# Argument Extraction Codebook

# 0. Basic Definition

**Task Overview:** Like an English reading comprehension, you are asked to read a paragraph and extract important information. These are some definitions you need to know.

**Abstracts** are the short and concise summary of at the start of a research paper, highlighting its background, methods, results, and implications.

**Event:** **Abstracts** are divided into paragraphs that convey one specific type of information. Each paragraph is named an **Event,** and depending on the type of information it conveys, it is classified by **Event Types**. Each **Event** describes a key aspect of the research clearly and independently.

**Event Types** are different types of **Events:**

**Background/Introduction:** Briefly outlines the context, motivation, and problem being addressed. It highlights the research gap and the paper's objectives or research questions.

**Method/Approach:** Summarizes the methodologies, frameworks, or techniques used to conduct the study, including experimental setups, algorithms, datasets, or analytical tools.

**Results/Findings:** Reports the main outcomes of the research, emphasizing key data, trends, or discoveries. Focuses on what was achieved or learned.

**Conclusions/Implications:** Discusses the significance of the findings, their impact on the field, potential applications, and how they address the initial problem or research gap. May include recommendations or future research directions.

**Summarization:** A summary that best describe the **Event** within 10 words

**Arguments:** Each **Event** of the abstract contains **arguments** to express information. You need to extract **Arguments** from **Events**.

# 1. Argument Definitions

Here we introduce brief definitions for all the Arguments you will use in the task. A more detailed definition is offered in **Section 4**.

## 1.1 Agent

**Definition:** The entity responsible for initiating or performing the **Main Action**, such as a person, system, method, or organization
**Examples:** Researchers applied a neural network to classify images.

## 1.2 Main Action

**Definition**: The main verb or verb phrase, including auxiliary verbs like "am," "is," "are".
**Examples:** Researchers applied a neural network to classify images.

## 1.3 Object(s)

**Definition:** The entity that receives, affected by, or is the focus of the **Main Action**, such as a concept, result, or entity being acted upon.
Primary Object: The primary receiver of the **Main Action**.
Primary Modifier: Describes Primary Object with attached information.
Secondary Object: The Secondary receiver of the **Main Action**.
Secondary Modifier: Describes Secondary Object with attached information.
**Example:** Researchers applied a neural network to classify images.
**Example:** Researchers applied a neural network with 256 layers to a transformer-based language model with adaptive learning rates for improving long-text summarization accuracy.
***More detailed examples in section 3.4***

## 1.4 Context

**Definition:** Provides foundational or situational information.
**Example:** Deep learning has revolutionized natural language processing tasks, enabling state-of-the-art results in translation, summarization, and question answering.

## 1.5 Purpose

**Definition:** Defines the purpose or aim.
**Example:** This study aims to develop a lightweight transformer model suitable for deployment on edge devices.

## 1.6 Method

**Definition:** Techniques, tools, or frameworks used.
**Example:** We employ a combination of knowledge distillation and parameter pruning to reduce model size while maintaining accuracy.

## 1.7 Results

**Definition:** Observations or outputs.
**Example:** The proposed method achieves a 40% reduction in model size with only a 1% drop in accuracy on the GLUE benchmark.

## 1.8 Analysis

**Definition:** Interpretation or explanation of other arguments.
**Example:** The slight decrease in accuracy can be attributed to the removal of redundant parameters that minimally impact overall model performance.

## 1.9 Challenge

**Definition:** Constraints or weaknesses.
**Example:** One significant limitation of the approach is its dependency on large-scale labeled datasets for effective knowledge distillation.

## 1.10 Ethical

**Definition:** Ethical concerns, implications, and justifications.
**Example:** The deployment of these models must address concerns about potential biases in training data that could unfairly disadvantage certain user groups.

## 1.11 Implication

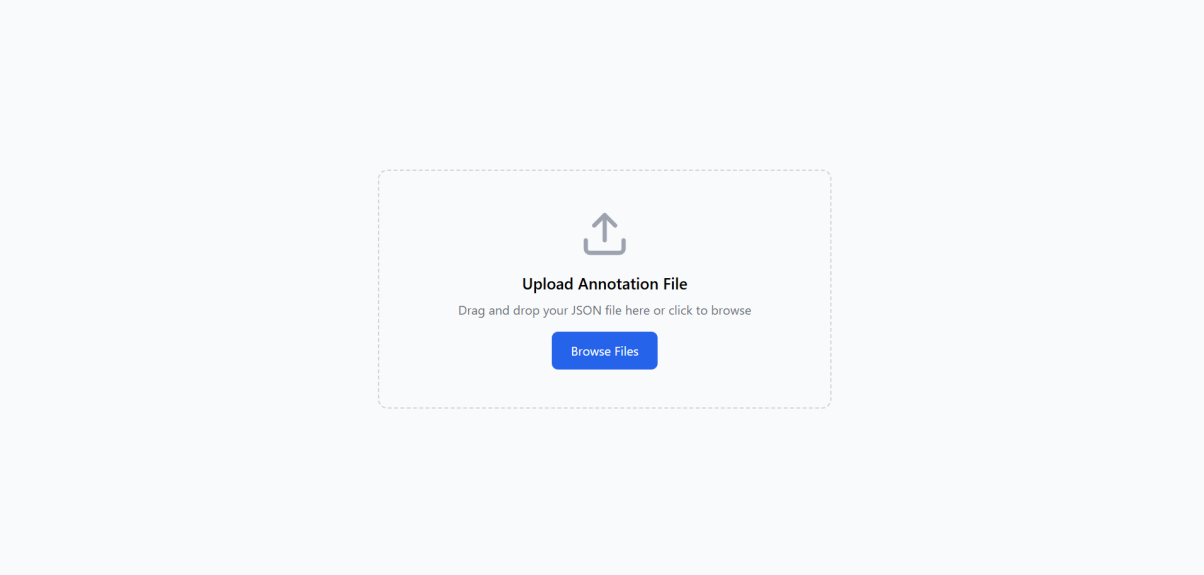
**Definition:** Broader applicability or significance and future research suggestion.
**Example:** Our approach opens the door for deploying advanced NLP models on low-power devices, paving the way for accessible AI in remote or resource-constrained environments.

## 1.12 Contradiction

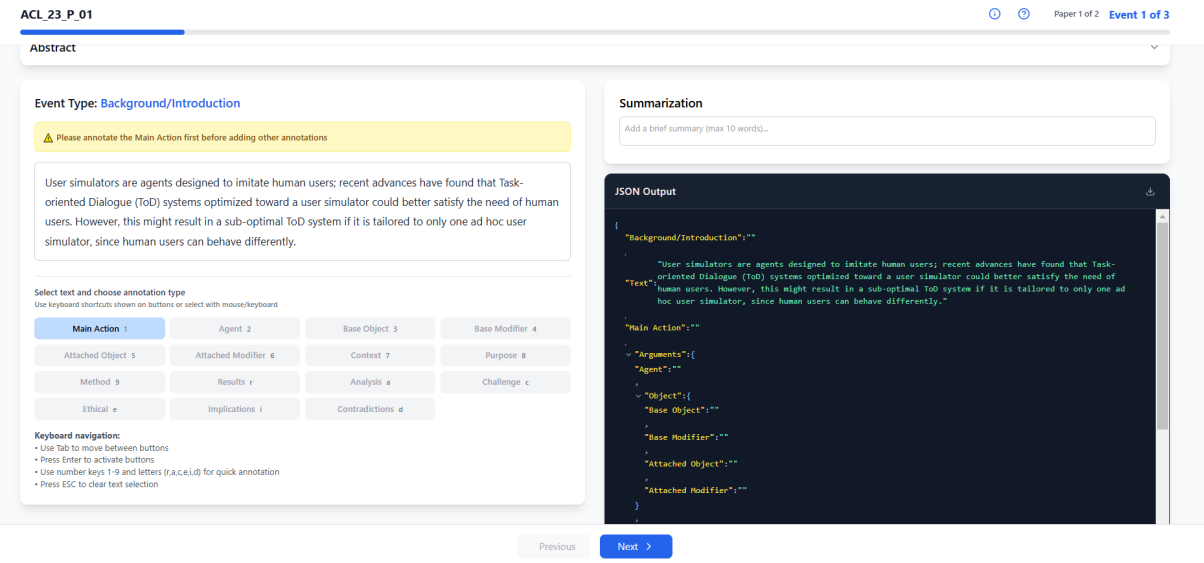
**Definition:** Disagreements challenging existing knowledge.
**Example:** Contrary to previous studies suggesting that parameter pruning significantly reduces accuracy, our results demonstrate minimal performance loss with careful pruning strategies.

# 2. Annotation Guidelines

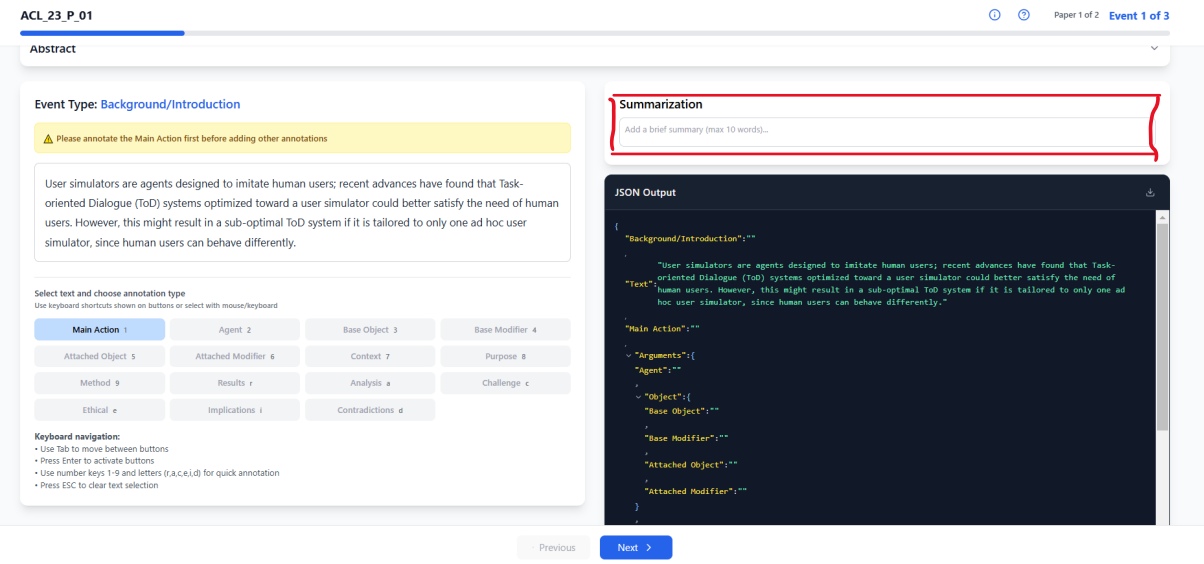
1. Click here to open the tool. The tool looks like this:




2. Open the annotation file provided to you, by clicking “Browse Files”.
3. After opening an annotation file in this tool, you will see something like the below image. Note, each of you will be working on different annotation file, so the content will vary from person to person.




4. After opening the file, first read the paragraph and provide a **Summarization** that best describes the paragraph, highlighted in the box shown below.




5. Then find the **Main Action** from the paragraph that leads the **Event**. It is usually related to the **Summarization** in semantic meaning. **Tip**s: How to choose the most important action word among all the actions? Check the **Event type**, **Main Action** usually helps to expand the **Event type**. (e.g., “propose” for Method/Approach, “demonstrate” for Result/Findings)

6. Then find the sentence where the **Main Action** is in and in most cases, the **Agent** and **all the Objects** are in this sentence. Note that the **Primary and Secondary Modifier** are always after **Primary and Secondary Objects** to provide additional information.
   Further, in the majority cases, there are only **Primary Object and Modifier** in the annotation.

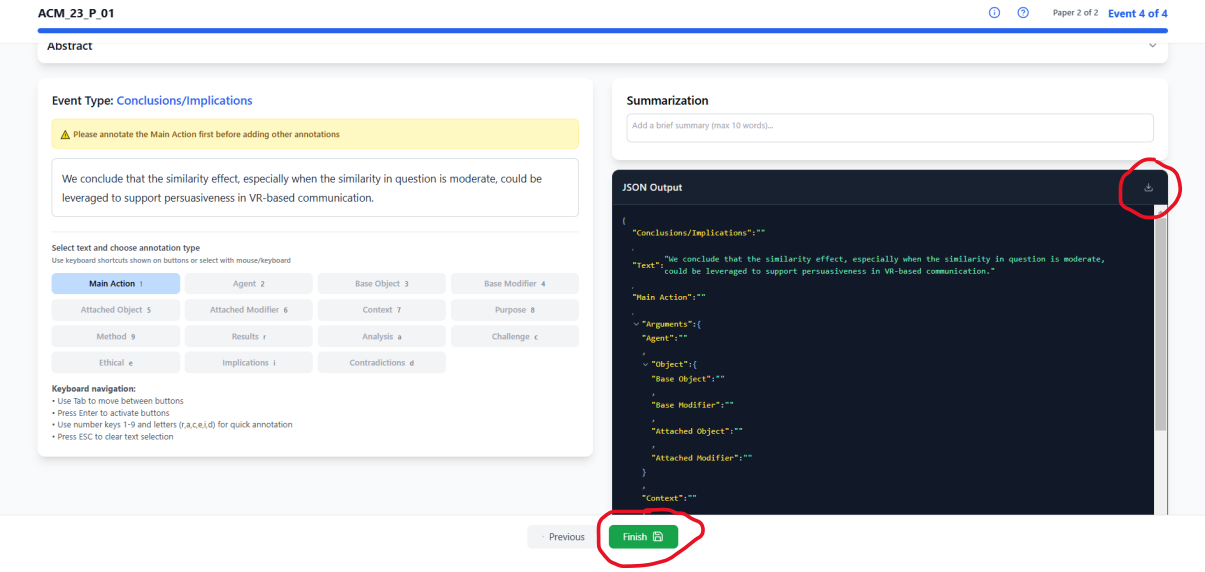
7. Then extract other **arguments**. Please refer to section 4 for detailed definition when you are unsure. Don’t infer any information, only find explicit arguments. You don’t need to fill in all the arguments for each paragraph (the **Main Action** is surely needed), but you are encouraged to cover as much information as possible.

8. All annotations in steps 5,6,7 can be done by selecting text and clicking the corresponding button on the left. All your annotations will be shown in JSON Output to the right as well.

9. You should always try to **breakdown long sentences** to annotate, unless one entire sentence is inseparable and is describing only one argument. Please see detailed explanation here in section 3.1.

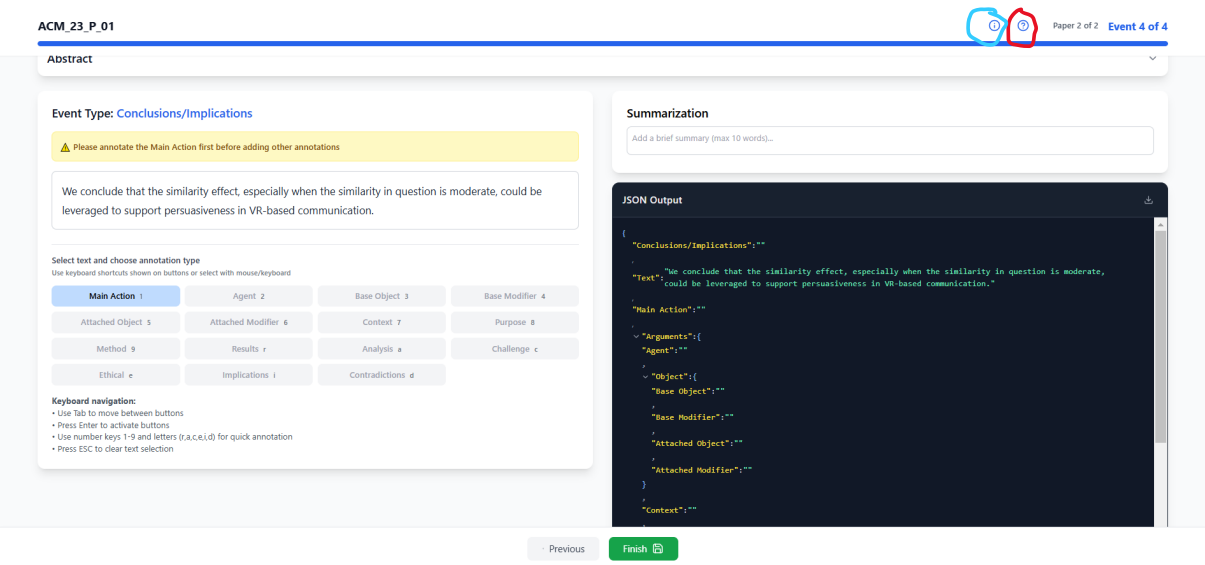
10. In total, you should annotate **Event Summarization** and **Arguments** (if explicit) for each **Event**.

11. Click on “Next” to annotate the next Event. In the last Event, the “Next” button will be replaced by “Finish” which automatically downloads the JSON annotation for you. This JSON file is what you need to submit to us. Alternatively, you can download with the download button to the top-right of the JSON Output.




12. Whenever you are unsure with annotation, please check section 3 and section 4 for clarification.

13. We highly recommend everyone watch the video demonstration in the annotation tool. (the question mark highlighted in red) Or if you want to access to this codebook, click on the “i” information button highlighted in blue.




# 3. Additional Notes

## 3.1 Annotate by breaking down sentences:

**Instruction:** Please annotate **segments of a sentence (a part of a sentence)** instead of a full sentence if **different segments of the sentence can be fit into different Arguments**. An example on this is give below:

**Example:**

However, this might result in a sub-optimal ToD system if it is tailored to only one ad hoc user simulator, since human users can behave differently.

**Annotation**:

<Challenge>: a sub-optimal ToD system

<Context>: tailored to only one ad hoc user simulator

<Analysis>: human users can behave differently

**Reasoning:**

❌ Instead of annotating full sentence as the argument <Challenge>

✅ Please breakdown the sentence and annotate the **segments of a sentence**.

## 3.2 Passive tense:

**Instruction:** In a passive tense structure: Something (Agent) + is done (Passive Verb) + by Someone/Something (Object)

**Example:** “The experiment was conducted by the researchers.”

**Annotation:**

<Agent>: The experiment

<Main Action>: was conducted
<Primary Object>: the researchers

***Please check the difference between the following similar examples***

**Example:** “Model performance can be improved by using a better prompting strategy.”
**Annotation:**
<Agent>: Model performance
<Main Action>: be improved
<Primary Object>: using a better prompting strategy

**Example:** “Model performance can be improved using a better prompting strategy.”
**Annotation:**
<Agent>: Model performance
<Main Action>: be improved
<Primary Object>: (not explicit, so leave this blank, refer to 3.5 for more information)
<Method>: using a better prompting strategy

## 3.3 Choosing proper Main Action:

**Instruction:** Use the Event Type when choosing the Main Action.
**Example:**
**Event:** background/introduction
**Text:** Addressing the issues of who saying what to whom in multi-party conversations (MPCs) has recently attracted a lot of research attention. However, existing methods on MPC understanding typically embed interlocutors and utterances into sequential information flows or utilize only the superficial of inherent graph structures in MPCs.”

**Annotation:**
<Main Action>: recently attracted

**Reasoning:**

✅ Why? It establishes the importance of the topic which is align with the purpose of background/introduction.

❌ Not choosing "embed " or "utilize" because they describe existing methods to show current <**Challenge**>, which doesn’t align with the purpose of background/introduction.

## 3.4 Primary and Secondary Objects:

**Instruction:** Recall definition and structure of <**Objects**> in section 1.3. This section discusses how you should annotate Primary Object and Secondary Object.

**Case 1: Parallel Objects**
When two objects are **parallel**, they follow this structure:
Main Action + Primary Object + as well as / and / also + Secondary Object
**Example:** Understanding the full spectrum of uptake factors is essential to identify ways in which policy makers and providers can facilitate the adoption of effective digital therapeutics within a health system, as well as the steps developers can take to assist in the deployment of products.
**Annotation:**
<Agent>: Understanding the full spectrum of uptake factors
<Main Action>: is essential to identify
<Primary Object>: ways
<Primary Modifier>: in which policy makers and providers can facilitate the adoption of effective digital therapeutics within a health system
<Secondary Object>: as well as the steps
<Secondary Modifier>: developers can take to assist in the deployment of products

**Reasoning:**
" ways " is the <Primary Object>: because it is the first and main thing being identified.
"as well as the steps " is <Secondary Object> because it is introduced later as an additional detail.

**Case 2: Transformation Objects**
When one <**object**> is mapped, translated, integrated, or moved into another, the structure is: Main Action + Primary Object + into / to + Secondary Object

✅ Primary Object → What is being transformed

✅ Secondary Object → Where it is placed or transformed into

**Example:** In this review, we aimed to map the most frequently discussed factors that determine the integration of digital therapeutics into health systems and practical use of digital therapeutics by patients and professionals.
**Annotation:**
<Agent>: we
<Main Action>: aimed to map
<Primary Object>: the most frequently discussed factors
<Primary Modifier>: that determine the integration of digital therapeutics

<Secondary Object>: into health systems and practical use
<Secondary Modifier>: of digital therapeutics by patients and professionals

**Reasoning:**

✅ The verb phrase "aimed to map" follows the above structure, making this a clear case of transformation objects.

✅ The <Primary Object> (“the most frequently discussed factors”) is being mapped into the <Secondary Object> (“into health systems and practical use”).

✅ Modifiers provide extra information of <**Objects**> but do not change the object relationships.

**Important:** Secondary Object should include the **proposition** (in, to, for, as well as, into, and, ….) before it, to indicate the relation between it and the Primary Object.

**Caution:** when you annotate the Objects, it requires reading and thinking rather than blindly following above structures:
**Example:** We modify the model parameter for a better performance.
<Agent>: We
<Main Action>: modify
<Primary Object>: the model parameter
<Primary Modifier>:
<Secondary Object>:
<Secondary Modifier>:
<Purpose>: a better performance
**Reasoning:**

❌ "better performance" is NOT a Secondary Object because:

1. It does not receive the <Main Action> (it is not being modified).
2. It is not a result of transformation.
3. It functions as a <Purpose> (explains why the modification happens).

## 3.5 Indirect Object:

**Instruction:** There is no **direct Object** sometimes, you should leave **Objects (all Objects and Modifiers)** empty:

**Case 1: verb in passive tense**
Recall 3.2 examples
**Example:** “Model performance can be improved by using a better prompting

strategy."
**Annotation:**
<Agent>: Model performance
<Main Action>: be improved
<Primary Object>: using a better prompting strategy

**Example:** "Model performance can be improved using a better prompting strategy."
**Annotation:**
<Agent>: Model performance
<Main Action>: be improved
<Primary Object>: (not explicit)
<Method>: using a better prompting strategy

**Reasoning:**

"By" → Identifies someone/something → ✅ <Primary Object> exists

"Using" → Describes how the action happens → ❌ No <Primary Object>

**Case 2: Verbs that don't need an object**
**Instruction:** <Agent> + <Main Action> + Preposition (by, in, through, …) + -ing Verb
**Example:**
We conclude by summarizing the key points.
**Annotation:**
<Agent>: We
<Main Action>: conclude
<Primary Object>:
<Primary Modifier>:
<Secondary Object>:
<Secondary Modifier>:
<Method>: by summarizing the key points

**Example:**
The system adapts through learning from interactions.
**Annotation:**
<Agent>: The system
<Main Action>: adapts
<Primary Object>:
<Primary Modifier>:
<Secondary Object>:
<Secondary Modifier>:
<Method>: through learning from interactions

## 3.6 Entire part of sentence as Object:

**Instruction:** In following structure, the entire part of sentence is <Primary Object>: <Agent> + <Main Action> + that / what / who / which / where / when / how / whether + part of sentence

**Example:**
Our experimental results on MultiWOZ show that the dialogue system trained by MUST achieves a better performance than those trained by a single user simulator.

**Annotation:**
<Agent>: Our experimental results on MultiWOZ
<Main Action>: show
<Primary Object>: that the dialogue system trained by MUST achieves a better performance than those trained by a single user simulator
<Primary Modifier>:
<Secondary Object>:
<Secondary Modifier>:

## 3.7 Verb Phrases:

**Definition:** A verb phrase is a **group of word** that **behave as a verb**.
**Instruction:** Annotate the **entire verb phrase**.
**Example:** It is important to validate the result.
**Annotation:**
<Agent>: It
<Main Action>: is important to validate
<Primary Object>: the result
<Primary Modifier>:
<Secondary Object>:
<Secondary Modifier>:

## 3.8 Text that fit multiple Arguments:

**Instruction:** If a text span can fit into many <Arguments>, follow:
**Order of importance**: Results > Purpose > Method > Analysis > Implication > Challenge > Contradiction > Context > Ethical.
Results is most important, and Ethical is the least.

**Examples:** Our result shows a difficult problem: Low-resource language models struggle with syntactic disambiguation.

**Annotation:**
<Agent>: Our result
<Main Action>: shows
<Primary Object>: a difficult problem
<Primary Modifier>:
<Secondary Object>:
<Secondary Modifier>:
<Results>: Low-resource language models struggle with syntactic disambiguation
<Challenge>:

**Reasoning: "**Low-resource language models struggle with syntactic disambiguation" fit both in <Results> and <Challenge>, but according to the order of importance, <Results> is more important than <Challenge>, and you shout annotate this text as <Results>.

### 3.9 Abbreviation

**Instruction:** You should use **both** original term and abbreviation
**Example**: Chain-of-Thought (CoT) improve the performance of LLMs.
**Annotation:**
<Agent>: Chain-of-Thought (CoT)
<Main Action>: improve
<Primary Object>: the performance
<Primary Modifier>: of LLMs
<Secondary Object>:
<Secondary Modifier>:

## 4. Event Arguments (Details)

Each **Arguments** (excluding the **Agent**, **Main Action** and **Object**) extracted should be categorized into one of the following 9 categories (4.1 to 4.9)

- **Underscored Indicator Words**: Indicator words are underscored for guiding annotation.
- **Highlighted Annotations**: Where to annotate is highlighted for reference.

### 4.1 Context

- **Definition**: Provides foundational or situational information.
- **Subarguments**:

- **Historical Context**
  **Definition:** Background from prior work or historical developments.
  **Rationale: Historical Context** provides foundational information by connecting the current study to prior developments, situating it within the broader **context** and framing its contributions and relevance**.**
  **Indicator word:** "builds on," "based on," "pioneered by," "founded on," "inspired by earlier work," "influenced by foundational methods."
  **Examples:**
  **(NLP)**: "Builds on Word2Vec by Mikolov et al. (2013), a foundational method for word representations in NLP, we propose our method."
  **(Todo: indicator word color, then also highlight what to annotate)**
  **(HCI)**: "This based on Fitts' Law (1954), a key model for predicting human movement in interaction tasks."
  **(Health Informatics):** "Our survey is designed by inspecting early electronic health records (EHRs) systems, pioneered in the 1960s."
  **(Digital Humanities)**: "This approach builds on early textual analysis methods developed in the mid-20th century, such as concordance creation."
  **(Computational Biology)**: "The study is founded on BLAST, a tool pioneered in 1990 for DNA sequence alignment."

- **Assumption:**
  **Definition:** Underlying ideas accepted as true without direct proof.
  **Rationale: Assumption** identifies pre-existing conditions or untested ideas that the study relies upon, providing critical **context** for understanding the scope and setup of the research.
  **Indicator Words:** "Assumes that", "presumes," "is based on the premise," "underlying assumption," "implicitly assumes," "accepted without validation."
  **Examples:**
  **(NLP)**: "The model assumes independence between tokens in sequence prediction tasks."
  **(HCI)**: "The study presumes that users have prior experience with touch-screen devices."
  **(Health Informatics)**: "The system relies on the assumption that patients accurately report symptoms in digital surveys."
  **(Digital Humanities)**: "This analysis presumes that the digitized manuscripts accurately represent the original texts."
  **(Computational Biology)**: "The algorithm is based on the premise that genetic variations in the sample set are representative of the larger population."

- **Condition**:
  **Definition**: Specific limits or situational factors present during the study that influence results and define where findings apply.
  **Rationale**: **Condition** clarifies the specific situational scope or constraints that define the study's results, ensuring the findings are interpreted within the appropriate **context**.
  **Indicator Words**: "Only applies to," "limited to," "restricted to," "valid for," "relevant for," "conducted under," "performed in," "with the following conditions," "in a controlled environment," "under standard conditions."
  **Examples**:
  **(NLP)**: "The evaluation was conducted on text data limited to publicly available news articles."
  **(HCI)**: "Findings are valid for desktop interfaces tested in environments with consistent lighting conditions."
  **(Health Informatics)**: "The study is restricted to patient records collected from urban hospitals over a six-month period."
  **(Digital Humanities)**: "The analysis was performed on digitized literary texts from the 19th century, under standard corpus linguistics procedures."
  **(Computational Biology)**: "The algorithm's performance is valid for genomic data sequenced using next-generation sequencing technologies."

- **Theoretical Framework**:
  **Definition**: Models underpinning the study.
  **Rationale**: **Theoretical Framework** links the study's methods or interpretations to established models or theories, ensuring that the **context** is known to readers.
  **Indicator Words**: "Based on," "grounded in," "applies the theory," "underpinned by," "draws on."
  **Examples**:
  **(NLP)**: "This work is based on the Transformer model, which uses self-attention for sequence processing."
  **(HCI)**: "The study applies Norman's theory of affordances to evaluate interface usability."
  **(Health Informatics)**: "The intervention is grounded in the Health Belief Model, which predicts behavior based on perceived risks and benefits."
  **(Digital Humanities)**: "The approach draws on Moretti's theory of distant reading to analyze large corpora of texts."
  **(Computational Biology)**: "This algorithm is underpinned by the

Neutral Theory of Molecular Evolution, which assumes genetic variations occur at constant rates over time."

## 4.2 Purpose

- **Definition**: Defines the purpose or aim.
- **Subarguments**:
    - **Goal/Aim:**
      **Definition**: The main purpose or objective of the study.
      **Rationale**: **Goal/Aim** describes the overarching intention or end result that the study seeks to achieve, providing clear **purposes** for the research by outlining what it aims to accomplish.
      **Indicator Words**: "Aims to," "seeks to," "goal is," "objective is," "intends to," "purpose of the study."
      **Examples**:
      **(NLP)**: "The study aims to improve sentiment classification for low-resource languages."
      **(HCI)**: "The goal is to enhance user engagement through adaptive interface designs."
      **(Health Informatics)**: "The aim is to evaluate the effectiveness of AI in early disease detection."
      **(Digital Humanities)**: "This research seeks to uncover patterns in 18th-century literary networks using computational methods."
      **(Computational Biology)**: "The objective is to predict protein structures more accurately using deep learning models."

    - **Hypothesis**:
      **Definition**: Statements proposed to explain a phenomenon, tested in the study.
      **Rationale**: **Hypothesis** specifies testable predictions that guide the study's methods, offering concrete **purposes** by explaining what phenomena the study investigates and seeks to validate.
      **Indicator Words**: "Proposes that," "tests whether," "is expected to," "hypothesizes that," "is assumed to result in," "predicts that."
      **Examples**:
      **(NLP)**: "Contextual embeddings are expected to improve entity recognition accuracy."
      **(HCI)**: "It is proposed that personalized feedback increases task completion rates."
      **(Health Informatics)**: "This study tests whether mobile health interventions reduce patient anxiety levels."
      **(Digital Humanities)**: "The hypothesis predicts that computational

stylistic analysis can distinguish between authors in collaborative texts."
**(Computational Biology)**: "It is hypothesized that integrating epigenetic data improves the accuracy of cancer subtype classification."

## 4.3 Method

- **Definition**: Techniques, tools, or frameworks used.
- **Subarguments**:
    - **Methodology**:
      **Definition**: General description of methods used.
      **Rationale**: **Methodology** outlines the overall approach or framework used in the study, providing a high-level description of the **methods** applied to achieve the study's goals.
      **Indicator Words**: "Was conducted using," "employed," "utilized," "was analyzed with," "applied methods include," "Updates based on," "refined iteratively," "adjusted using feedback," "loop includes."
      **Examples**:
      **(NLP)**: "Transformer-based models were fine-tuned on annotated datasets."
      **(HCI)**: "Usability testing was conducted with think-aloud protocols."
      **(Health Informatics)**: "Patient surveys were analyzed using mixed-methods approaches."
      **(Digital Humanities)**: "Text analysis was conducted using natural language processing techniques to extract recurring themes."
      **(Computational Biology)**: "Genome-wide association studies were conducted using statistical regression models."

    - **Subsequent Methodology**:
      **Definition**: Additional or follow-up techniques and procedures applied after the primary methodology.
      **Rationale**: **Subsequent methodologies** are critical for extending, validating, or refining initial findings. They establish a logical progression in the **method** and often bridge initial approaches with additional results or validation steps**.**
      **Indicator Words**: "Following this," "subsequent steps include," "in the next phase," "was further refined by," "extended through."
      **Examples**:
      **(NLP)**: "Following this, subsequent steps included fine-tuning the model on domain-specific datasets to improve accuracy further."
      **(HCI)**: "Subsequent methods involved iterative prototyping and user feedback collection to refine the interface."

**(Health Informatics)**: "In the next phase, machine learning models were retrained with additional features to enhance diagnostic precision."
**(Digital Humanities)**: "After the initial text analysis, subsequent steps included network analysis to map relationships between authors and works."
**(Computational Biology)**: "The study extended through additional gene expression analyses to validate the primary findings from sequence alignment."

- **Algorithm/Process Description**:
  **Definition**: Step-by-step technical details.
  **Rationale**: **Algorithm/Process Description** provides a detailed breakdown of the procedural steps or workflows involved, explaining how the **methods** were implemented technically..
  **Indicator Words**: "Preprocesses by," "includes steps such as," "consists of," "process involves," "workflow includes."
  **Examples**:
  **(NLP)**: "The system preprocesses text by tokenizing, normalizing, and embedding inputs before classification."
  **(HCI)**: "The interface design process included wireframing, prototyping, and iterative user testing."
  **(Health Informatics)**: "The prediction pipeline involves feature extraction, model training, and validation on patient datasets."
  **(Digital Humanities)**: "The workflow includes transcribing digitized manuscripts, cleaning OCR errors, and conducting semantic analysis."
  **(Computational Biology)**: "The process involves sequence alignment, motif discovery, and phylogenetic tree construction to understand evolutionary relationships."

- **Benchmarking**:
  **Definition**: Comparisons against standards or other methods.
  **Rationale**: **Benchmarking** evaluates the effectiveness of the study's methods by comparing results against established baselines or alternative approaches, demonstrating the **methods**' relative performance.
  **Indicator Words**: "Evaluated against," "compared to," "benchmark results include," "was measured relative to," "tested alongside."
  **Examples**:
  **(NLP)**: "The model was evaluated against BERT and GPT baselines on standard benchmarks."
  **(HCI)**: "User performance was compared to established metrics from prior usability studies."

**(Health Informatics)**: "Diagnostic accuracy was benchmarked against results from clinical experts."
**(Digital Humanities)**: "The proposed algorithm was compared to existing text-mining tools to assess its efficiency in processing historical corpora."
**(Computational Biology)**: "The new protein-folding model was tested alongside AlphaFold to evaluate predictive accuracy."

- **Implementation**:
  **Definition**: Practical realization of theories or tools.
  **Rationale**: **Implementation** describes how the theories, methods, or frameworks were practically realized, providing concrete details on the **methods**' real-world execution.
  **Indicator Words**: "Implemented using," "was developed as," "deployed with," "realized in," "translated into."
  **Examples**:
  **(NLP)**: "The model was implemented using PyTorch and trained on a distributed GPU cluster."
  **(HCI)**: "A prototype application was developed using React and deployed for user testing."
  **(Health Informatics)**: "The intervention was implemented as a mobile app integrated with existing EHR systems."
  **(Digital Humanities)**: "The text analysis tool was developed as a Python package and deployed on Jupyter Notebooks for interactive use."
  **(Computational Biology)**: "The genome annotation pipeline was implemented using Snakemake and deployed on a high-performance computing cluster."

- **Optimization**:
  **Definition**: Refinements for better outcomes.
  **Rationale**: **Optimization** focuses on adjustments or refinements made to improve the efficiency, accuracy, or performance of the **methods** used in the study.
  **Indicator Words**: "Optimized by," "tuned using," "refined to," "adjusted for," "improved to maximize."
  **Examples**:
  **(NLP)**: "Hyperparameters were tuned using grid search to maximize classification accuracy."
  **(HCI)**: "Interaction flows were streamlined to reduce task completion time."
  **(Health Informatics)**: "The algorithm was optimized for faster processing of large-scale patient datasets."

**(Digital Humanities)**: "The text-mining model was refined to handle OCR errors, improving the analysis of digitized historical documents."
**(Computational Biology)**: "The protein-structure prediction pipeline was adjusted for better runtime performance on high-throughput datasets."

- **Mechanism**:
  **Definition**: The step-by-step process explaining how something works.
  **Rationale**: **Mechanism** provides a detailed breakdown of how methods or processes function, offering a systematic **methods** of why specific approaches produce the observed outcomes.
  **Indicator Words**: "Works by," "operates through," "step-by-step process," "mechanism involves," "functioning of."
  **Examples**:
  **(NLP)**: "The attention mechanism assigns weights to tokens, prioritizing more relevant context for predictions."
  **(HCI)**: "The recommendation system adjusts displayed content based on real-time user interactions."
  **(Health Informatics)**: "The wearable device measures heart rate variability to detect early signs of stress."
  **(Digital Humanities)**: "The mechanism involves parsing historical texts, tagging named entities, and mapping relationships to create a knowledge graph."
  **(Computational Biology)**: "The algorithm works by aligning DNA sequences, identifying conserved motifs, and predicting functional domains."

- **Validation**:
  **Definition**: Checking to make sure the results are correct and reliable.
  **Rationale**: **Validation** ensures the correctness and reliability of the research by using verification **methods** to confirm their accuracy and robustness.
  **Indicator Words**: "Validated through," "confirmed by," "cross-checked with," "verified using."
  **Examples**:
  **(NLP)**: "The model's predictions were validated using cross-validation on three datasets."
  **(HCI)**: "Usability findings were confirmed through follow-up user testing sessions."
  **(Health Informatics)**: "The algorithm's outputs were validated against expert-annotated clinical data."
  **(Digital Humanities)**: "The text classification system was validated through manual review by domain experts."

**(Computational Biology)**: "Predicted protein structures were verified using experimentally resolved crystal structures from the PDB database."

- **Metrics**:
  **Definition:** Measures used to evaluate the performance or quality of a system or method.
  **Rationale: Metrics** show what **method** is used to measure the results by providing measurable criteria that assess the performance, quality, or effectiveness of the study's methods.
  **Indicator Words:** "Measured by," "evaluated using," "performance was assessed with," "scores indicate."
  **Examples:**
  **(NLP):** "The model's performance was evaluated using accuracy, F1 score, and BLEU for translation tasks."
  **(HCI):** "User satisfaction was measured through SUS (System Usability Scale) scores and task completion time."
  **(Health Informatics):** "The system's effectiveness was assessed using sensitivity, specificity, and patient adherence rates."
  **(Digital Humanities)**: "The system's text-mining accuracy was evaluated using precision, recall, and F1 scores."
  **(Computational Biology)**: "The effectiveness of the genome analysis pipeline was assessed using ROC-AUC and false discovery rates."

## 4.4 Results

- **Definition**: Observations or outputs.
- **Subarguments**:
  - **Observation**:
    **Definition**: Important findings or patterns noticed during the study.
    **Rationale**: **Observation** identifies key patterns or findings that emerge from the study's analysis, providing direct **results** that highlight what was discovered.
    **Indicator Words**: "Observed that," "noticed a trend where," "found that," "results showed."
    **Examples**:
    **(NLP)**: "The model struggled with rare words, showing lower accuracy for low-frequency tokens."
    **(HCI)**: "Users preferred interfaces with fewer menu options, reporting higher satisfaction scores."
    **(Health Informatics)**: "We found that patients responded more quickly to reminders sent in the morning compared to the evening."
    **(Digital Humanities)**: "The analysis showed a trend where 18th-

century novels contained increasingly complex social networks over time."
**(Computational Biology)**: "We observed that specific gene clusters were consistently expressed under high-stress conditions in bacterial samples."

- **Evidence/Support**:
  **Definition**: Data or results used to back up conclusions.
  **Rationale**: **Evidence/Support** provides concrete data that substantiates the study's findings, presenting measurable **results** to validate conclusions.
  **Indicator Words**: "Achieved," "supported by," "data shows," "results confirm."
  **Examples**:
  **(NLP)**: "The model achieved 92% accuracy on the test set, outperforming all baselines."
  **(HCI)**: "80% of participants completed the task faster using the new interface design."
  **(Health Informatics)**: "Survey data showed a 30% reduction in reported symptoms after the intervention."
  **(Digital Humanities)**: "Topic modeling results supported by manual annotations revealed recurring themes across historical documents."
  **(Computational Biology)**: "Gene expression analysis confirmed a 40% increase in specific proteins under experimental conditions."

- **Classification/Categorization**:
  **Definition**: Organizing data into specific groups or labels.
  **Rationale**: **Classification/Categorization** organizes the study's findings into meaningful groups, providing structured **results** that simplify analysis and interpretation.
  **Indicator Words**: "Organized into," "grouped by," "categorized as," "labeled with."
  **Examples**:
  **(NLP)**: "Text data was categorized into sentiment labels: positive, negative, or neutral."
  **(HCI)**: "Users were grouped based on their interaction styles: novice, intermediate, or expert."
  **(Health Informatics)**: "Patients were classified by risk level: low, moderate, or high."
  **(Digital Humanities)**: "Historical texts were categorized by genre: fiction, biography, and political essays."
  **(Computational Biology)**: "Samples were grouped into clusters based

on gene expression patterns under different experimental conditions."

- **Novelty**:
  **Definition**: New and original contributions or discoveries.
  **Rationale**: **Novelty** highlights unique findings or innovations introduced by the study, presenting **results** that demonstrate its originality and contribution to the field.
  **Indicator Words**: "Introduces," "presents a new," "first to," "novel contribution."
  **Examples**:
  **(NLP)**: "This study introduces a new algorithm for real-time translation of code-mixed languages."
  **(HCI)**: "We present a novel interface that adapts dynamically based on user emotions."
  **(Health Informatics)**: "Our work is the first to integrate wearable device data into predictive health models."
  **(Digital Humanities)**: "This research presents a novel computational approach to trace the evolution of literary styles over centuries."
  **(Computational Biology)**: "This study is the first to use multi-omics data integration to predict cancer patient outcomes with high accuracy."

- **Comparison**:
  **Definition**: Comparing of findings or approaches.
  **Rationale**: **Comparison** evaluates the study's **results** relative to other methods or baselines, highlighting its performance, advantages, or differences.
  **Indicator Words**: "Outperforms," "compared to," "similar to," "results indicate better than."
  **Examples**:
  **(NLP)**: "Our model outperforms GPT-3 in accuracy but requires less computational resources."
  **(HCI)**: "This design resulted in faster task completion compared to traditional menu-based interfaces."
  **(Health Informatics)**: "The proposed algorithm shows similar precision to expert diagnoses but operates significantly faster."
  **(Digital Humanities)**: "Our method processes historical corpora more efficiently compared to existing text-mining tools."
  **(Computational Biology)**: "This pipeline performs better than existing models in predicting gene-disease associations with lower false-positive rates."

- **Scalability**:
  **Definition**: The ability of a system to handle larger or more complex

tasks effectively.
**Rationale**: **Scalability** evaluates how well the **results** hold up as the system is scaled to handle larger datasets, tasks, or environments.
**Indicator Words**: "Scales to," "adapts seamlessly," "handles large-scale," "maintains performance at scale."
**Examples**:
**(NLP)**: "The architecture scales efficiently to datasets with billions of sentences."
**(HCI)**: "The interface adapts seamlessly to multi-device environments, including tablets and smartphones."
**(Health Informatics)**: "The platform supports integration with healthcare systems managing millions of patient records."
**(Digital Humanities)**: "The system processes large-scale digitized archives, maintaining performance across collections spanning multiple languages."
**(Computational Biology)**: "The algorithm scales to genomic datasets containing millions of sequences without significant loss of speed or accuracy."

- **Causation**:
  **Definition**: Direct relationships where one thing causes another.
  **Rationale**: **Causation** identifies cause-effect relationships observed in the study, showing how changes in one variable directly influence **results**.
  **Indicator Words**: "Results show that," "caused by," "led to," "directly influenced by."
  **Examples**:
  **(NLP)**: "Increasing training data size directly improves the model's accuracy."
  **(HCI)**: "Reducing menu complexity led to faster task completion times."
  **(Health Informatics)**: "Timely medication reminders significantly increased patient adherence rates."
  **(Digital Humanities)**: "Providing historical context for documents led to more accurate classification in text analysis tasks."
  **(Computational Biology)**: "Mutations in this gene directly caused a loss of function in the associated protein under experimental conditions."

- **Correlation**:
  **Definition**: Relationships observed between variables, without proving one causes the other.
  **Rationale**: **Correlation** highlights relationships or associations observed in the **results**, without establishing direct causality.

**Indicator Words**: "Associated with," "linked to," "correlates with," "relationship between."
**Examples**:
**(NLP)**: "Word frequency correlates with the model's confidence in its predictions."
**(HCI)**: "Longer session times are associated with higher user satisfaction ratings."
**(Health Informatics)**: "Higher social media activity is linked to increased patient engagement with health interventions."
**(Digital Humanities)**: "The frequency of religious motifs in texts is associated with the political climate of the era."
**(Computational Biology)**: "Gene expression levels are correlated with changes in environmental conditions such as temperature."

- **Disagreement**:
  **Definition**: Unexpected differences between observed and expected results.
  **Rationale**: **Disagreement** identifies inconsistencies between predictions and actual **Results**.
  **Indicator Words**: "Contrary to expectations," "unexpectedly," "differences between," "observed versus expected," "discrepancies in."
  **Examples**:
  **(NLP)**: "The model performed worse on shorter sentences than anticipated."
  **(HCI)**: "Contrary to expectations, users preferred the non-adaptive interface in time-sensitive tasks."
  **(Health Informatics)**: "Predicted recovery rates were significantly lower than clinical observations."
  **(Digital Humanities)**: "Unexpectedly, thematic patterns in the dataset did not align with established historical narratives."
  **(Computational Biology)**: "Contrary to expectations, certain gene knockouts showed no effect on cell growth under stress conditions."

## 4.5 Analysis

- **Definition**: Interpretation or explanation of other arguments.
- **Subarguments**:
  - **Explanation**:
    **Definition**: Reasons given to clarify why something happens.
    **Rationale**: **Explanation** provides interpretations of the study's observations by clarifying why certain methods or phenomena occur, contributing to the overall **analysis** of the study's choices and findings.
    **Indicator Words**: "Because," "due to," "as a result of," "this happens

when," "explained by."
**Examples**:
**(NLP)**: "The model struggles with rare words because the training data lacks sufficient examples."
**(HCI)**: "Users prefer larger buttons due to improved visibility and easier interaction."
**(Health Informatics)**: "Patients responded more quickly to shorter reminders because they require less cognitive effort."
**(Digital Humanities)**: "The prevalence of certain phrases in 19th-century novels is explained by the influence of serialized publication formats."
**(Computational Biology)**: "The algorithm misclassified some genes due to insufficient training data for rare functional categories."

- **Examples (as a Subargument)**:
  **Definition**: Specific instances or scenarios used to illustrate findings, phenomena, or explanations.
  **Rationale**: **Examples** help contextualize and clarify abstract findings or explanations, making them more tangible and relatable. They enhance the **analysis** by providing real-world relevance or detailed case scenarios.
  **Indicator Words**: "For instance," "such as," "an example is," "illustrated by," "demonstrated through."
  **Examples**:
  **(NLP)**: "An example of this is the model's performance improvement on domain-specific text, such as medical records."
  **(HCI)**: "Survey shows button size has an impact of people's interaction satisfaction, for instance, older adults demonstrated higher interaction satisfaction when larger buttons were introduced."
  **(Health Informatics)**: "This can be seen in cases where early intervention prevented adverse health outcomes, such as reduced hospital readmissions."
  **(Digital Humanities)**: "An example is the application of computational text analysis to identify recurring themes in Renaissance literature."
  **(Computational Biology)**: "For instance, gene mutation patterns were demonstrated through comparative analysis of cancer and normal tissues."

- **Elaborations**:
  **Definition**: Detailed expansions or additional explanations that clarify findings, methods, or theoretical insights.
  **Rationale**: **Elaborations** provide deeper insights or further interpret findings and observations, enhancing understanding. They enrich the

**analysis** by breaking down complex concepts into more digestible details or exploring implications in depth.
**Indicator Words**: "This means that," "in more detail," "can be expanded as," "to elaborate further".
**Examples**:
**(NLP)**: "This means that increasing the training corpus size directly affects performance for low-resource languages."
**(HCI)**: "To elaborate further, user satisfaction improves when interfaces are tailored to individual preferences."
**(Health Informatics)**: "This can be expanded as a multi-stage process where wearable devices contribute both real-time and historical data for predictions."
**(Digital Humanities)**: "In more detail, thematic patterns in 19th-century novels reveal evolving societal norms reflected in character dialogues."
**(Computational Biology)**: "To elaborate further, integrating transcriptomic and proteomic data enables a comprehensive understanding of gene regulation networks."

## 4.6 Challenge

- **Definition**: Constraints or weaknesses.
- **Subarguments**:
    - **Limitation**:
      **Definition**: Factors that may affect the reliability or validity of the results.
      **Rationale**: **Limitation** highlights constraints that influence the reliability or applicability of the study's findings, making it a critical part of the overall **Challenge** by identifying weaknesses that frame the study's scope.
      **Indicator Words**: "However", "Limited by," "constrained by," "restricted to," "affected by," "scope includes."
      **Examples**:
      **(NLP)**: "The model's performance is limited by the lack of diverse training data."
      **(HCI)**: "Findings are constrained by the use of a controlled lab environment, which may not reflect real-world conditions."
      **(Health Informatics)**: "The analysis is limited to data collected from a single healthcare provider."
      **(Digital Humanities)**: "The study is constrained by the limited availability of digitized archives for non-Western literary traditions."
      **(Computational Biology)**: "The algorithm's effectiveness is restricted to datasets with high-quality sequence annotations."

- **Error/Artifact**:
  **Definition**: Issues or anomalies that may affect the accuracy or interpretation of results.
  **Rationale**: **Error/Artifact** addresses inaccuracies or irregularities in the data or methods, contributing to the **Challenge** by identifying potential distortions that impact result interpretation.
  **Indicator Words**: "May include errors," "affected by artifacts," "issues in," "noise in data," "anomalies from."
  **Examples**:
  **(NLP)**: "The model occasionally misclassifies due to noise in the training data."
  **(HCI)**: "User testing results may include artifacts from inconsistent internet speeds during sessions."
  **(Health Informatics)**: "Measurement errors in wearable devices could impact the accuracy of health predictions."
  **(Digital Humanities)**: "Text analysis outputs may be affected by OCR errors in digitized historical manuscripts."
  **(Computational Biology)**: "Sequencing errors in raw genomic data could introduce artifacts into mutation detection algorithms."

## 4.7 Ethical

- **Definition**: Ethical concerns, implications, and justifications.
- **Subarguments**:
  - **Ethical Risks**:
    **Definition**: Potential moral or societal risks arising from the study or system.
    **Rationale**: **Ethical Risks** identify specific negative ethical implications, such as bias, fairness, privacy violations, or societal harm, which need to be addressed as part of **Ethical Considerations**.
    **Indicator Words**: "Raises concerns about," "risks of," "ethical implications," "must ensure," "could perpetuate," "potential harm."
    **Examples**:
    **(NLP)**: "Deploying biased language models risks reinforcing societal stereotypes and marginalizing vulnerable groups."
    **(HCI)**: "Persuasive interface designs may exploit users' cognitive biases without their informed consent."
    **(Health Informatics)**: "Opaque AI-driven diagnoses can undermine patient trust and lead to inequitable healthcare outcomes."
    **(Digital Humanities)**: "Automated text analysis tools could misrepresent historical narratives, perpetuating biased interpretations of cultural texts."

**(Computational Biology)**: "The use of genomic data without explicit consent risks violating privacy and could disproportionately impact marginalized populations."

- **Ethical Justification**:
  **Definition**: Reasons given to support ethical choices in a study or system.
  **Rationale**: **Ethical Justification** explains the steps taken to mitigate ethical risks and ensure the study aligns with accepted moral principles, contributing to the broader **Ethical Considerations** argument.
  **Indicator Words**: "Ensures privacy by," "justified by," "prioritizes fairness," "adheres to ethical guidelines," "promotes inclusivity," "balances utility and ethics."
  **Examples**:
  **(NLP)**: "Training data was anonymized to protect user privacy while ensuring model utility."
  **(HCI)**: "The system prioritizes accessibility features to promote inclusivity for users with disabilities."
  **(Health Informatics)**: "Patient consent was obtained before using data to ensure adherence to ethical guidelines."
  **(Digital Humanities)**: "The digitization process adheres to copyright laws and respects the cultural significance of historical artifacts."
  **(Computational Biology)**: "Genomic data usage was justified by obtaining participant consent and adhering to international bioethics standards."

- **Stakeholders/Beneficiaries**:
  **Definition**: Groups or individuals affected by or benefiting from the study or system.
  **Rationale**: **Stakeholders/Beneficiaries** identify the groups impacted by the study or system, ensuring that **Ethical Considerations** account for fairness in distributing benefits and acknowledging societal impacts.
  **Indicator Words**: "Benefits include," "stakeholders impacted are," "beneficiaries gain from," "affects groups such as," "end-users benefit from."
  **Examples**:
  **(NLP)**: "Developers of multilingual applications benefit from the improved translation model."
  **(HCI)**: "End-users of e-learning platforms gain from the enhanced interface usability."
  **(Health Informatics)**: "Patients with chronic conditions benefit from real-time monitoring and predictive alerts."

**(Digital Humanities)**: "Historians and educators benefit from automated analysis of large text corpora, enabling new insights into historical narratives."
**(Computational Biology)**: "Researchers in precision medicine gain from improved gene-disease association predictions, advancing personalized treatment strategies."

## 4.8 Implications

- **Definition**: Broader applicability or significance and future research suggestion.
- **Subarguments**:
    - **Implication**:
      **Definition**: The broader significance or potential impact of the findings.
      **Rationale**: **Implication** highlights the broader applicability or real-world impact of the study's findings, demonstrating the relevance and significance of the research within the **Implications** argument.
      **Indicator Words**: "Suggests that," "could improve," "may lead to," "potential impact includes," "significant for," "enhances."
      **Examples**:
      **(NLP)**: "Improved sentiment analysis can enhance content moderation on social media platforms."
      **(HCI)**: "Findings suggest that adaptive interfaces could significantly improve productivity in workplace applications."
      **(Health Informatics)**: "Real-time health monitoring may reduce hospital readmission rates through early intervention."
      **(Digital Humanities)**: "Automated analysis of historical texts could broaden access to previously understudied literary works, enhancing cultural understanding."
      **(Computational Biology)**: "Better protein-structure predictions may lead to advancements in drug discovery and personalized medicine."

    - **Risk/Uncertainty**:
      **Definition**: Potential challenges or unknown factors that could affect outcomes.
      **Rationale**: **Risk/Uncertainty** identifies potential limitations or unknowns regarding the applicability or success of the study's findings, ensuring that the **Implications** account for real-world challenges and variability.
      **Indicator Words**: "May fail when," "uncertainty about," "risks include," "unknown factors," "could be affected by," "introduces challenges."
      **Examples**:
      **(NLP)**: "The model may fail when applied to languages not represented

in the training data."
**(HCI)**: "There is uncertainty about how users will adapt to the new interface in high-stress environments."
**(Health Informatics)**: "Reliance on wearable devices introduces the risk of inaccurate health data due to hardware malfunctions."
**(Digital Humanities)**: "The automated analysis could be affected by inaccuracies in OCR processing of historical texts."
**(Computational Biology)**: "The algorithm's predictions may fail when applied to novel species with incomplete genomic data."

- **Future Work/Questions**:
  **Definition**: Suggested directions for further research or development.
  **Rationale**: **Future Work/Questions** proposes extensions, refinements, or new research directions, showing how the study's findings can inspire continued exploration under the **Implications** argument.
  **Indicator Words**: "Future work includes," "further studies are needed," "could explore," "suggested directions include," "future research could."
  **Examples**:
  **(NLP)**: "Future work could explore adapting the model for low-resource languages."
  **(HCI)**: "Further studies are needed to evaluate interface performance in real-world settings."
  **(Health Informatics)**: "Future research could investigate the integration of genomic data into predictive health models."
  **(Digital Humanities)**: "Future studies could explore applying these methods to analyze underrepresented cultural archives."
  **(Computational Biology)**: "Future research could include expanding the algorithm to incorporate epigenomic data for more precise predictions."

## 4.9 Contradictions (to current Knowledge)

- **Definition**: Disagreements challenging existing knowledge.
- **Subarguments**:
  - **Contradiction**:
    **Definition**: Findings that conflict with or challenge prior research.
    **Rationale**: **Contradiction** highlights results that directly challenge or deviate from existing knowledge, contributing to the **Contradictions to Knowledge** argument by questioning established assumptions or findings.
    **Indicator Words**: "Unlike prior studies," "contrary to previous research," "our findings challenge," "in conflict with," "disagrees with."

**Examples**:
**(NLP)**: "Unlike prior studies, our results show that pretraining on domain-specific data is not always beneficial."
**(HCI)**: "Contrary to previous research, we found that users preferred static menus over adaptive ones."
**(Health Informatics)**: "Our findings contradict earlier claims that wearable devices always improve patient adherence."
**(Digital Humanities)**: "Our analysis challenges previous interpretations by showing that key literary themes emerged earlier than previously thought."
**(Computational Biology)**: "Contrary to prior research, the study found that certain gene mutations do not correlate with increased disease risk in specific populations."

- **Alternative/Counterargument**:
  **Definition**: Differing or opposing perspectives on an issue.
  **Rationale**: **Alternative/Counterargument** introduces opposing viewpoints or perspectives that challenge the study's claims, enhancing the **Contradictions to Knowledge** argument by presenting alternative explanations or frameworks.
  **Indicator Words**: "An alternative view suggests," "some argue that," "a counter perspective is," "opposing viewpoints include," "it has been proposed that."
  **Examples**:
  **(NLP)**: "An alternative view suggests that rule-based systems are more interpretable than neural models for certain tasks."
  **(HCI)**: "Some argue that traditional desktop interfaces remain more efficient than mobile-first designs for complex workflows."
  **(Health Informatics)**: "A counterargument is that patient-reported data might be less reliable than clinical observations."
  **(Digital Humanities)**: "It has been proposed that manual annotation of historical texts provides more nuanced interpretations than computational methods."
  **(Computational Biology)**: "A counter perspective is that single-cell data might not fully capture population-level dynamics in gene expression studies."

- **Rebuttal**:
  **Definition**: Responses addressing critiques or opposing views.
  **Rationale**: **Rebuttal** provides responses or justifications to counter opposing views, strengthening the **Contradictions to Knowledge** argument by defending the study's claims against alternative perspectives.

**Indicator Words**: "While it is true that," "although some argue," "despite concerns".
**Examples**:
**(NLP)**: "While rule-based systems are interpretable, they lack the flexibility and scalability of neural models."
**(HCI)**: "Although desktop interfaces may be efficient, mobile-first designs better meet the needs of on-the-go users."
**(Health Informatics)**: "Despite concerns about reliability, patient-reported data provides valuable insights when combined with clinical records."
**(Digital Humanities)**: "Although manual annotation provides depth, computational methods enable analysis at a scale that would be impractical otherwise."
**(Computational Biology)**: "While population-level studies capture broad trends, single-cell data provides finer granularity essential for understanding cellular heterogeneity."