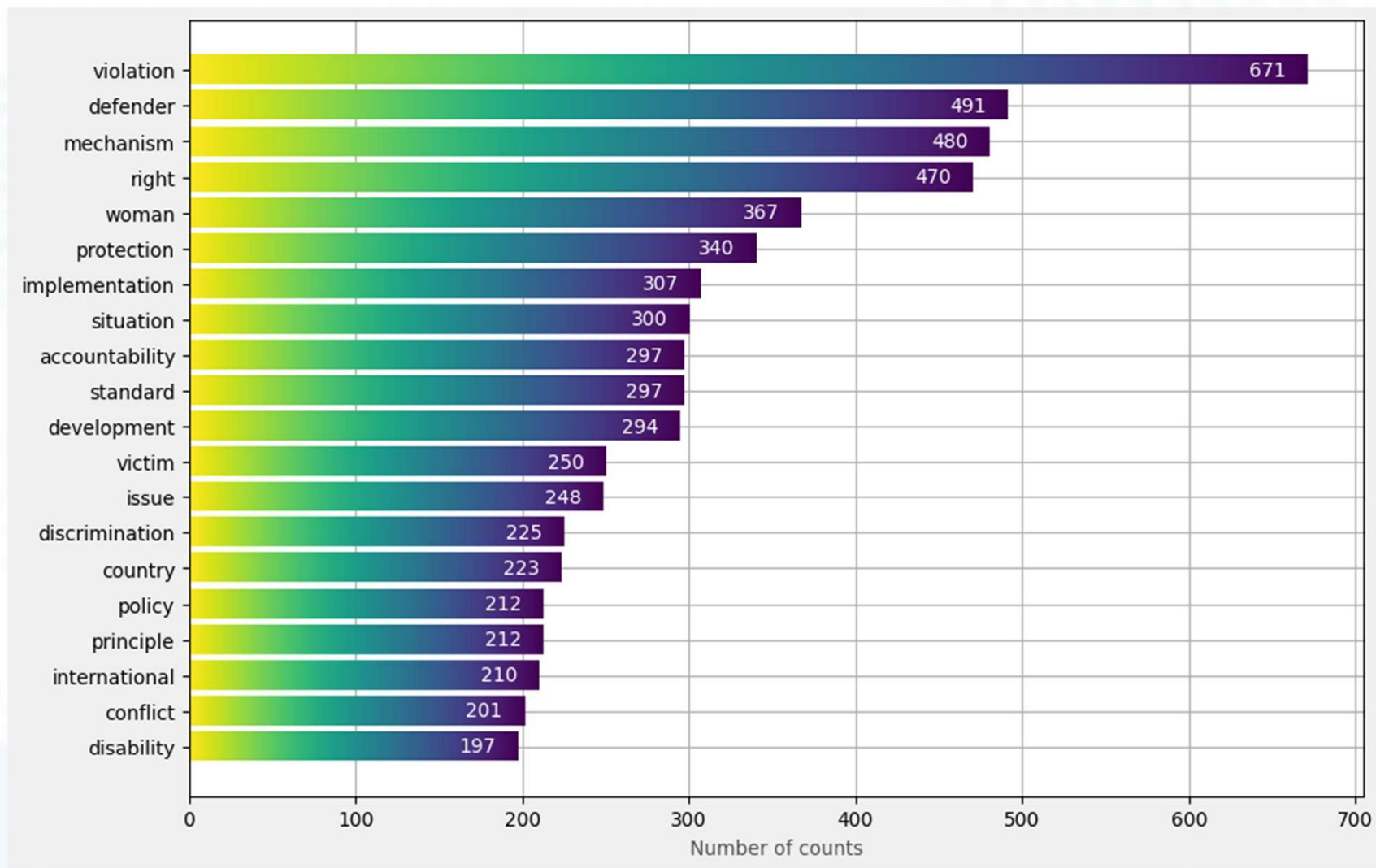
## Machine learning



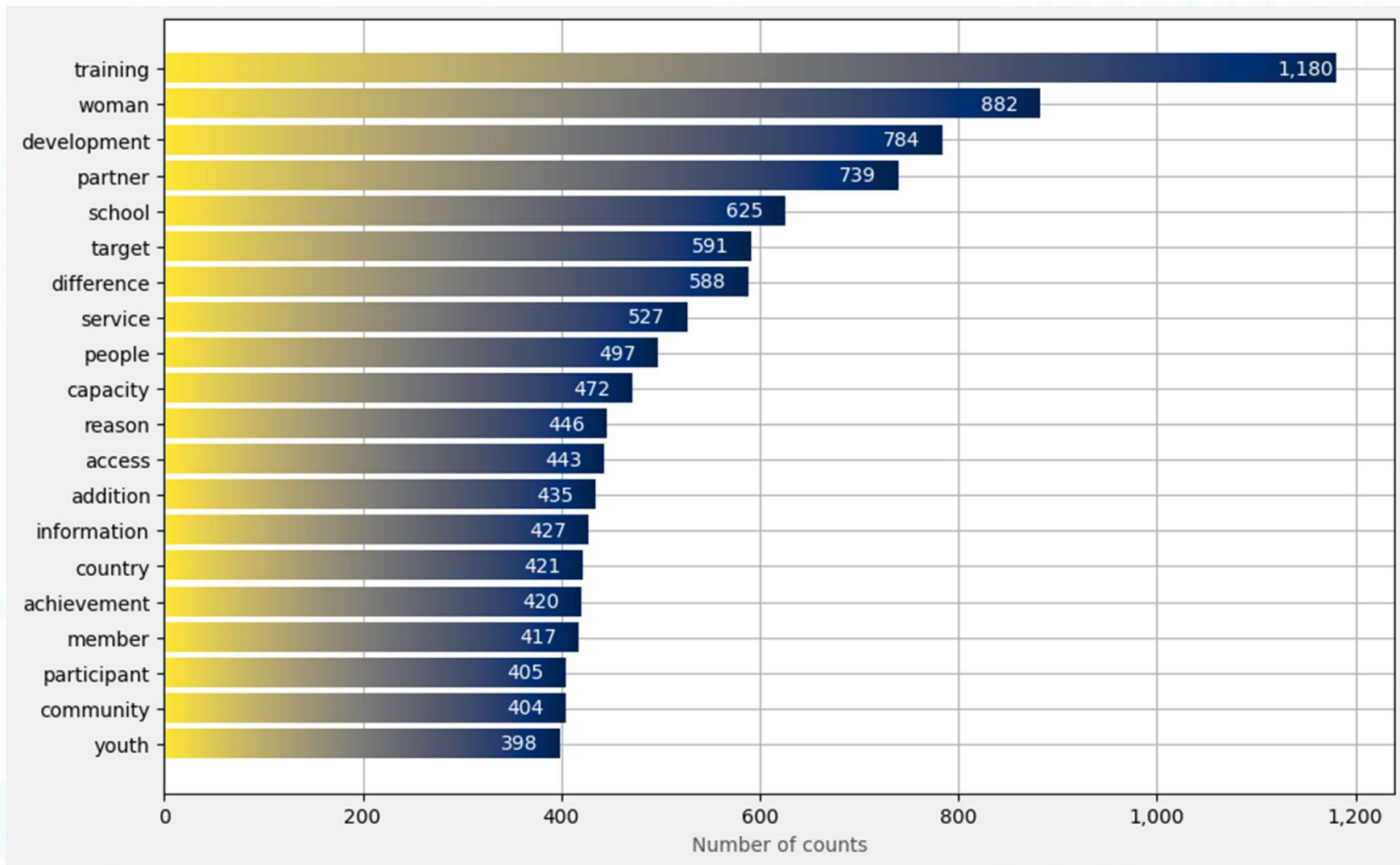
# Content analysis



## Top 20 nouns in reference to human rights in English documents



## Top 20 nouns in reference to human rights in English documents with low HRBA ambitions



**Q&A**



**dav|consulting**



# Data science and AI in monitoring and evaluations

Data science and AI approaches offer huge opportunities to transform M&E systems by enabling practitioners to process and analyse vastly more data in less time at reduced cost and enrich their methodological toolkits through computer-assisted automation.

# Rules-based approach

Final Draft: 23.11.2022  
EVA-11

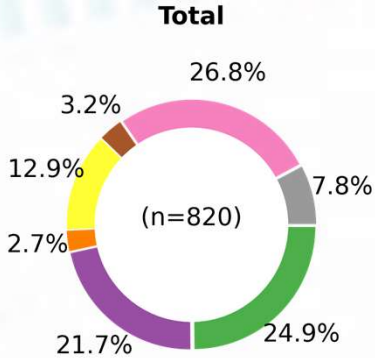
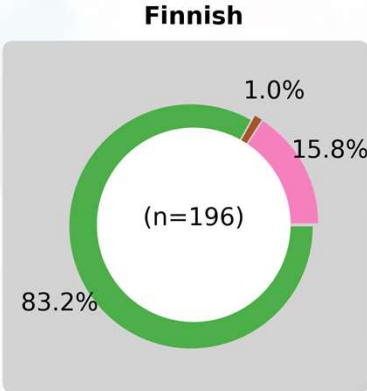
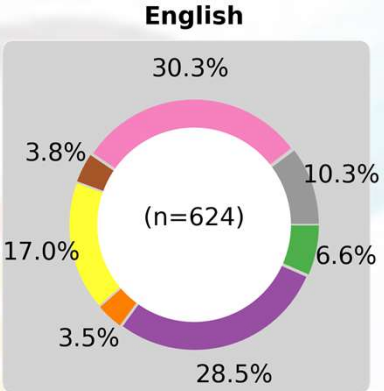
## Analytical framework for data science - labels and criteria for levels of HRBA ambition used by MFA

Computed assisted rule-based analysis on selected document (project plans/document/proposals and latest annual narrative report/final report/progress report). The labels represent the evidence expected to be found as the shared minimum “common denominators” irrespective of cooperation instrument, channel or type of partner organisation.

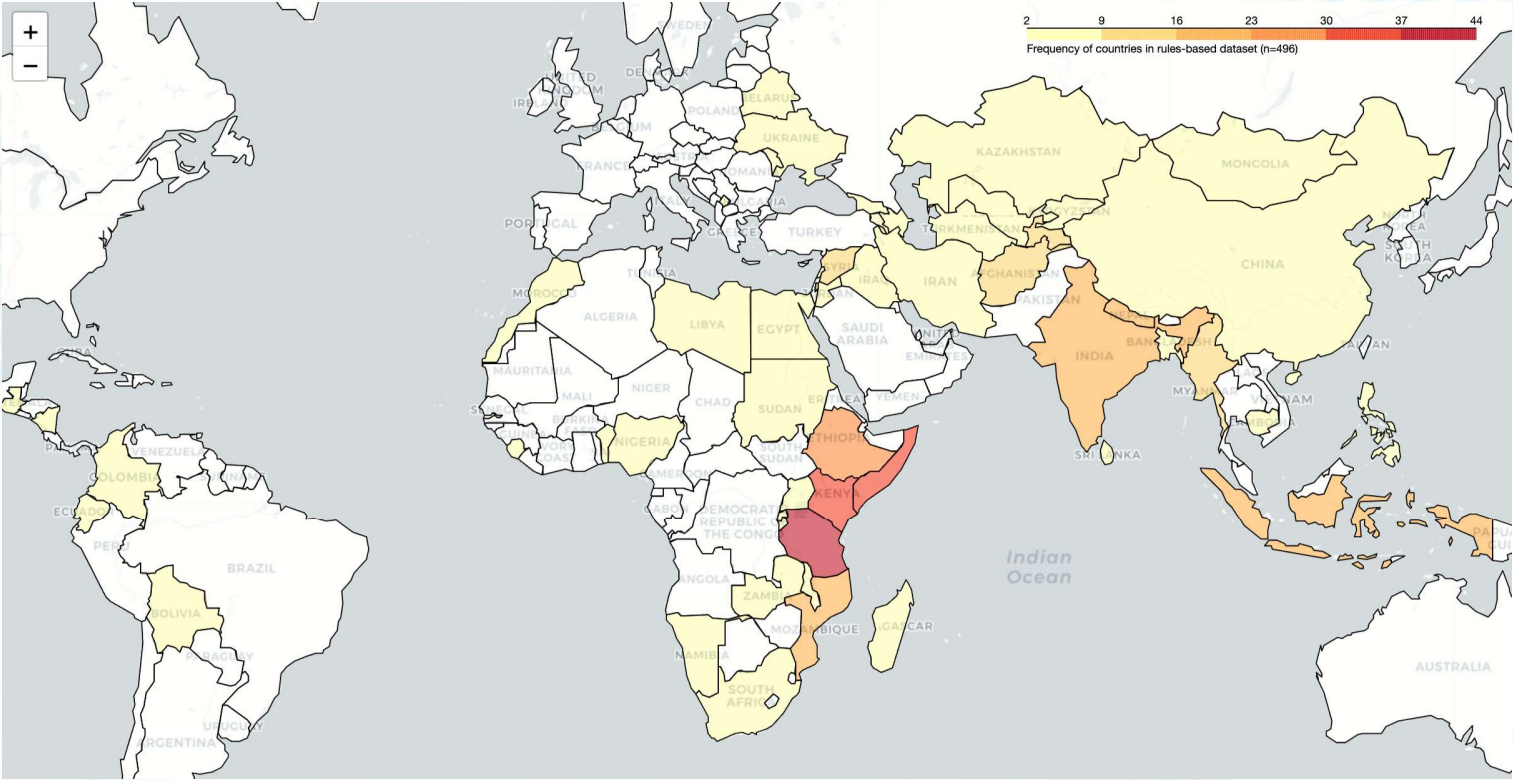
Criteria	EQ1.1 (project plan); EQ 1.2 (annual report)	Data labels (English)	Data labels (Finnish)
<b>Human rights blind</b> – i.e. cases that do not fulfill the minimum criteria for human rights sensitive  Description: The development intervention is ignorant of human rights and the risk of unintentional harmful effects has not been assessed.	No mention of the labels at sensitive level	(see the one below)	(see the one below)
<b>Human rights sensitive</b>  Description: Human rights principles guide the programming, implementation, monitoring and evaluation of the intervention. A basic human rights assessment has been carried out in order to be sufficiently aware of the human rights situation. This is done to avoid unintentional negative effects on the enjoyment of human rights and to ensure that the intervention does not contribute to discriminatory structures, norms and practices. The intervention does not have an explicit commitment to human rights in terms of expected results. Elements related to capacity development or advocacy may be included in the intervention.	<b>At least one mention</b> of label a)  <b>or</b> minimum of 1 of b)  in project plan (EQ 1.1.) in annual report (EQ 1.2)	a) [human rights principle*] <b>OR</b> b) [universal*] [interrelated*] [indivisibility] [inalienable] [indivisible] [accountab*] [transparency] [participation] [inclus*] [equality] [*discriminat*] [principle*]	a) [ihmisoikeuseriaa*] <b>OR</b> b) [universaali*] [yleismaailmallisuus*] [keskinäisriippuvuus*] [jakamattomuus*] [luovuttamattomuus*] [vastuuvollisuus*] [tilivelvollisuus*] [läpinäkyvyys*] [osallisuus*] [osallistumi*] [osallistavuus*] [tasa-arvo*] [yhdenvertaisuus*] [syrjit*] [syrjintä*]

# Data coverage by cooperation instrument and language

- Bilateral support
- CSO and INGO cooperation
- Country programmes
- Funds for Local Cooperation
- Institutional cooperation
- Multilateral support
- Private sector and development policy investments



# Geographical coverage of interventions analysed



## Means and modes of HRBA levels by cooperation instrument and document type

Rules-based

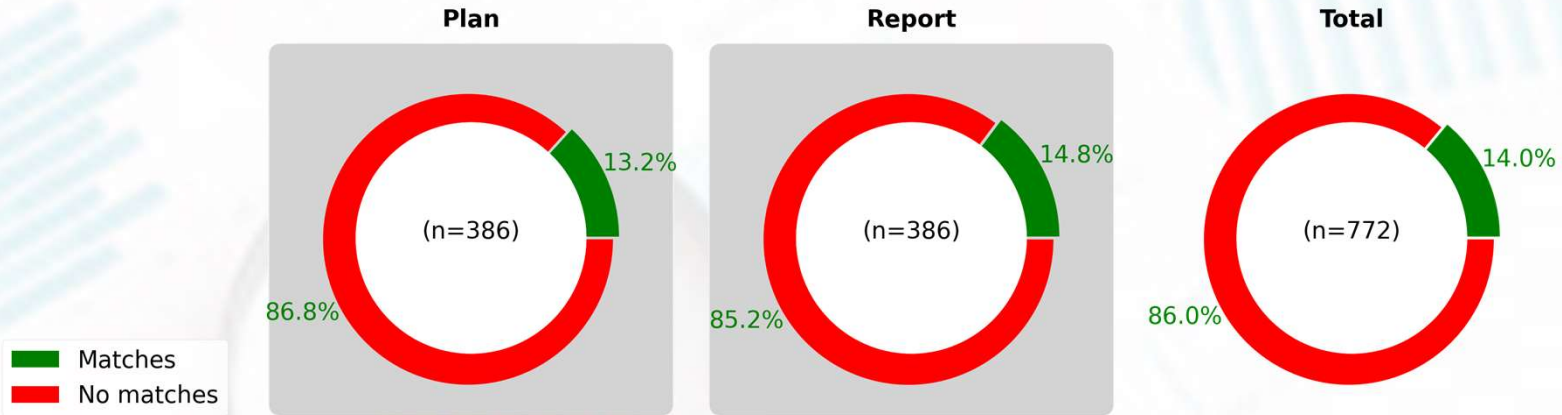
	Plan		Report		Total	
	mean	mode	mean	mode	mean	mode
Bilateral support	3.5	(4)	3.8	(4)	3.6	(4)
CSO and INGO cooperation	3.7	(4)	3.4	(4)	3.5	(4)
Country programmes	3.4	(4)	4.0	(4)	3.7	(4)
Funds for Local Cooperation	3.8	(4)	3.2	(3)	3.5	(4)
Institutional cooperation	3.2	(3)	3.0	(3)	3.1	(3)
Multilateral support	3.5	(4)	3.5	(4)	3.5	(4)
Private sector and development policy investments	1.2	(1)	1.0	(1)	1.1	(1)
<b>Total</b>	3.0	(4)	2.7	(4)	2.8	(4)

Machine learning

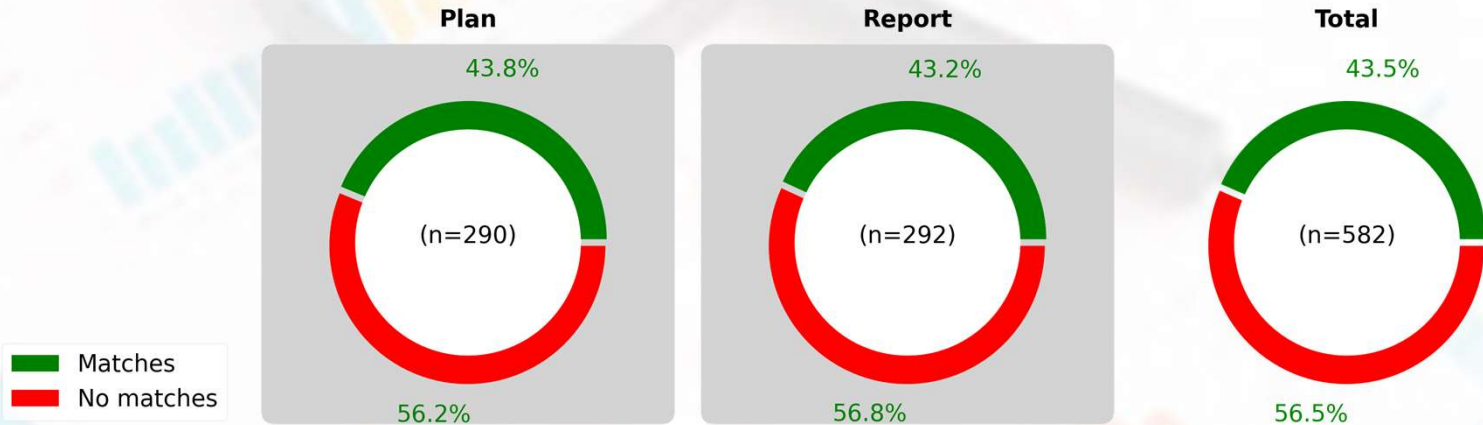
	Plan		Report		Total	
	mean	mode	mean	mode	mean	mode
Bilateral support	3.1	(4)	3.0	(3)	3.1	(3)
CSO and INGO cooperation	2.9	(2)	2.7	(3)	2.8	(2)
Country programmes	3.3	(4)	2.5	(2)	2.9	(2, 3)
Funds for Local Cooperation	3.1	(4)	2.9	(3)	3.0	(3)
Institutional cooperation	2.1	(2)	2.0	(2)	2.0	(2)
Multilateral support	3.2	(4)	3.1	(3)	3.1	(4)
Private sector and development policy investments	2.1	(2)	1.9	(2)	2.0	(2)
<b>Total</b>	3.0	(2)	2.8	(3)	2.9	(2)

# Share of HRBA levels matching self-assessment by document type

Rules-based



Machine learning



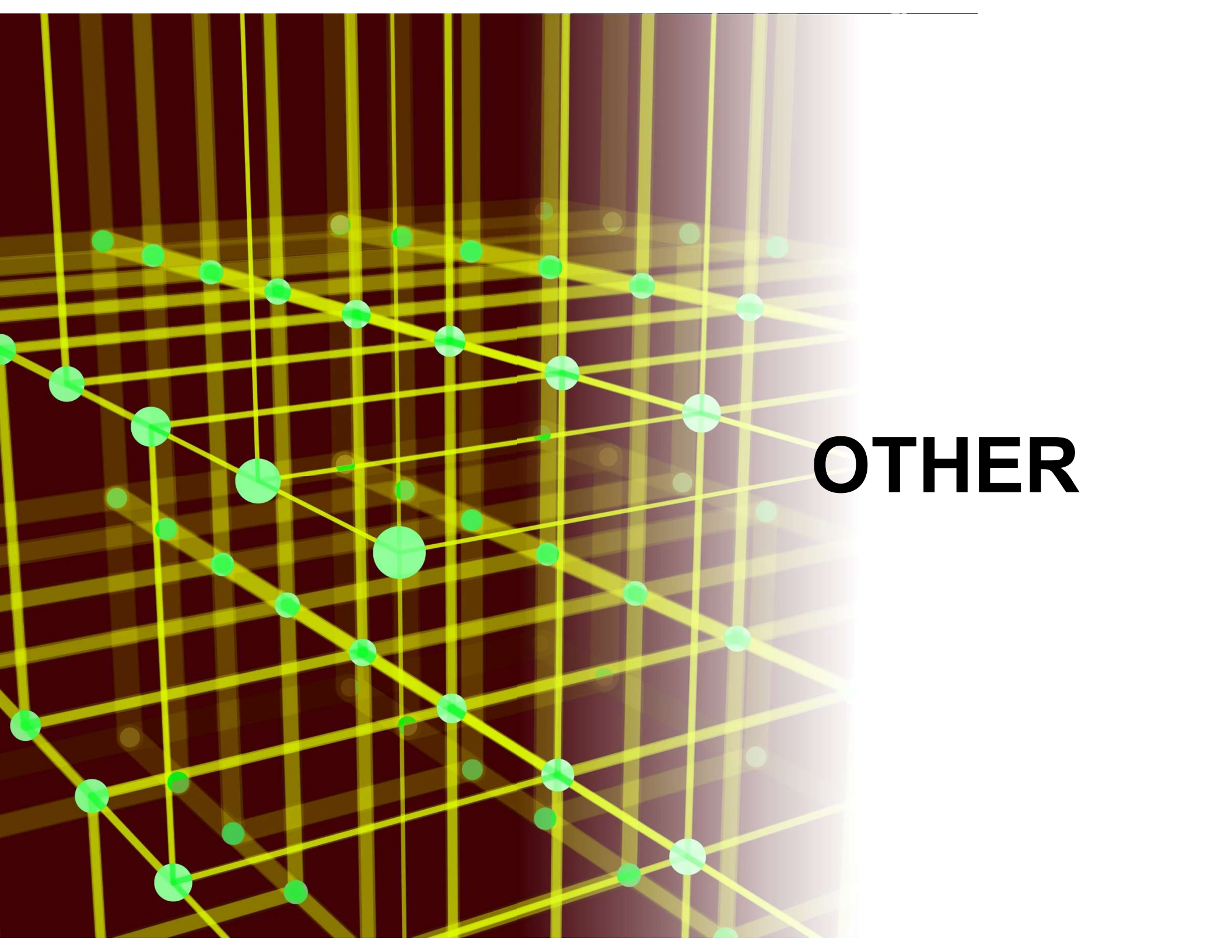
## Limitations and caveats

- Quality of the mentioned co-designed analytical framework and how it reflects the context at hand. Several rounds of iteration and refinement between the consultant and the evaluation management team, also in consultation with other MFA staff, were made.
- Reliability of the data that has been processed: A) Designed algorithm's ability to parse out relevant excerpts (with limited noise). B) Language use. It is the reporting entities' reports (and their language) that has been processed, which might not necessarily reflect the underlying operations perfectly.
- Volume, quality and representation of the HRBA example texts for the pre-trained language model.
- The degree of difficulty. The number of HRBA classes/categories and their nature will affect the accuracy of both applied approaches - rules-based and machine learning.
- Robustness of the results. There is no clear-cut way to determine the exact accuracy of the results from the different approaches. This results from the fact that there is no single available source of truth.

## Food for thought

- Good automated approach for handling large corpus of text – boost speed and consistency.
- Efficient iteration research.
- Accuracy of the classification not determined. Results should be seen as model estimates (which varies). Use overlapping estimates as best estimates. The lack of a single source of truth makes training and testing these types of assessments difficult. In the future, it is suggested to contemplate whether this can be established for a subset of the data.
- Use best estimates as steppingstone to probe deeper and find sample cases in the qualitative evaluation.
- Limit the number of classes and/or establish classes that are semantically and more objectively different. In this case, the used labels - HRBA levels of ambitions - are very similar and, to a certain extent, relatively arbitrary. A rule of thumb is that if human struggles with the interpretation of the categories, an NLP approach will also.
- Use report templates for planning and reporting of funded interventions if possible. This will improve the performance of data science techniques.





**OTHER**



Requirements for adopting and  
deploying data science and AI in  
monitoring & evaluation

# DATA SCIENCE LANDSCAPE

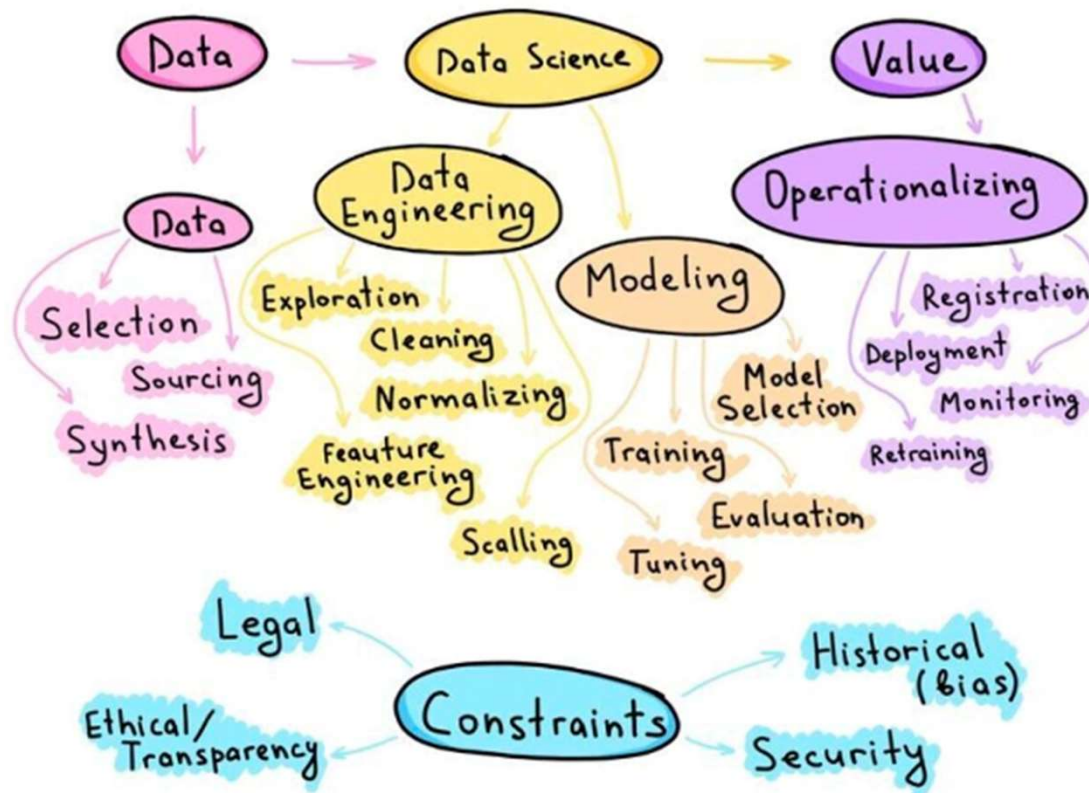


@code.know  
Credit: @data.professor

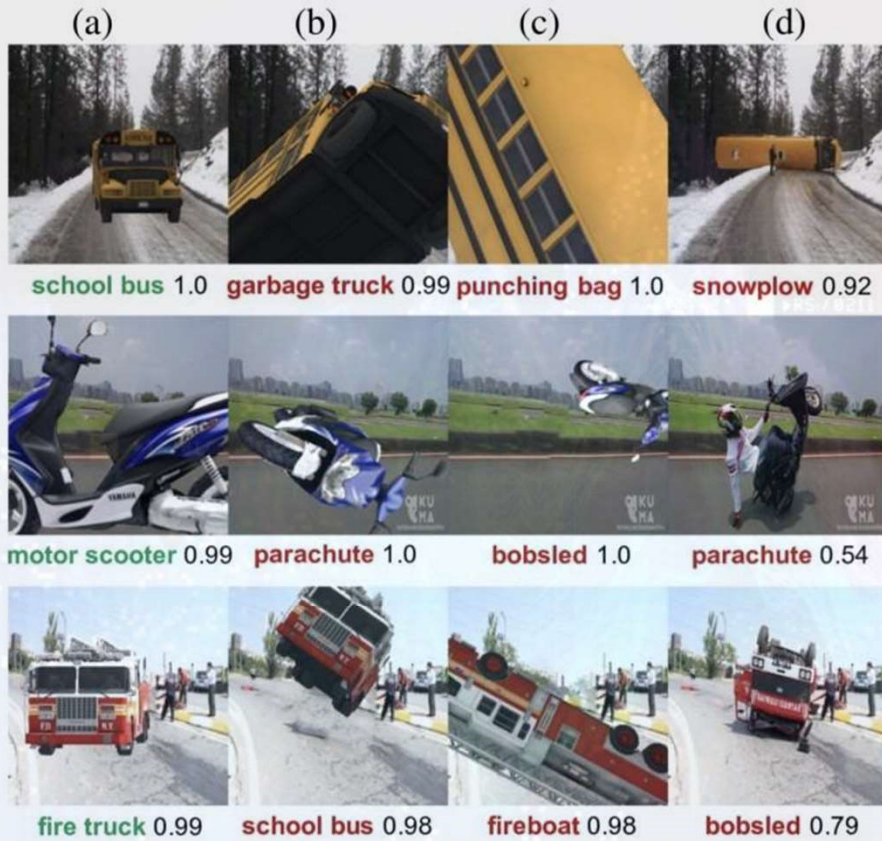
# WHAT COMPANIES THINK A.I. LOOKS LIKE



# WHAT IT ACTUALLY IS



# Caveats



## Areas where we work

Providing data science and fit-for-purpose analytic services and products:

- Data collection (*surveys, text mining, web extraction, API, etc*).
- Analysis (*data manipulation, inference, text analysis, connectivity assessment etc*).
- Communication (*reports, dashboards, interactive online applications*).
- Advisory (*incl. developing processes for commissioning/delivering data analytics services*).