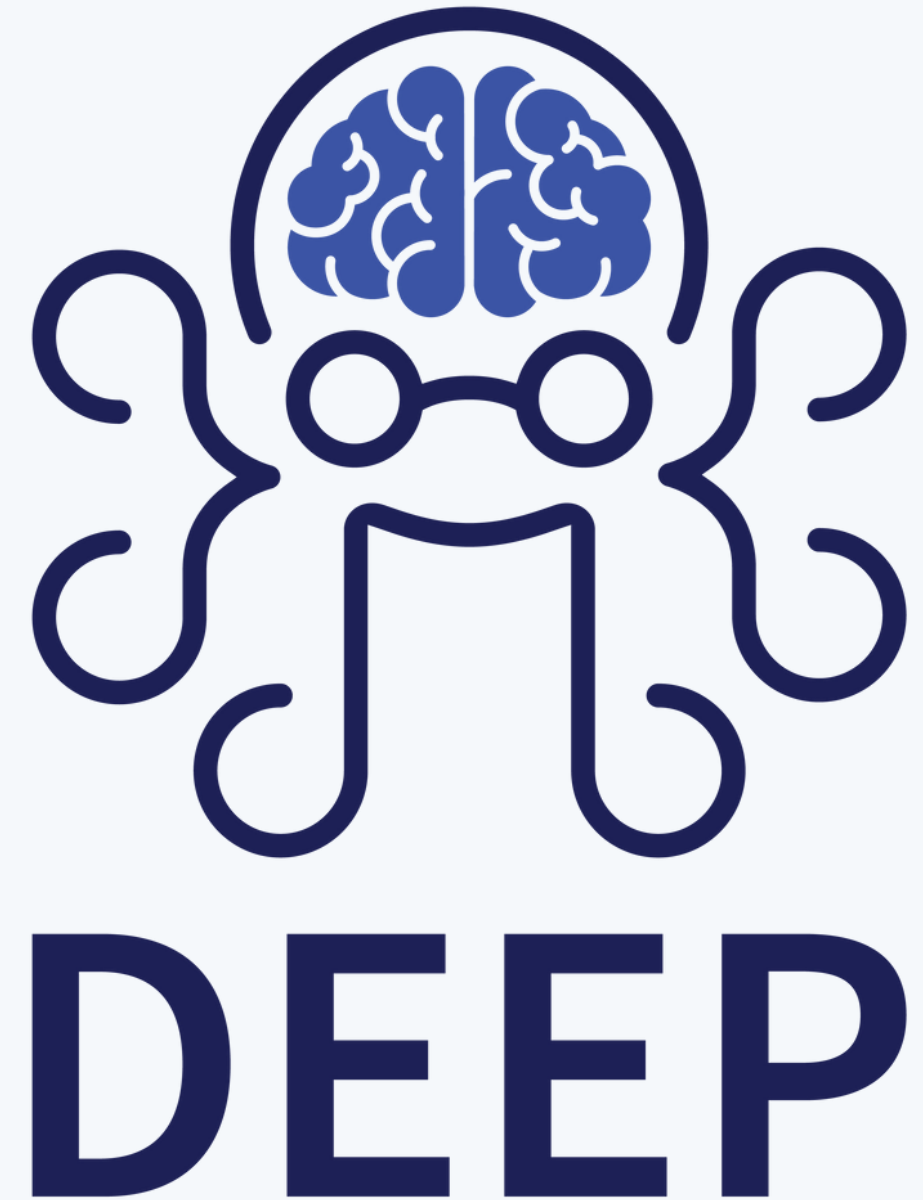


# DATA ENTRY & EXPLORATION PLATFORM

HuMEL

Emerging Technologies in humanitarian Settings

July 16<sup>th</sup>, 2024





# ABOUT DEEP

DEEP, the Data Entry and Exploration Platform, is a joint open-source initiative offering:



An intelligent online platform supporting collaborative humanitarian analysis



A suite of sense making tools for qualitative and unstructured data promoting transparency and data sharing



Customizable analytical framework allowing for diverse uses: secondary data review, media monitoring, protection analysis, operational learning

# GOVERNANCE BOARD



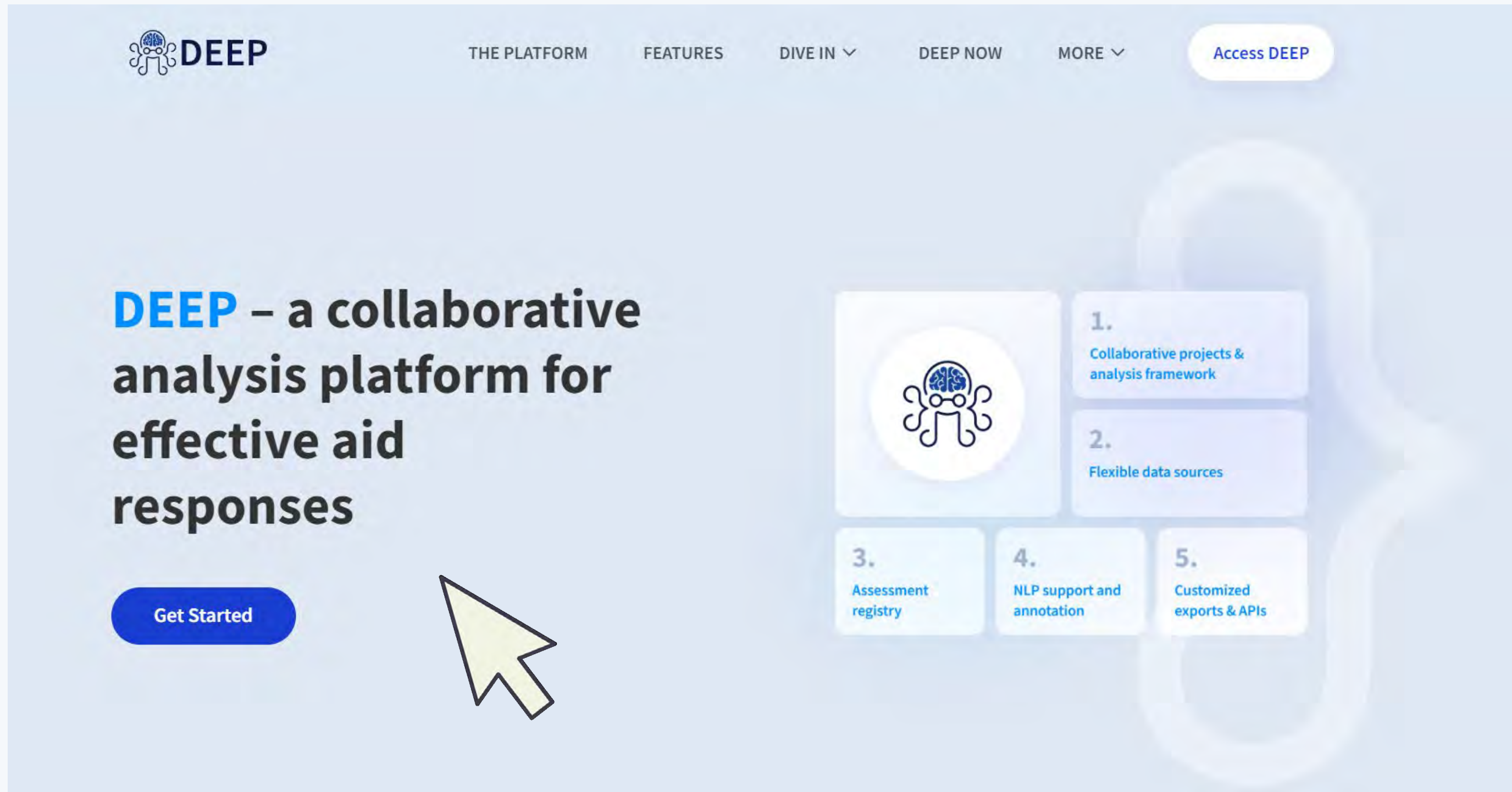
SECRETARIAT:



TECHNICAL PARTNER:



# THE DEEP PLATFORM



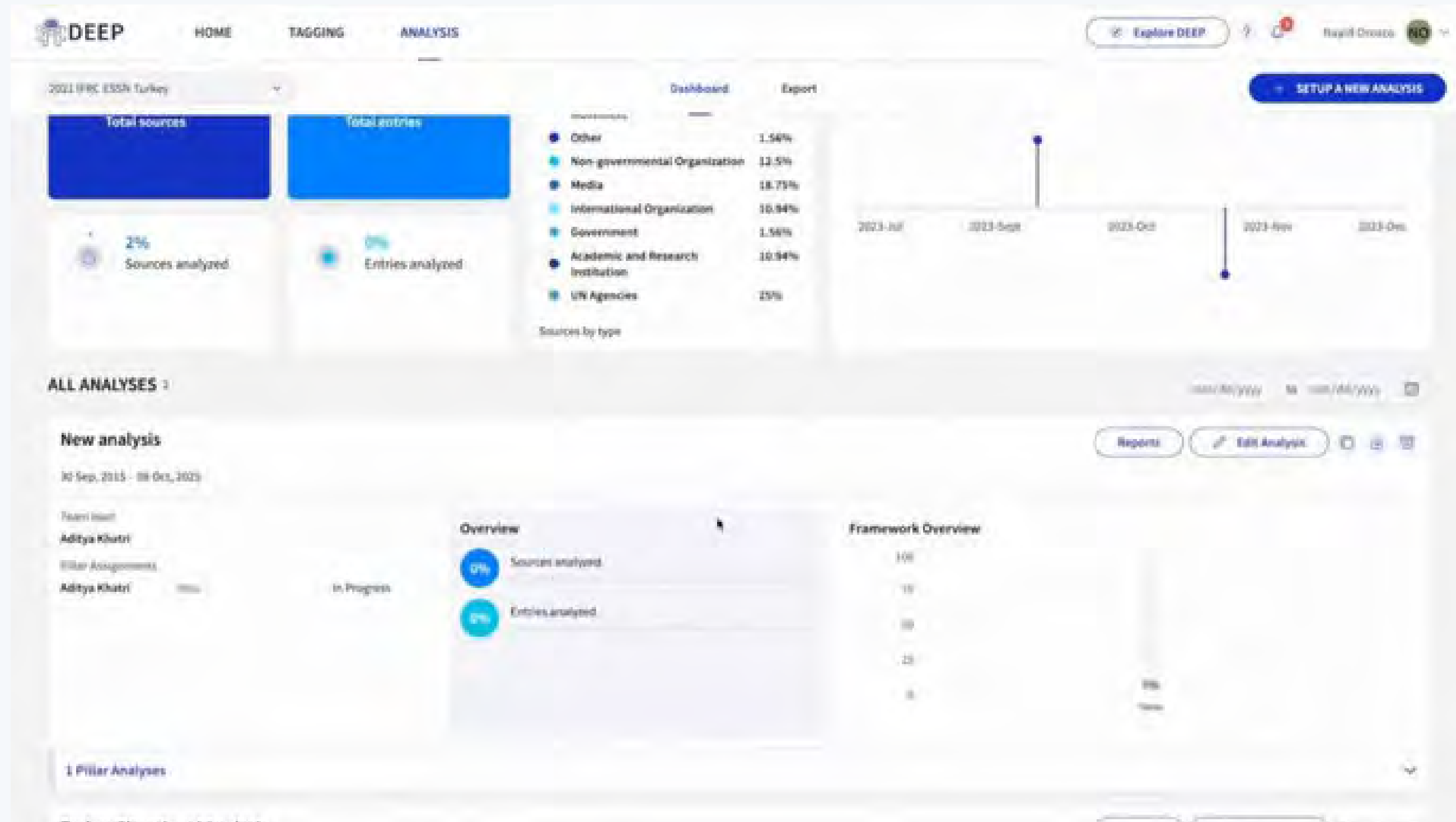
The image shows a screenshot of the DEEP platform website. At the top left is the DEEP logo, which consists of a stylized brain icon with circuit-like lines extending from it, followed by the word "DEEP" in a bold, sans-serif font. To the right of the logo is a navigation menu with the following items: "THE PLATFORM", "FEATURES", "DIVE IN" (with a downward arrow), "DEEP NOW", "MORE" (with a downward arrow), and a white button with rounded corners labeled "Access DEEP".

Below the navigation menu, on the left side, is a large text block: "DEEP – a collaborative analysis platform for effective aid responses". The word "DEEP" is in blue, and the rest is in black. Below this text is a blue button with rounded corners labeled "Get Started". A yellow mouse cursor is pointing at the "Get Started" button.

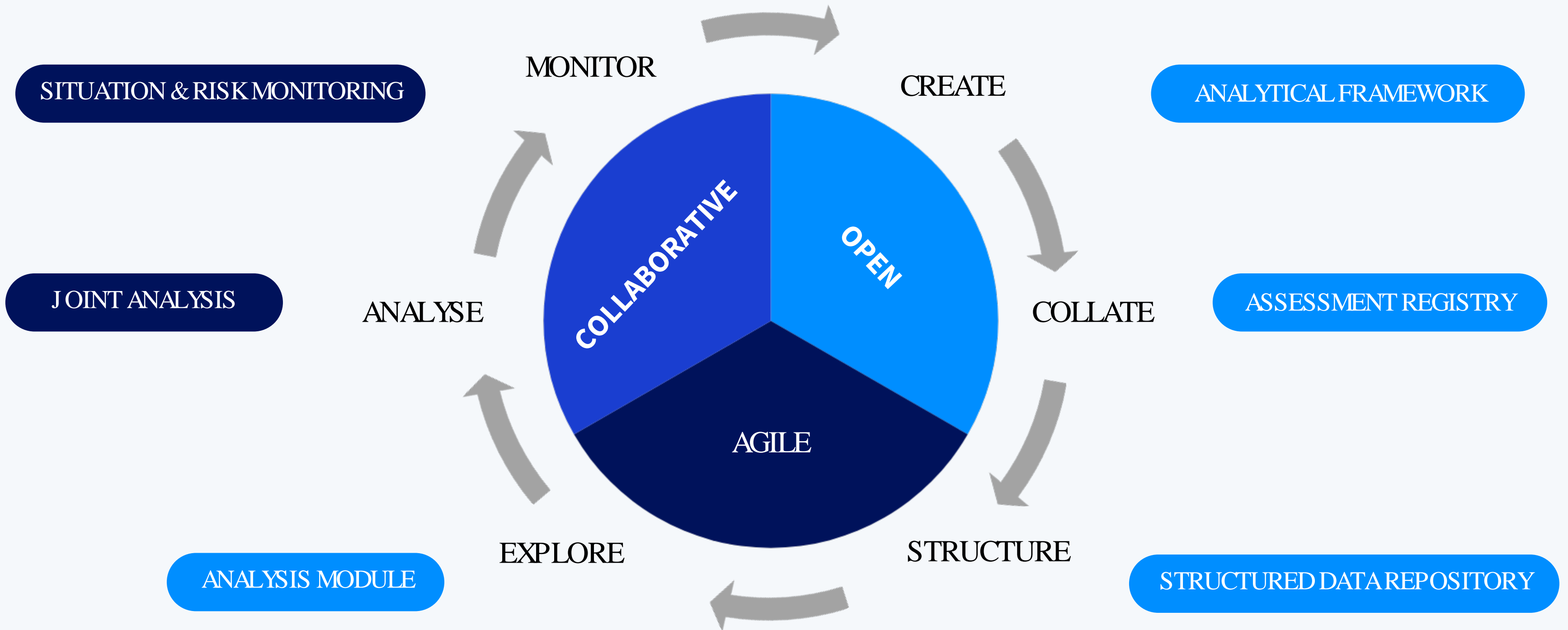
On the right side of the page, there is a grid of five feature cards. The first card is larger and contains the DEEP brain icon. The other four cards are smaller and arranged in a 2x2 grid. Each card has a number and a description:

- 1. Collaborative projects & analysis framework
- 2. Flexible data sources
- 3. Assessment registry
- 4. NLP support and annotation
- 5. Customized exports & APIs

# THE DEEP PLATFORM

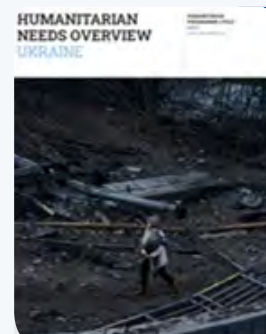


# PRINCIPLES & ANALYSIS FLOW



● UNDER DEVELOPMENT    ● IMPLEMENTED

# USES OF DEEP AND ITS DATA



## [Ukraine: Humanitarian Needs Overview \(HNO\) 2023](#)

- DEEP used as data repository
- DEEP-based products used as secondary sources



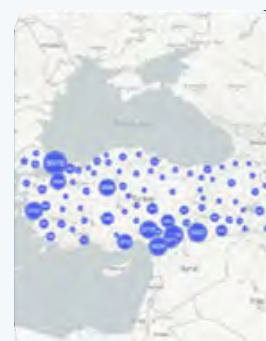
## [Sudan: Situation Analysis reports](#)

- [Situation reports](#)
- Scenario building
- DEEP used as data repository
- Analysis module and NLP features



## [Ukraine: Assessments mapping](#)

- DEEP used as assessment registry



## [Turkey: Media monitoring](#)

- Most extensive refugee-related library worldwide
- Coverage of humanitarian, academic and media information



## [Sahel: Protection monitoring](#)

- PAF-DEEP Project, led by DRC
- Data repository for Sahel countries
- SDR to inform PAU (Protection Analysis Updates)
- Informing joint analysis processes and HNO

# DEEP's Open Source and AI Innovations



## DATA COLLECTION

- Chrome extension
- Online vs local sources
- Connectors



## DATA STRUCTURING

- Analytical frameworks
- Tagging & Assessments registry
- NLP in DEEP: Automatic entry extraction



## VISUALIZATION & EXPORT

- DEEP Dashboards
- Excel & Word exports



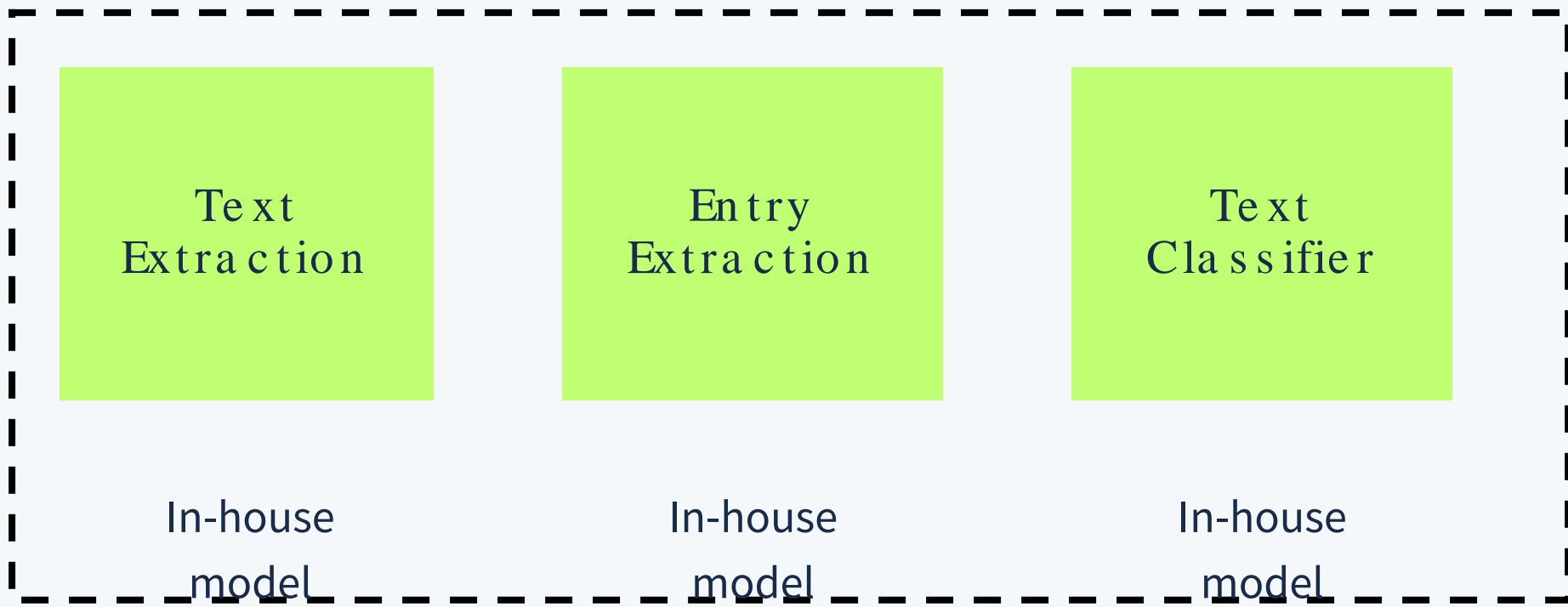
## ANALYSIS MODULE

- Concept
- NLP Features
- Reporting Module

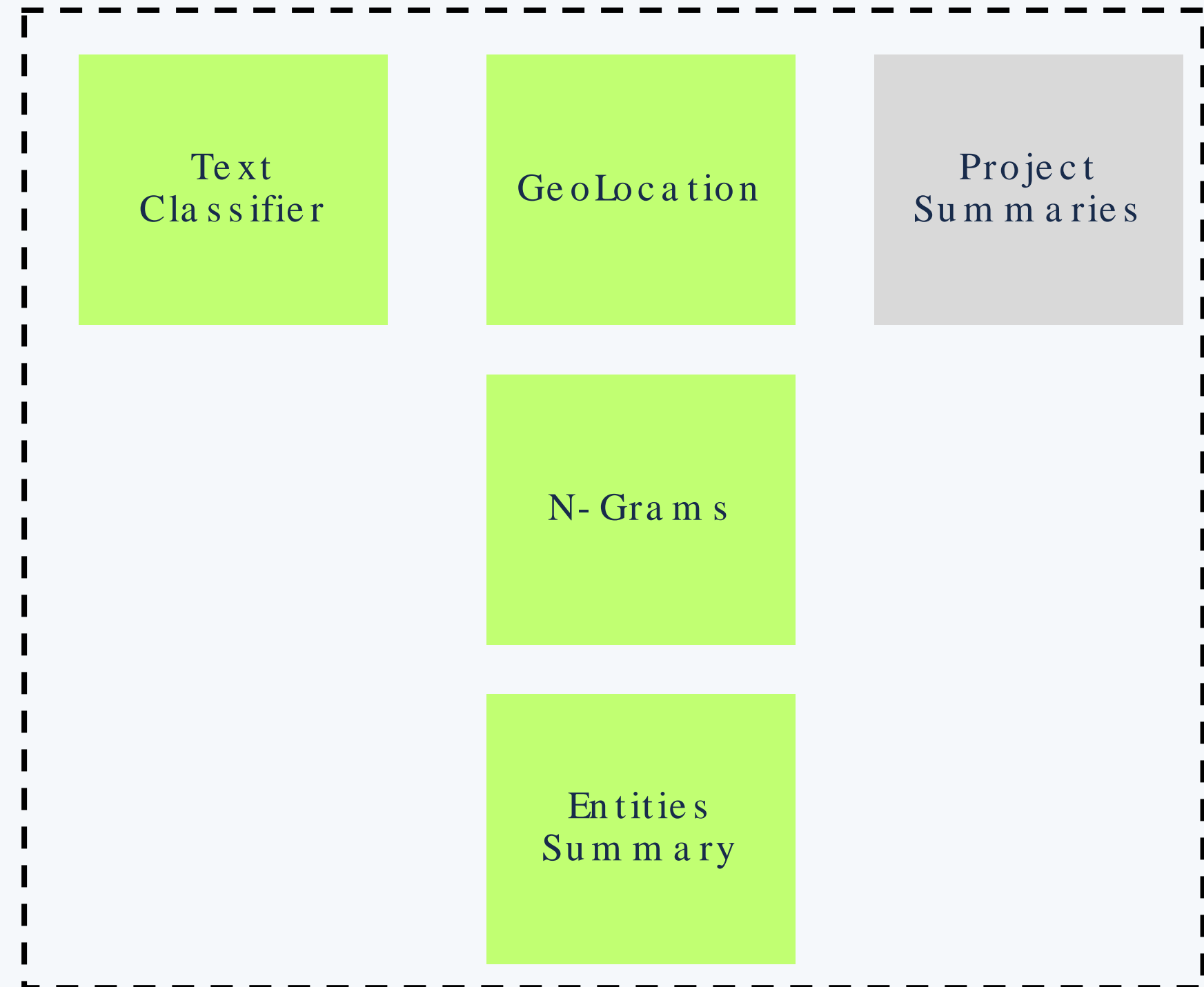


# DEEP NLP features

## Tagging pipeline



## Analysis pipeline



# DEEP NLP features



## AUTOMATIC EXTRACTION AND CLASSIFICATION

This feature streamlines text extraction and classification, boosting efficiency for secondary data analysis.



## TOPIC MODELING

DEEP's analysis module utilizes topic modeling to automatically group entries into distinct stories. Entries are clustered based on the similarity of the topics they refer to, enhancing data organization.



## AUTOMATIC SUMMARIZATION

Within the analysis module of DEEP, the auto summarization task automatically generates concise summaries from a selected number of entries within the same story. This feature provides humanitarian analysts with efficient and informative summaries.



## AUTOMATIC GEO-TAGGING

Optimizing a model that extracts geographic information from the text, making it more accurate and appropriate for the humanitarian sector.

Ongoing  
research



Summarization



Zero-shot classification



Quantitative extraction  
from text

# DEEP NLP Research

## Humanitarian BERT

HumBert is the first language model adapted to humanitarian topics. It is an XLM-Roberta model trained on humanitarian texts - approximately 50 million textual examples (roughly 2 billion tokens) from public humanitarian reports, law cases and news articles.

On Hugging Face: [huggingface.co/nlp-the-deep/humbert](https://huggingface.co/nlp-the-deep/humbert)



## HumSet

HumSet is a novel and rich multilingual dataset curated by humanitarian analysts and covers various disasters around the globe that occurred from 2018 to 2021. The dataset consists of approximately 17,000 annotated documents from 46 humanitarian response projects.

Manually extracted and tagged by humanitarian analysts

46 projects

~17k documents

~149k snippets of text

3 languages: English, French, Spanish

# USEFUL RESOURCES

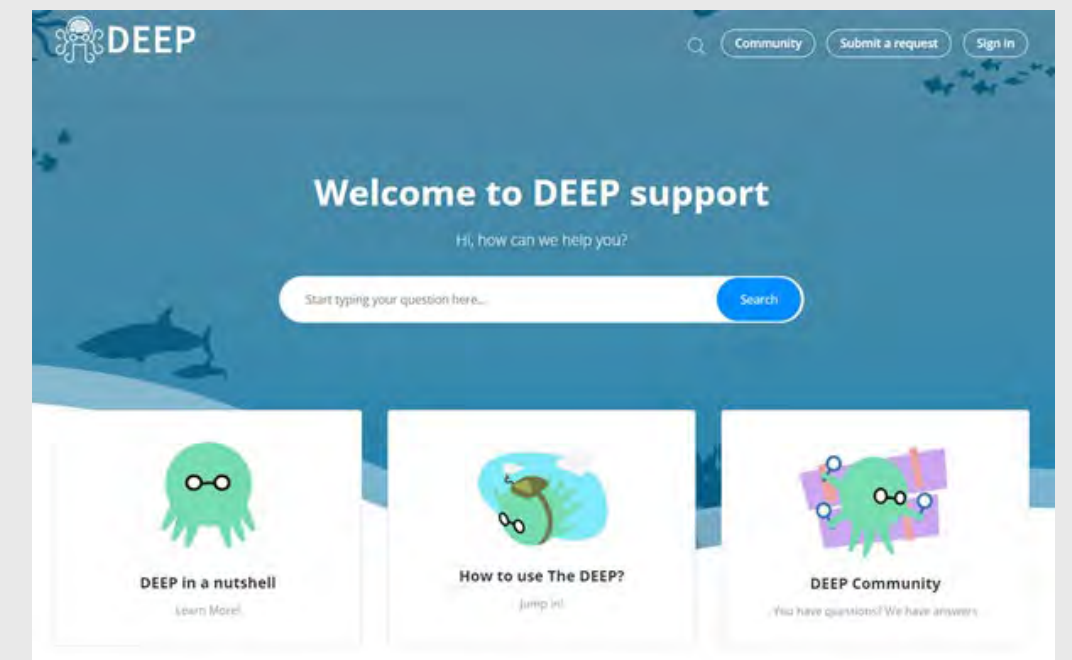


## TRAINING VIDEOS

- [DEEP Support Desk](#)
- [Explore Dashboard](#)
- [Creating a Project](#)
- [Analytical Framework](#)
- [Adding Sources and Tagging](#)
- [Understanding and Exporting Data](#)
- [Analysis Module](#)
- [Reporting Module](#)
- [Automatic Entry Extraction](#)
- [New Assessment Registry](#)

## USER GUIDE

<https://deephelp.zendesk.com/hc/en-us>





# DEEP'S LATEST NEWS

## IN THE MAKING

- Improved Assessment Registry

1  
Analysis Module  
with automatic  
summarization

2  
Support & training

- [New Zendesk portal](#)
- DEEP Trainer's Package
- [Kaya Course](#)

3  
Reporting Module  
[Introduction to the Data Entry and Exploration Platform \(DEEP\)](#)

4  
Automatic  
Entry  
Extraction



# CONTACT US

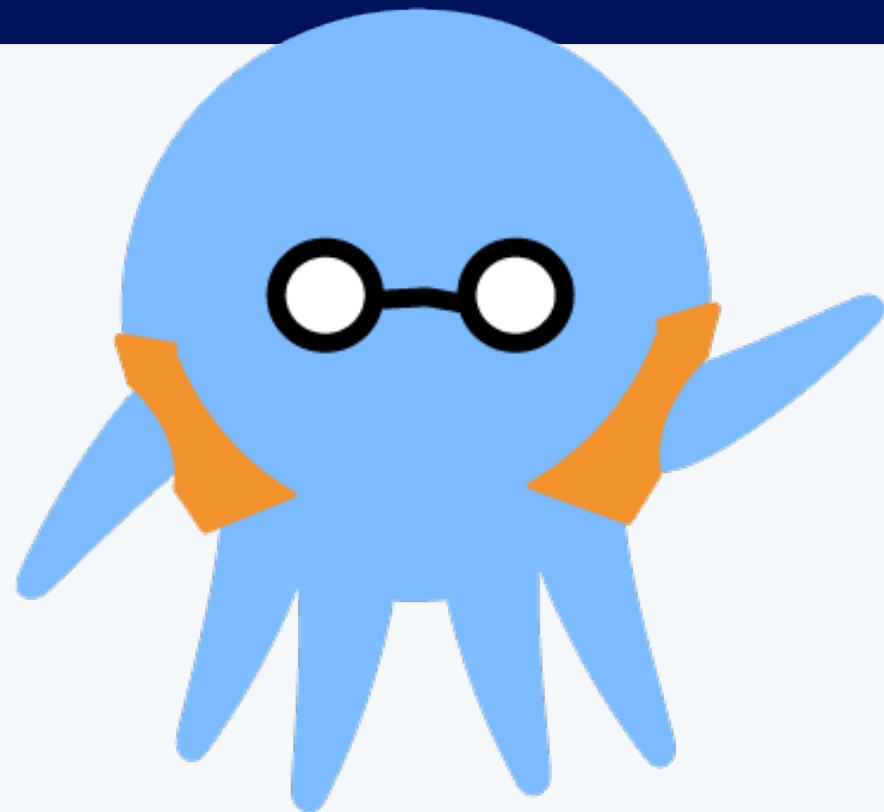
DEEP Project Manager: [pm@thedeep.io](mailto:pm@thedeep.io)  
Technical Support: [support@thedeep.io](mailto:support@thedeep.io)

Doubts, questions, suggestions, informal comms:  
[jose@datafriendlyspace.org](mailto:jose@datafriendlyspace.org)  
[nayid@datafriendlyspace.org](mailto:nayid@datafriendlyspace.org)



# THANK YOU

FOLLOW US  
TO STAY UPDATED



[thedeep.io](https://thedeep.io)  
[app.thedeep.io](https://app.thedeep.io)  
[nlp.thedeep.io](https://nlp.thedeep.io)



[@theDeep\\_project](https://twitter.com/theDeep_project)



[theDEEP](https://www.linkedin.com/company/theDEEP)

# DEMONSTRATION OF DEEP

DEEP Capacity Building  
Lead – Jose 'Pepe' COBOS





# HuMEL 'Emerging Technologies in humanitarian Settings' with DEEP



Ebru EREN WEBB

German Red Cross International Cooperation  
Division Information Management Team DEEP Expert

# The DEEP IFRC Turkiye Earthquake : Sudden Onset Crisis Project Example 1



**Goal:** collect, analyze and report on the ongoing crisis



**Target audience:** IFRC Turkiye Country Office, **all involved humanitarian actors**



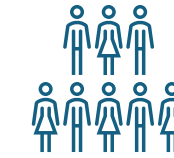
**Coverage:** all relevant humanitarian and media documents,



**Coverage: 1,068 documents** at the end of 6 months from February to July 2023



**Outputs:** Weekly, then bi-weekly situational updates, 19 regular briefings, 4 thematic reports



**Personel:** 5 taggers, 2 analysts, 2 IFRC Staff and 1 Surge



**Success story** with a system **to be replicated in all similar sudden onset crisis situations**

# The DEEP IFRC Turkiye Earthquake : Sudden Onset Crisis Project Example 1

## IFRC IM & OPERATIONS BRIEFING PROCESS

Established framework to enable Turkiye Earthquake operational response



### Product Request and feedback

**HEOps** requests regular morning briefings based on SDR and analysis products – provides regular feedback on utility for operation

### Delivery

**IM team lead** receives briefing deck in morning delivers situation briefing at start of every coordination meeting and Joint Task Force



### Data Collection and Analysis

**IM Qualitative Analysis team** produce secondary data reviews and situation briefs released publicly through joint DEEP activation with DFS



### Synthesis and Reporting

Reports handed over to **SIMS Coordinator** to synthesize together with visualizations and other available info to produce briefing deck for ops. coordination meeting during "night shift"

# The DEEP Hurricane Beryl Project: Sudden Onset Crisis Project Example 2



**Goal:** Collect, categorize, report data to inform operations



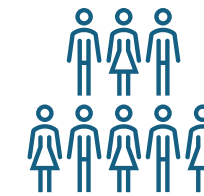
**Target audience:** IFRC & HNSs and the entire humanitarian sector



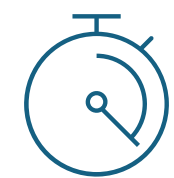
GRC IM involvement with two IM staff



**Outputs: 600 documents, three situational update reports published publicly in two weeks**



Professional tagging team of **5 people and 2 analysts** funded by the DEEP itself



**Current stage:** Active daily coordination with IFRC via the UN Analysis Cell