

# NZINGA: Leveraging LLMs to navigate complex language in clinical trial informed consents forms

Mbangula Lameck Amugongo | Biostatistics & Data Sciences (Boehringer Ingelheim)

Belfast, Northern Ireland | 16 June 2026

# Disclaimer

© 2026 Boehringer Ingelheim International GmbH. All rights reserved.

This presentation and its contents are property of Boehringer Ingelheim and are, inter alia, protected by copyright law. Complete or partial passing on to third parties as well as copying, reproduction, publication or any other use by third parties is not permitted.

# Agenda



Introduction

---



Problem statement

---



Methods

---



Results

---



Takeaway

---



Conclusion

# Introduction

## What is Informed Consent Form (ICF) review for data reuse?

- A structured review of ICFs to determine whether clinical trial data may be reused beyond the original study purpose
  - Identifies permissions, restrictions and conditions for secondary use and data sharing
  - Ensures that any reuse of data is aligned with what participants have explicitly consented to



### Why it matters?

- Protects participant rights by ensuring data is used only within the scope of consent
- Ensures legal and regulatory compliance, reducing risk related to GDPR, ethics committees and country-specific rules
- Enables data sharing by making secondary use legally and operationally possible

# Problem Statement

## Manual ICF review is labour and time intensive process

- Review take appr. 15 minutes per ICF

**01**

ICF format  
not standardised

**02**

Different  
languages

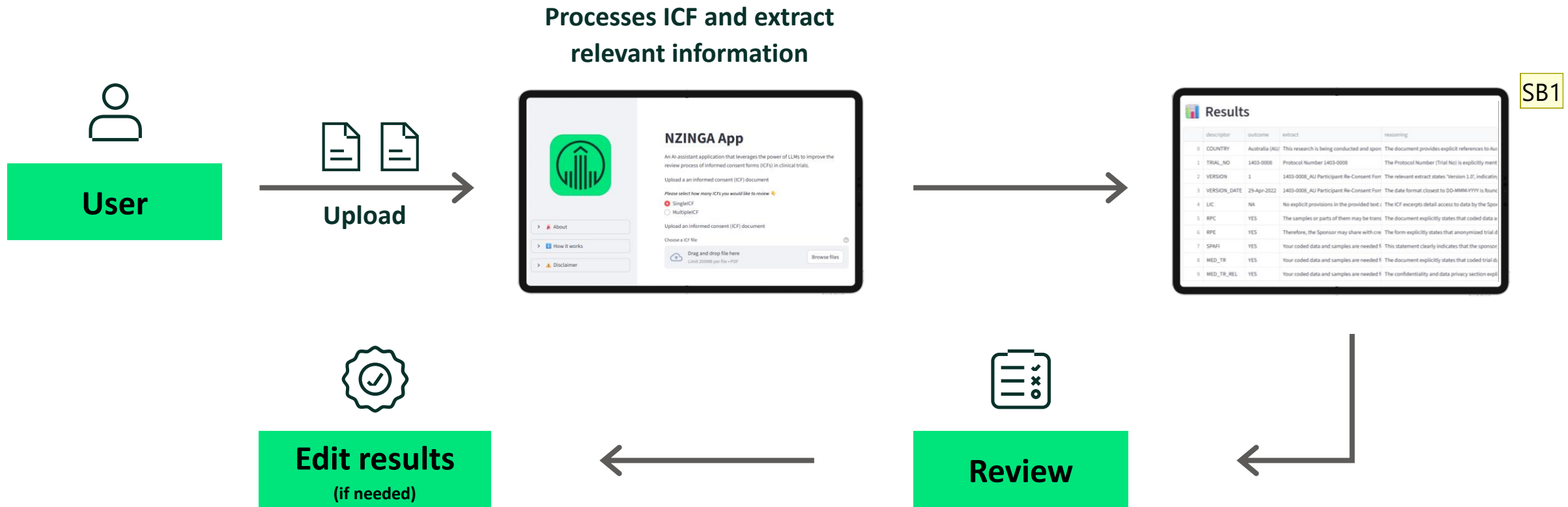
**03**

Variability, i.e., IRB,  
country requirements

- 2400 ICF documents are reviewed annually
  - 600 hours (75 working days)
- ICF Review is essential but only the first step in enabling data sharing

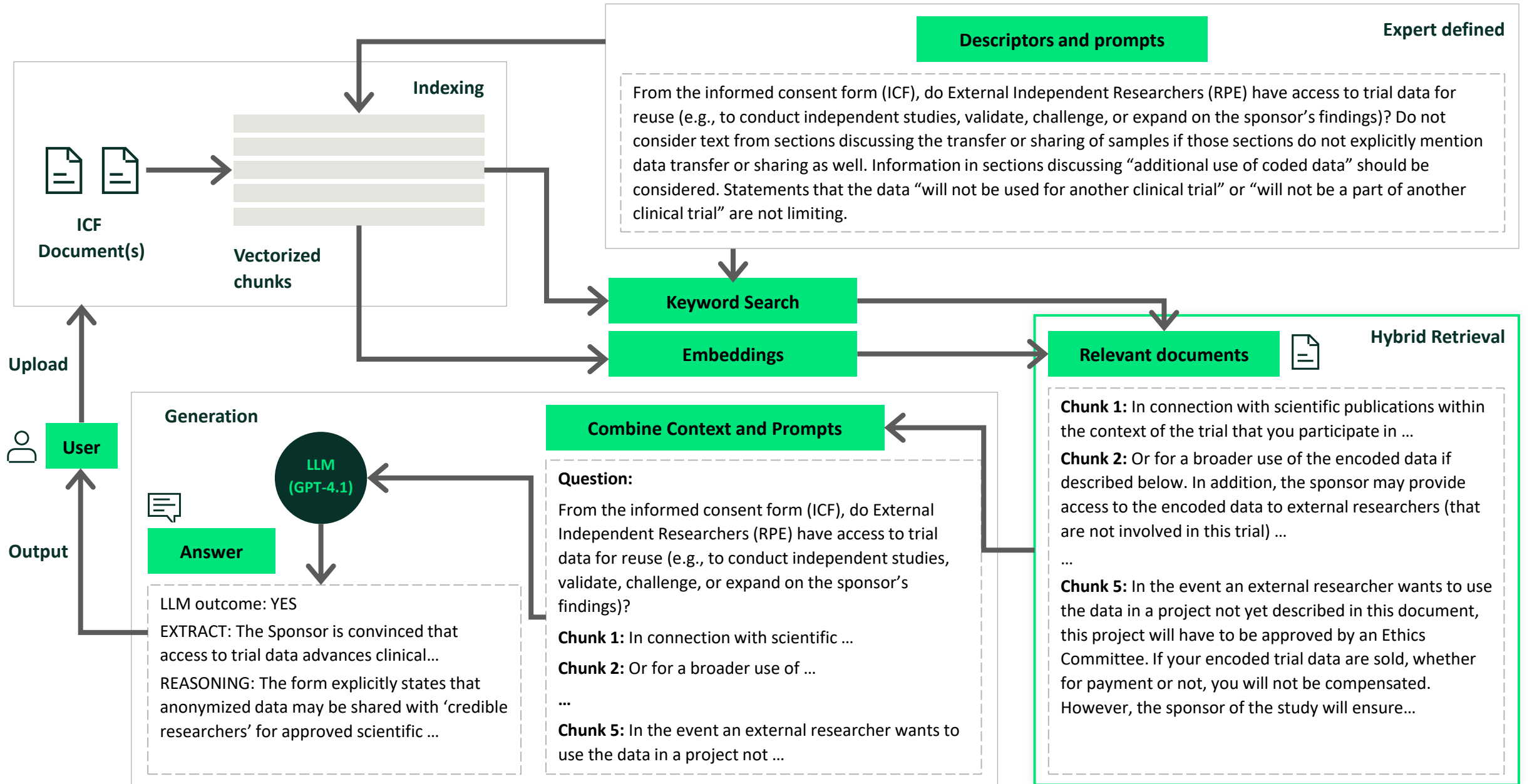
# Nzinga

An AI-assistant tool to improve the review process of **Informed Consent Forms (ICFs)** in clinical trials



SB1

# Methods



# What did we evaluate?

Category	Descriptor	Definition	Outcome
Recipient	LIC	Owners of a licensed product.	YES   NO   NA
Recipient	RPC	Research partners and collaborators of the Sponsor.	YES   NO   NA
Recipient	RPE	External independent researchers.	YES   NO   NA
Recipient	SPAFI	Sponsors, affiliates, or third-party processors.	YES   NO   NA
Purpose	MED_TR	Learn about the trial medication.	YES   NO   NA
Purpose	MED_TR_REL	Learn about medications related to the trial medication.	YES   NO   NA
Purpose	MED_ALL	Learn about any medications.	YES   NO   NA
Purpose	DIS_ALL	Learn about any disease.	YES   NO   NA
Purpose	DIS_TR	Learn about the disease studied in the trial.	YES   NO   NA
Purpose	DIS_TR_REL	Learn about diseases related to the trial disease.	YES   NO   NA
Purpose	DIS_TR_TA	Learn about diseases in a specific therapeutic area.	YES   NO   NA
Purpose	DPROD_ALL	Develop any diagnostic tools.	YES   NO   NA
Purpose	DPROD_TR_REL	Develop diagnostic tools related to the trial disease.	YES   NO   NA
Purpose	TRPROD_ALL	Develop any therapeutics.	YES   NO   NA
Purpose	TRPROD_TR_REL	Develop therapeutics related to the trial disease.	YES   NO   NA
Purpose	QUAL_ALL	Improve the quality of future clinical trials.	YES   NO   NA
Data Retention	DLINK30	Subject re-identification link deleted after 30 years.	YES   NO
Data Retention	RR50	Data retained for up to 50 years.	YES   NO
Data Retention	RR80	Data retained for up to 80 years.	YES   NO
Specific Text	RESTRICTIVE_WORDING	Text includes other restrictions on data reuse.	YES   NO
Specific Text	GENOMIC_INFO	Text mentions genomic data.	YES   NO

Overall accuracy computed per descriptor

Class-level - YES/NO/NA

- YES:** Text supports data reuse
- NO:** Explicitly prohibits reuse
- NA:** No relevant context provided

## Binary class (YES / NO)

- **Data retention:** DLINK30, RR50, RR80 **Specific language:** GENOMIC\_INFO, RESTRICTIVE\_WORDING
- **NO = prohibition or no mention**

## Why it Matters

- Class-level analysis reveals how well the system distinguishes
- Affirmative     Ambiguous     Negative statements

# Statistical Analysis

## Evaluation

**01** 438 ICF documents



*Assess the performance sensitivity to various chunk sizes*

Chunk sizes: 500 | 1000 | 2000

### Metrics

- **Accuracy:** LLM outputs compared against **expert-validated ground truth**
- **Completeness:** No ground truth for supporting text
  - Manual expert review

## Validation

**02** 488 ICF documents



*Validate RAG pipeline performance across diverse ICF documents and regions*

### Metrics

- **Accuracy:** LLM outputs compared against expert-validated ground truth
- **Completeness:** No ground truth for supporting text
  - Manual expert review
- **Reasoning Quality:** Soundness and interpretability of LLM explanation
- **Relevance:** Appropriateness of the extracted information

# Results

## Evaluated the application

- Processing 438 ICF documents
- Nzinga achieved average 81.6% across all descriptors



### Validation

Further optimisation, processed 488 documents

**01**

Accuracy of 90%  
across descriptors

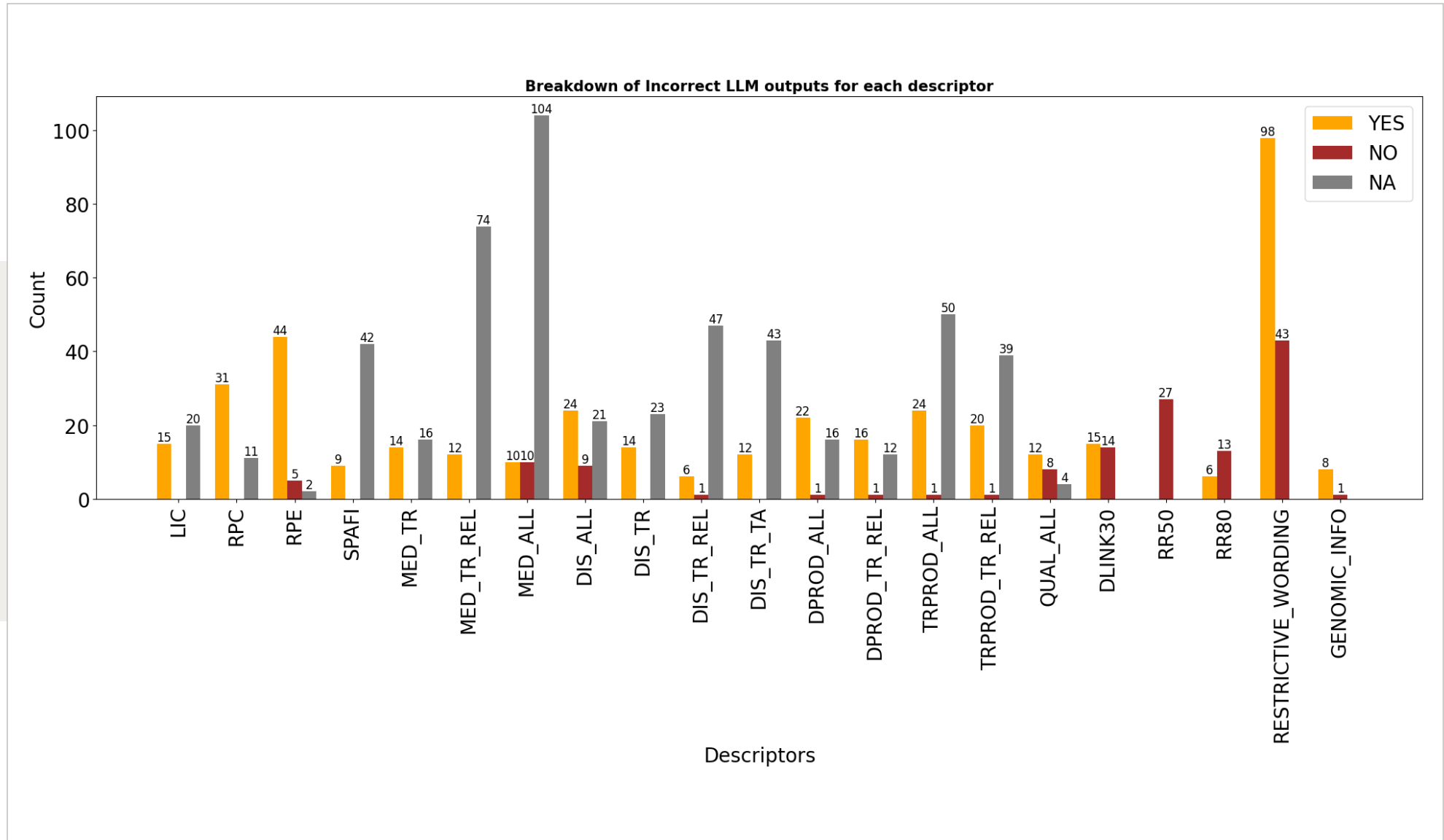
**02**

186 ICF completely  
correctly analysed

- ✓ Process optimization: Less than 3 minutes to process and extract information
- ✓ Reducing evaluation time and SME can focus on extracted relevant chunks

# Results

✓ 488 ICF  
processed





> About

> How it works

> Disclaimer

## Preview of the uploaded document

Biobanking Assent Form

127%

PDF

Extract descriptors

✓ Processing completed successfully!

✦ Processing descriptor: OTHER\_DATA\_RET

## Results

	descriptor	outcome	extract	reason
4	LIC	NA	No information available	The p
5	RPC	NA	No information available	The p
6	RPE	NA	No information available	I care
7	SPAFI	NA	No information available	The p
8	MED_TR	NA	No information available	The p
9	MED_TR_REL	NA	No information available	The p
10	MED_ALL	NA	No information available	The p
11	DIS_ALL	NA	No information available	The p
12	DIS_TR	NA	No information available	The p
13	DIS_TR_REL	YES	The purpose of this optional study is to collect blood samples, and information from	The f
14	DIS_TR_TA	NA	No information available	The

## Final Results after applying business rules

	lic	rpc	rpe	spafi	med_tr	med_tr_rel	med_all	dis_all	dis_tr	dis_tr_rel
	na	na	na	na	na	na	na	na	na	yes

Data frame to edit LLM outcomes

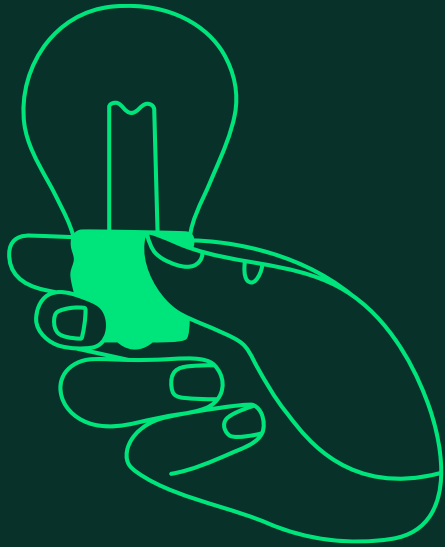
Download Original Results

You can edit the results below. If you make changes, a new download option will appear with comparison sheets.

Download Original Results

Button to download results

# Takeaway



## Scalable AI-assisted ICF review is feasible:

- Workflow efficiently analysed 900+, supporting data-reuse

## Showing strong descriptor performance

- Prompt optimization: Accuracy improved 81.6% to 90%

## Human-in-the-loop by Design

- Workflow enhances expert review rather than replacing it, maintaining oversight, reliability, and regulatory trust



## Ambiguity challenge for some descriptors

- Non-standard language — even humans

Low document level performance: 186 of 488 ICFs were fully correct

- Descriptor accuracy does not automatically translate to flawless document-level performance



Human-centred workflow remains essential for data sharing decisions

# Conclusion



**01**

Nzinga Speed up ICF review process

**02**

Consistent high-quality review of data-reuse language

**03**

Possibilities for adaptability to various business use cases that requires ICF review

# Thank you

## Why not reach out?

[Mbangula\\_lameck.amugongo@boehringer-ingelheim.com](mailto:Mbangula_lameck.amugongo@boehringer-ingelheim.com)



### Acknowledgments

#### Boehringer Ingelheim

- Claudia Neumann
- Andreas Freisinger
- Jaroslaw Deska

#### Staburo

- Lena Schaller
- R. Maarten van Dijk
- Helene Wendt
- Enrica Zanuttigh



Nzinga pre-print paper