

Simple Italian and AI: strengths and weaknesses

Giuliana Fiorentino, Vittorio Ganfi, **Marco Russodivito**,
Alessandro Cioffi, Maria Ausilia Simonelli

University of Molise, Italy



Introduction

Introduction: the PRIN project

A_Preparation: Italst and s-Italst corpora, ATS, selecting the metrics

B_ Experimental phase 1: comparing manual (human) and automatic (AI) simplification. Outcomes.

C_ Experimental phase 2: Validation of AI simplification with 2 datasets (juridical evaluation and text comprehension). Outcomes.

D_ Experimental phase 3: Validation of AI simplification with a new dataset and new prompts. Results

Conclusions and perspectives

VerbACxSS: on verb analysis and complexity, synthesis and simplification, for accessibility

(June 2022 – May 2025)

Website

<https://verbacxss.it>



Italst Corpus

<https://huggingface.co>



Research project

	Rifiuti		Sanità		Servizi pubblici		
	Carta dei servizi	Bandi di gara	Atti generali di prog.	Accreditamento	Carta dei servizi	Razion. partecip. pub.	
Basilicata	3	5	1	2	4	5	20
Calabria	5	6	1	4	4	5	25
Campania	5	9	1	6	4	5	30
Lazio	4	5	1	2	4	5	21
Lombardia	10	5	1	2	4	7	29
Molise	4	6	1	6	4	5	26
Toscana	5	12	1	3	4	8	33
Veneto	6	3	1	4	4	6	24
93			37		78		208
42			29		46		

RQ1:

To what extent can AI simplify text compared to a human?

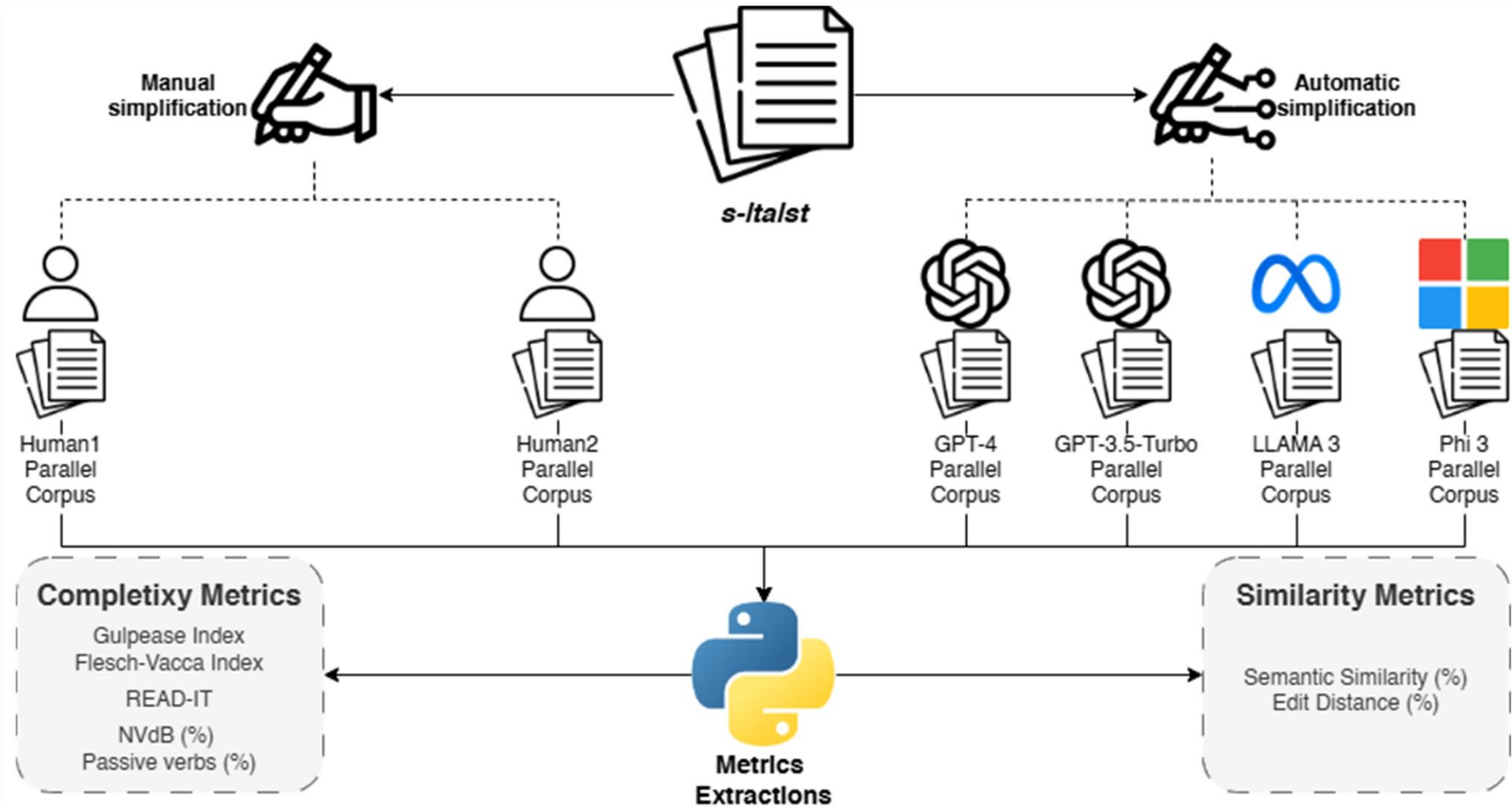
RQ2:

How much time does a human require to perform text simplification?

RQ3:

What simplification rules are applied by humans?

Experimental phase 1 - RQs



Experimental Phase 1 - Pipeline

Sei un dipendente pubblico che deve riscrivere dei documenti istituzionali italiani per renderli **semplici e comprensibili per i cittadini**. Ti verrà fornito un documento pubblico e il tuo compito sarà quello di riscriverlo applicando regole di semplificazione **senza** però **modificare il significato** del documento originale. Ad esempio potresti rendere le **frasi** più **brevi**, **eliminare le perifrasi**, **esplicitare** sempre il **soggetto**, utilizzare **parole** più **semplici**, trasformare i verbi **passivi** in verbi di forma **attiva**, spostare le **frasi parentetiche** alla **fine** del periodo.

GPT-4 Prompt

You are a public servant tasked with rewriting Italian institutional documents to make them **simple and understandable for citizens**. You will be provided with a public document, and your task will be to rewrite it by applying simplification rules **without altering** the original **meaning**. For example, you might **shorten sentences, eliminate periphrases**, always make the **subject explicit**, use **simpler words**, convert **passive** verbs into **active** ones, and move **parenthetical phrases** to the **end** of the sentence.

GPT-4 Prompt (translated)

	<i>Original</i>	<i>Human1</i>	<i>Human2</i>	<i>GPT-3.5-Turbo</i>	<i>GPT-4</i>	<i>LLaMA 3</i>	<i>Phi 3</i>
<i>Tokens</i>	33,295	34,135	29,755	30,032	31,722	36,035	36,056
<i>Sentences</i>	1,314	1,506	1,744	1,515	1,840	1,944	1,900
<i>Tokens per Sentences</i>	25.33	22.66	17.06	19.53	17.24	18.53	18.97
<i>Sentences per Documents</i>	164.25	188.25	218.00	189.37	230.00	243.00	237.50
<i>Gulpease Index</i>	44.31	49.72	50.64	48.49	51.34	50.26	50.16
<i>Flesch Vacca Index</i>	19.97	34.23	33.63	30.33	36.75	34.09	33.75
<i>NVdB (%)</i>	73.28	80.44	76.89	78.28	81.07	80.18	80.16
<i>Passive (%)</i>	20.87	15.78	17.71	13.99	12.00	15.81	15.72
<i>READ-IT BASE (%)</i>	75.91	68.62	51.00	66.61	55.00	58.37	57.69
<i>READ-IT LEXICAL (%)</i>	93.64	85.37	89.71	91.96	90.29	77.13	75.74
<i>READ-IT SYNTACTIC (%)</i>	63.72	53.14	40.09	38.42	29.92	40.97	41.24
<i>READ-IT GLOBAL (%)</i>	86.48	69.24	61.34	68.69	54.60	59.26	58.37
<i>Semantic Similarity (%)</i>	-	96.52	97.26	96.06	95.80	94.96	94.96
<i>Edit distance (%)</i>	-	35.84	29.20	49.21	52.14	55.48	55.44

Parameters

GPT-4 simplifications can be comparable to human simplifications.

GPT-4 simplifications are negligibly better for complexity metrics, moderately worse for semantic similarity, and largely rephrased (Edit distance) compared to human simplifications.

Experimental Phase 1 - Outcome

Paragrafo 26

Original

La trasparenza che costituisce uno degli obiettivi permanenti del Corpo di Polizia Municipale, definisce e rende note le modalità operative adottate, i tempi, nonché i criteri che portano all'erogazione del servizio nel suo complesso, rivedendolo periodicamente in un'ottica di miglioramento continuo.

Reviewer1 (251.0s)

La trasparenza è uno degli obiettivi permanenti del Corpo di Polizia Municipale, definisce e rende note il modo di operare adottato, i tempi e i criteri con cui si offre il servizio nel suo complesso, rivedendolo periodicamente per migliorare sempre.

Reviewer2 (3.0s)

La trasparenza è uno degli obiettivi del Corpo di Polizia Municipale. La trasparenza chiarisce i tempi e i modi di erogazione del servizio nel suo complesso. Il servizio viene periodicamente modificato per migliorarlo di continuo.

GPT-4

Il Corpo di Polizia Municipale mira sempre alla trasparenza. Questo significa che rende noti i metodi di lavoro, i tempi e i criteri usati per fornire i suoi servizi. Inoltre, controlla e migliora questi aspetti regolarmente.

Words:23 Token:43 Chars:301
Frase:1 Punteggiatura:5
Nomi:12 Avverbi:1 Pronomi:3 Articoli:6 Aggettivi:8
Verbi:5 Verbi attivi:5 Verbi passivi:0

VdB: 75.51 %
Gulpease: 32.7 Flesch Reading Ease: 4.48
Flesch Kincaid Grade: 22.8 Automated Readability Index: 29.1

Words:20 Token:40 Chars:250
Frase:1 Punteggiatura:4
Nomi:9 Avverbi:2 Pronomi:3 Articoli:6 Aggettivi:3
Verbi:8 Verbi attivi:8 Verbi passivi:0

VdB: 88.64 %
Gulpease: 43.8 Flesch Reading Ease: 22.42
Flesch Kincaid Grade: 20.1 Automated Readability Index: 23.4

similarity:0.97 edit_distance:83
added_tokens:13 added_vdb_tokens:13
deleted_tokens:16 deleted_not_vdb_tokens:7

Words:21 Token:35 Chars:230
Frase:3 Punteggiatura:3
Nomi:12 Avverbi:1 Pronomi:1 Articoli:6 Aggettivi:1
Verbi:4 Verbi attivi:3 Verbi passivi:1

VdB: 87.5 %
Gulpease: 58.7 Flesch Reading Ease: 51.14
Flesch Kincaid Grade: 9.0 Automated Readability Index: 10.8

similarity:0.93 edit_distance:136
added_tokens:8 added_vdb_tokens:7
deleted_tokens:20 deleted_not_vdb_tokens:9

Words:20 Token:38 Chars:225
Frase:3 Punteggiatura:5
Nomi:9 Avverbi:3 Pronomi:1 Articoli:7 Aggettivi:2
Verbi:7 Verbi attivi:7 Verbi passivi:0

VdB: 88.49 %
Gulpease: 61.2 Flesch Reading Ease: 50.84
Flesch Kincaid Grade: 9.2 Automated Readability Index: 9.4

similarity:0.89 edit_distance:199
added_tokens:20 added_vdb_tokens:17
deleted_tokens:29 deleted_not_vdb_tokens:14

Paragrato 41

Original

Fatte salve le eventuali conseguenze di legge, possono essere esclusi dal trasporto ed allontanati, senza diritto ad alcun risarcimento, i passeggeri molesti e quelle persone che non prestassero ascolto alle osservazioni ed agli inviti del personale di servizio tendenti a far rispettare le presenti norme.

Reviewer1 (213.0s)

I cittadini che sull'autobus sono pericolosi, dannosi o fastidiosi per gli altri, e che non obbediscono alle richieste degli autisti o dei controllori (che fanno rispettare le norme) possono essere mandati via dall'autobus e non hanno diritto a riavere i soldi del biglietto. Oltre a questo i cittadini pericolosi e dannosi possono anche avere altre conseguenze previste dalla legge.

Reviewer2 (140.0s)

Ricordate le eventuali conseguenze di legge, i passeggeri molesti e quelle persone che non prestassero ascolto alle osservazioni ed agli inviti del personale di servizio tendenti a far rispettare le presenti norme possono essere esclusi dal trasporto ed allontanati, senza diritto ad alcun risarcimento, .

GPT-4

I passeggeri che disturbano gli altri o non ascoltano le indicazioni del personale possono essere esclusi dal trasporto e allontanati. Non avranno diritto a ricevere risarcimenti. Questo è valido salvo diverse disposizioni di legge.

Words:20 Token:45 Chars:306
Frase:1 Punteggiatura:4
Nomi:13 Avverbi:1 Pronomi:1 Articoli:5 Aggettivi:5
Verbi:6 Verbi attivi:5 Verbi passivi:1

VdB: 85.71 %
Gulpease: 37.4
Flesch Kincaid Grade: 22.0
Flesch Reading Ease: 17.34
Automated Readability Index: 28.5

Words:22 Token:61 Chars:383
Frase:2 Punteggiatura:6
Nomi:13 Avverbi:5 Pronomi:5 Articoli:5 Aggettivi:6
Verbi:9 Verbi attivi:8 Verbi passivi:1

VdB: 86.76 %
Gulpease: 44.1
Flesch Kincaid Grade: 16.0
Flesch Reading Ease: 33.07
Automated Readability Index: 19.3

similarity:0.78 edit_distance:273
added_tokens:37 added_vdb_tokens:22
deleted_tokens:29 deleted_not_vdb_tokens:9

Words:20 Token:44 Chars:305
Frase:1 Punteggiatura:4
Nomi:13 Avverbi:1 Pronomi:1 Articoli:5 Aggettivi:4
Verbi:6 Verbi attivi:5 Verbi passivi:1

VdB: 85.42 %
Gulpease: 36.5
Flesch Kincaid Grade: 21.6
Flesch Reading Ease: 18.36
Automated Readability Index: 28.5

similarity:0.97 edit_distance:190
added_tokens:1 added_vdb_tokens:1
deleted_tokens:2 deleted_not_vdb_tokens:0

Words:21 Token:34 Chars:232
Frase:3 Punteggiatura:3
Nomi:8 Avverbi:2 Pronomi:3 Articoli:3 Aggettivi:3
Verbi:7 Verbi attivi:6 Verbi passivi:1

VdB: 88.89 %
Gulpease: 56.9
Flesch Kincaid Grade: 10.1
Flesch Reading Ease: 43.09
Automated Readability Index: 11.8

similarity:0.89 edit_distance:212
added_tokens:15 added_vdb_tokens:12
deleted_tokens:25 deleted_not_vdb_tokens:8

Testo 6 Campania

Limitations:

- The documents are much simpler
- The style was deeply changed and less similar to original document
- Possible loss of juridical value
- Documents are juridically equivalent but with some details removed

Experimental Phase 1 - Open problems

RQ1:

How does automatic simplification meet the legal quality of texts?

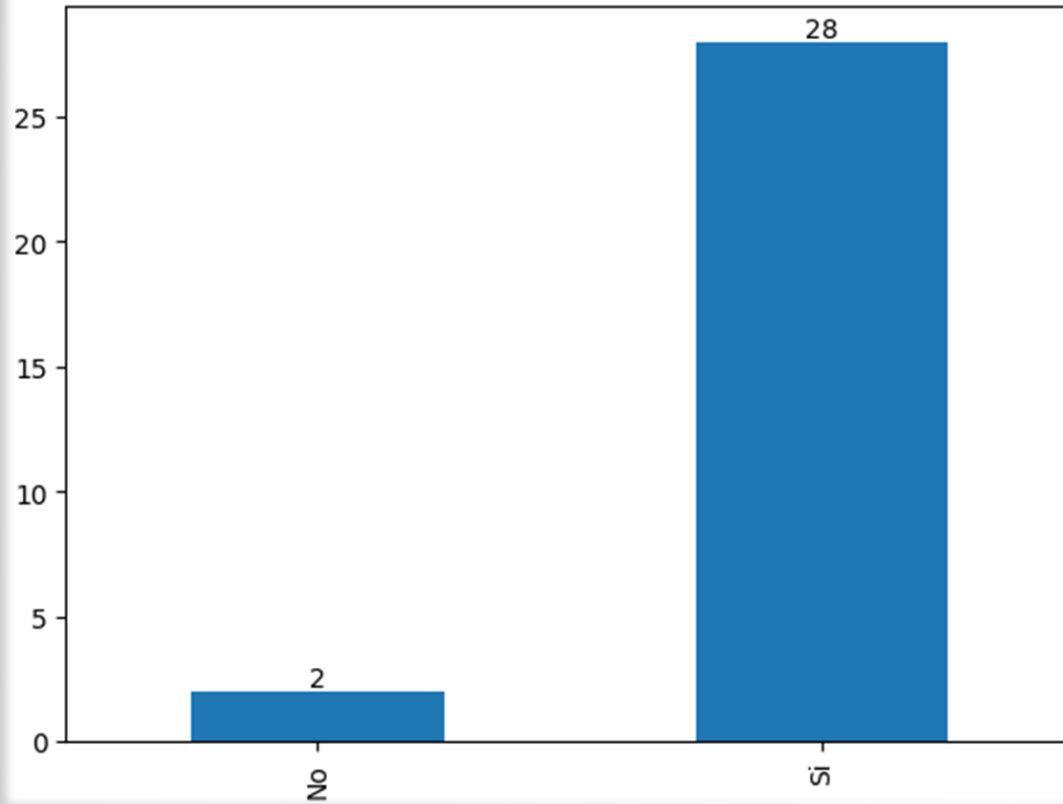
RQ2:

How does automatic simplification improve text comprehension for various categories of readers/users?

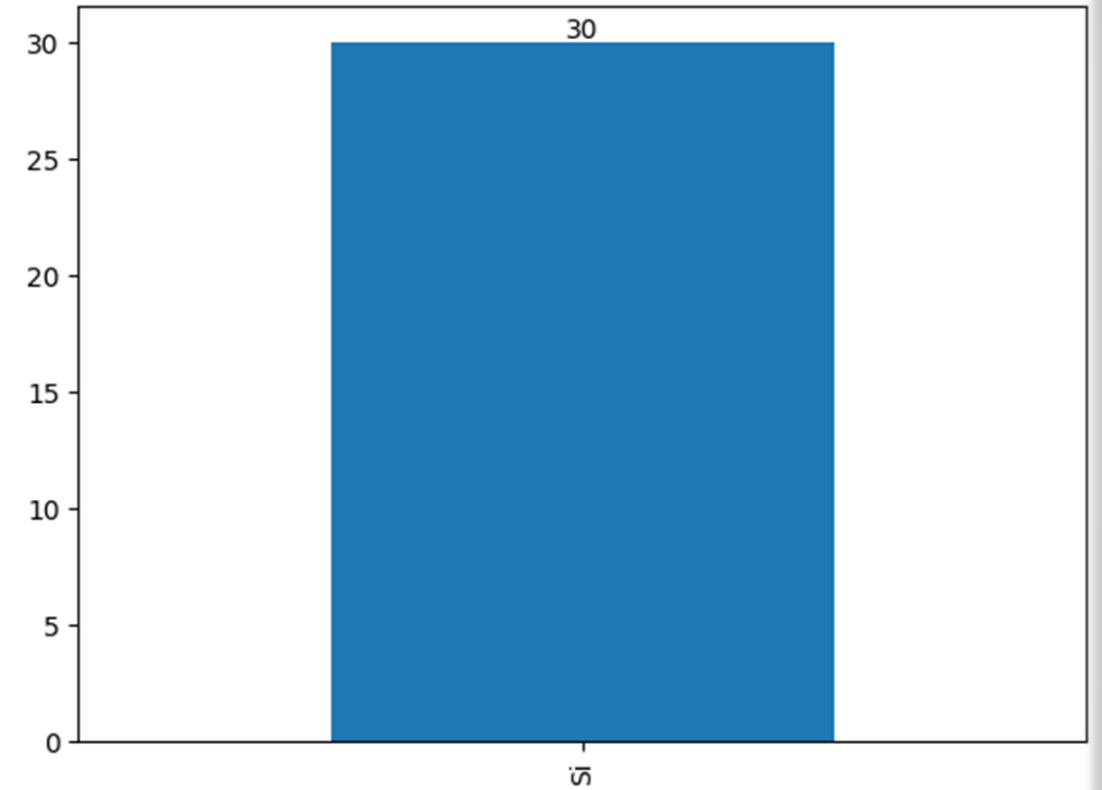
Experimental Phase 2 - RQs

Q1 - Sono presenti tutte le informazioni essenziali?

Biased (1 - 30)



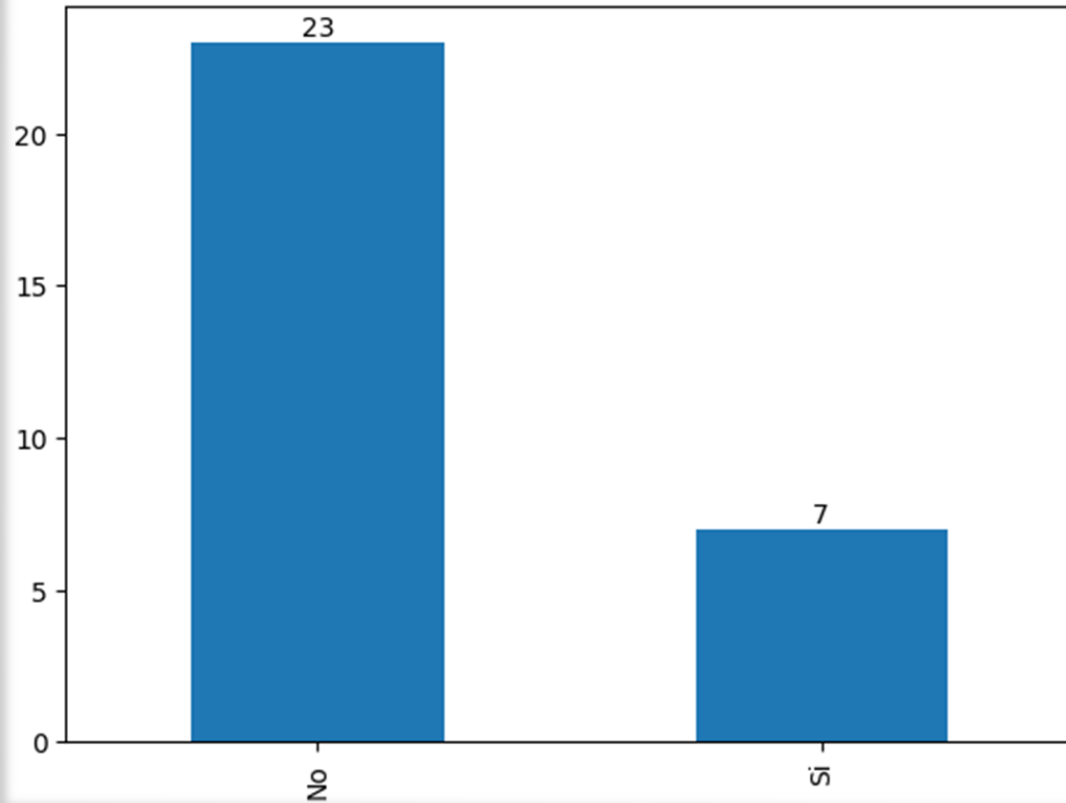
Unbiased (31 - 60)



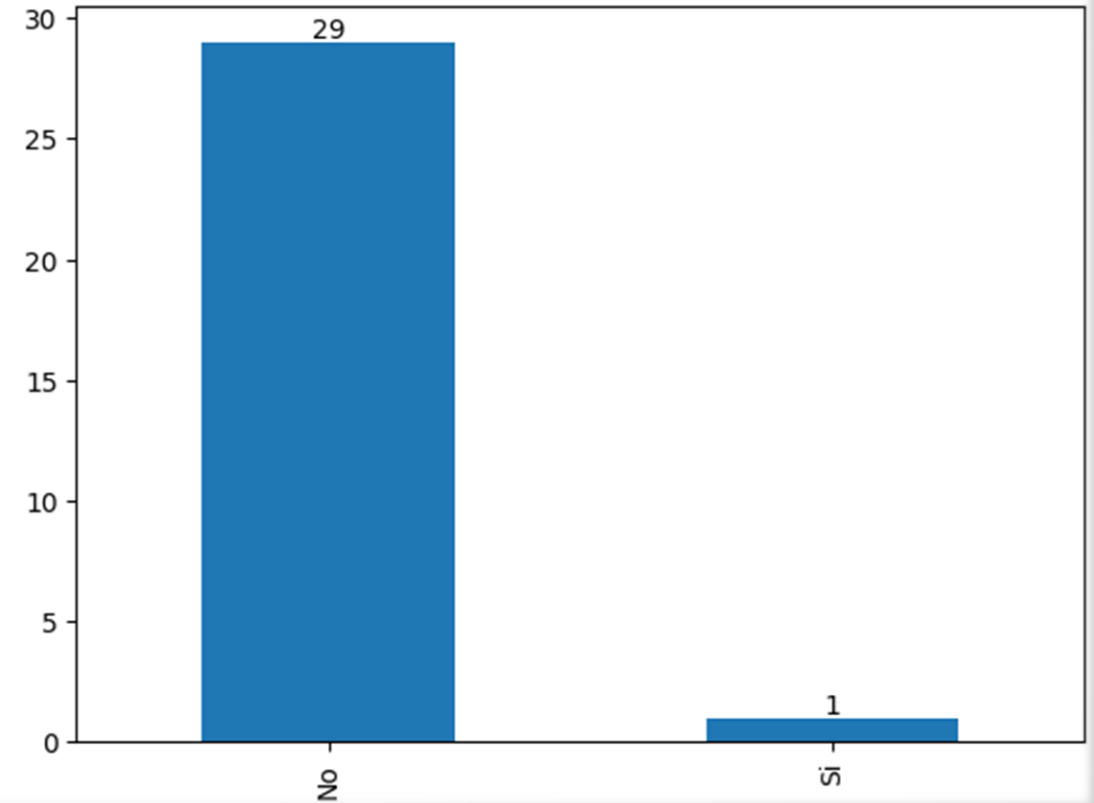
Experimental Phase 2

Q3 - Sono presenti parole (tecnicismi) con effetto giuridico che sono state cancellate creando problematicità al testo semplificato?

Biased (1 - 30)

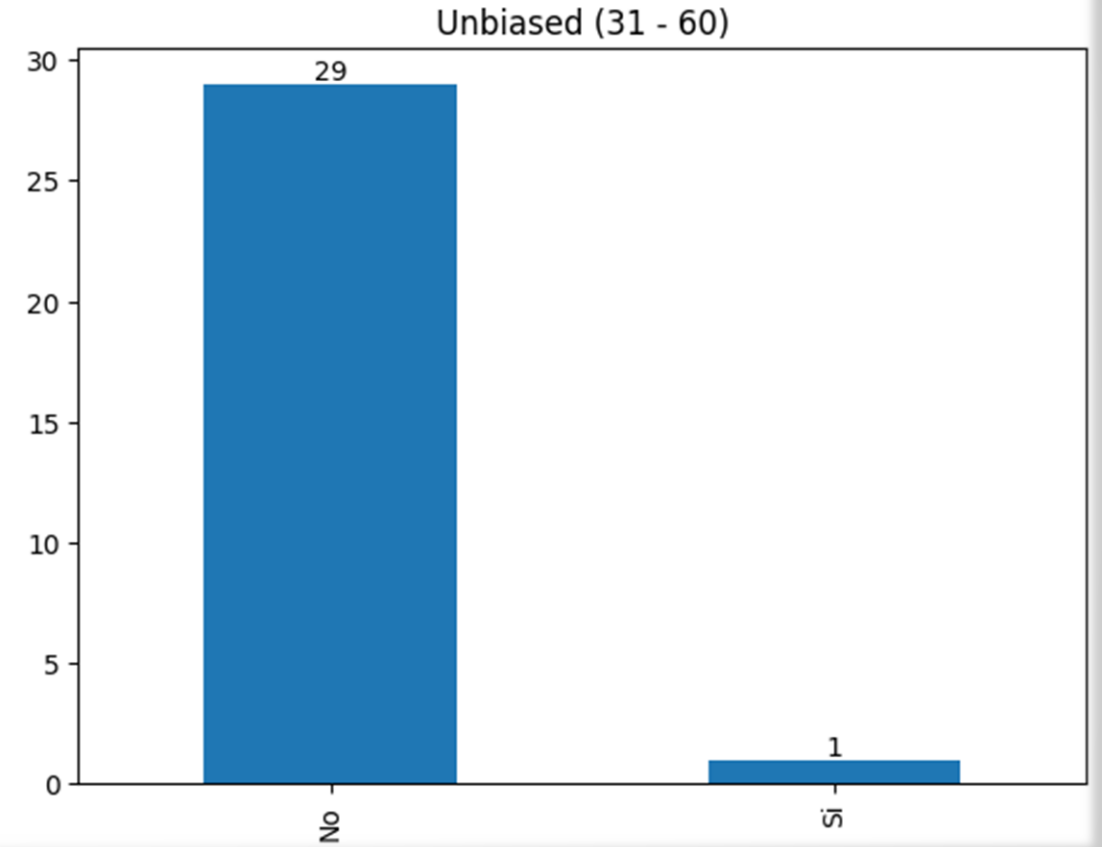
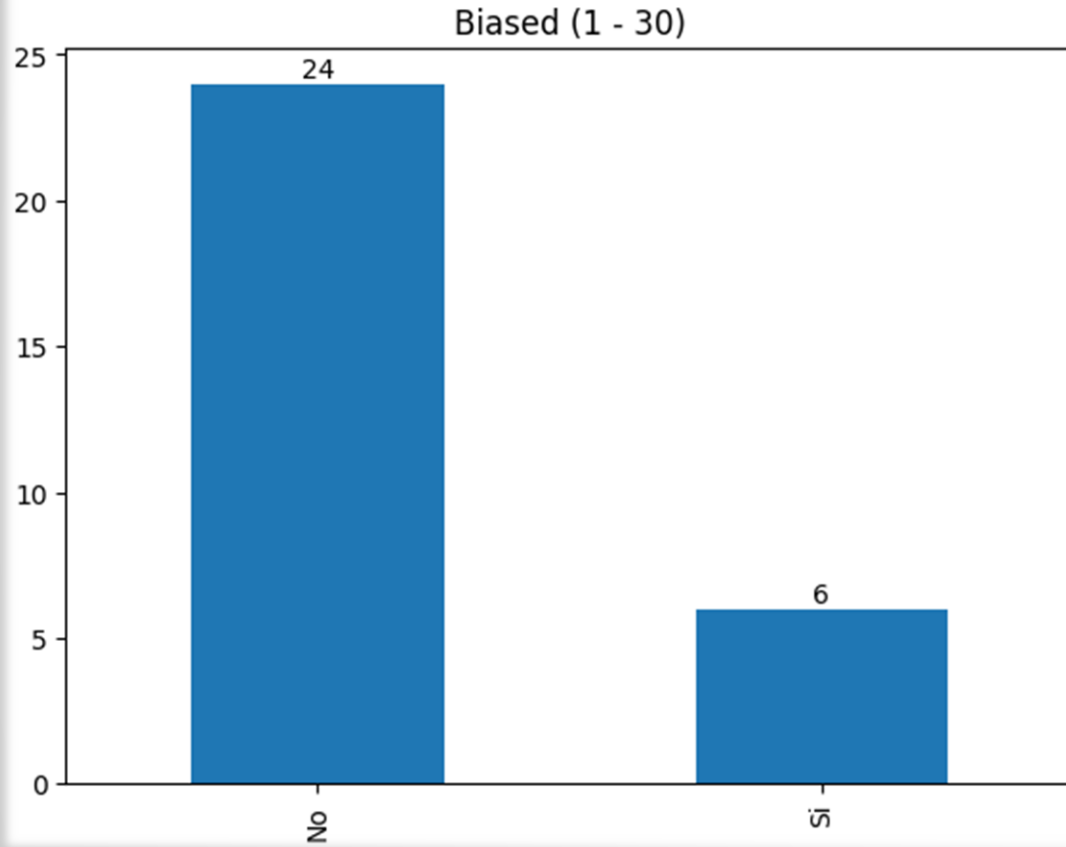


Unbiased (31 - 60)



Experimental Phase 2

Q4 - Sono stati introdotti errori interpretativi?



Experimental Phase 2

1. Select administrative documents
2. Select a short excerpt from the document
3. Simplify the selected document
4. Formulate text comprehension questions
5. Administer the test



Experimental phase 2: Comprehension assessment

 The reading and test execution times are consistent with the participant's experience

- There is no clear improvement in the correctness of the comprehension test



 There is an improvement in the quantitative assessment of the text's difficulty

Experimental phase 2: RQ2 outcome

RQ 1: automatic simplification meets the legal quality of the texts in a highly acceptable way with few reservations regarding the lexicon.
POSSIBLE CORRECTIVE: GLOSSARIES

RQ 2: based on the comprehension tests, automatic simplification does not seem to drastically improve the readers / users' understanding of the texts.

Future Prospects:

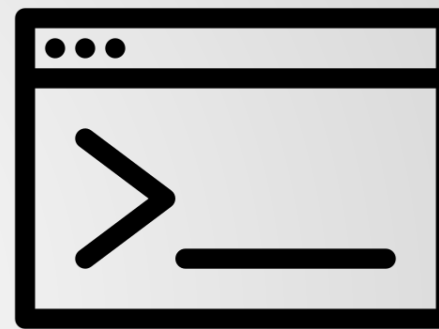
- perform other comprehension tests using texts in which the metrics of the simplified texts are significantly better than those of the original texts,
- modify the prompt to obtain a more significant simplification
- work on the 'factors' (e.g. lexical) that impact understanding.

Experimental phase 2: Outcome

**Preserve the juridical meaning of
the simplified administrative
documents**



**OpenAI
LLMs**



**Prompting
techniques**

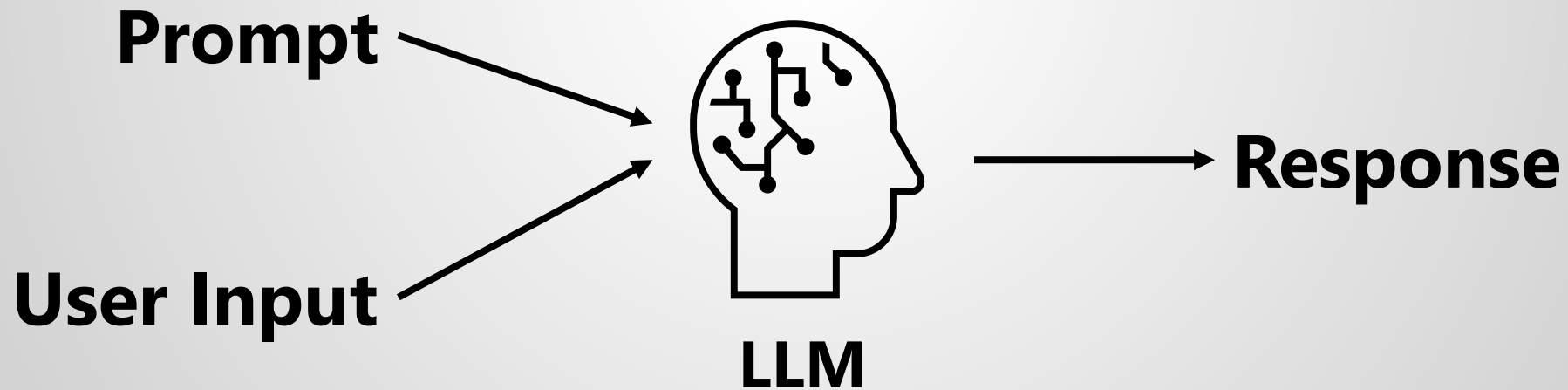
Proposal

LLMs & Prompting

LLMs are models trained on vast amounts of text data to understand and generate human-like language.

When you interact with LLMs you need to write a **prompt** to instruct the model to perform the desired task.

Prompt engineering is the process of crafting an effective prompt to get an accurate response from an LLM.



Prompting engineering

System prompt:

You are an expert **government employee**.

Your task is to **rewrite** an Italian institutional document to make it **simple** and **understandable** for citizens.

You can apply the following rules:

- **Shorter sentences**
- Use **simple words**
- Transform **passive** verbs **to active**
- Makes the **subject explicit**
- ...

ROLE

TASK

CONTEXT &
INSTRUCTIONS



User:

[Original document to be simplified]

LLM:

[Simplified document]

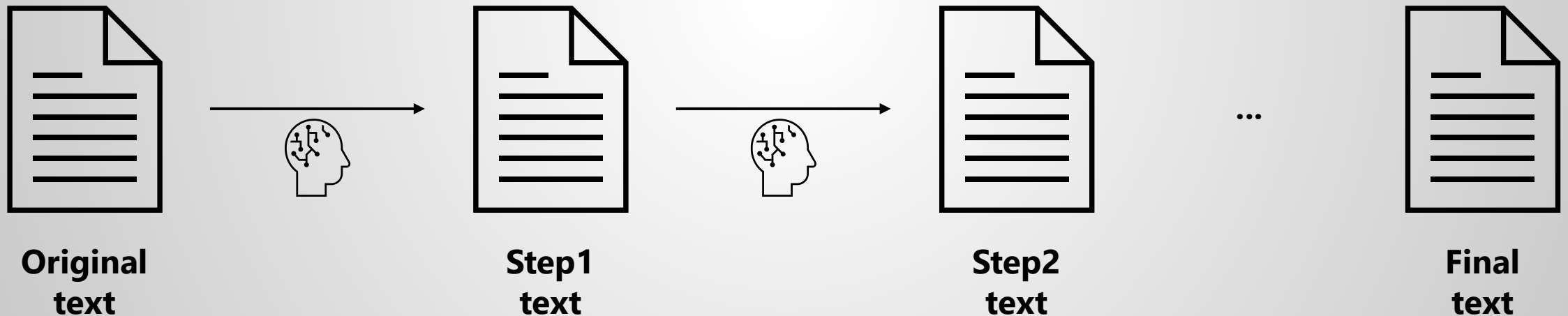


BASIC Prompt

The simplification of an administrative document is a **complex task**.

To preserve juridical meaning and to maintain control over the simplification process, it is possible to exploit the “**divide et impera**” principle

Prompt Chaining ensures more accurate, clear, and **progressively simplified** outputs.



Prompt chaining - idea

Apply state-of-the-art guidelines proposed over the years.

- 1) Proofreading
- 2) Replace bureaucratic and formal connectives
- 3) Replace bureaucratic and formal expressions
- 4) Divide long and complex sentences
- 5) Reorganize sentences (SVO, explicit subject)
- 6) Verbs (passive to active, indefinite to finite form)
- 7) Replace nominalizations**

Prompt chaining – Simplification step



Original



Step 1
Proofreading



Step 2
Connectives

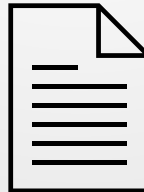
...



Step 6
Verbs



Step 7
Nominalizations



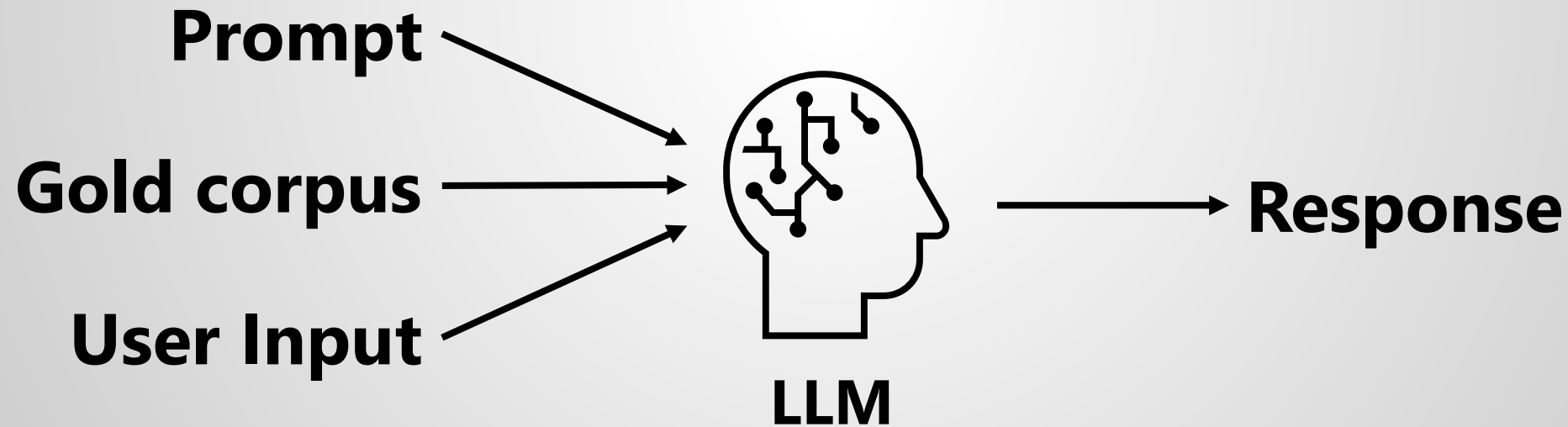
Simplified

Prompt chaining – Simplification step

For each simplification step, we wrote a **detailed prompt** to instruct the LLM to perform only that specific operation.

We manually simplified a small **corpus of 8 paragraph** (from an administrative corpus) applying all simplification steps.

The **gold corpus** is a parallel corpus that can be used to improve the accuracy of each simplifications step.



Prompt chaining – Gold Corpus

System prompt:

You are an expert **government employee**.

Your task is to [...]

Follow these instructions:

- ...

Examples:

- ...

ROLE

TASK

CONTEXT &
INSTRUCTIONSFEW-SHOT
PROMPTING**User:**

[Document to simplify]

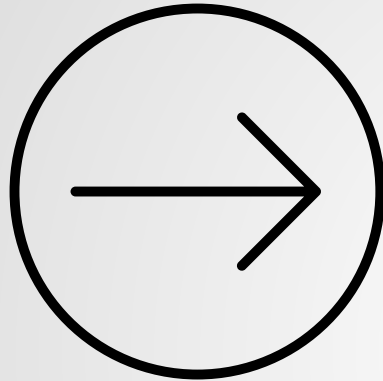
LLM:

[Document simplified]

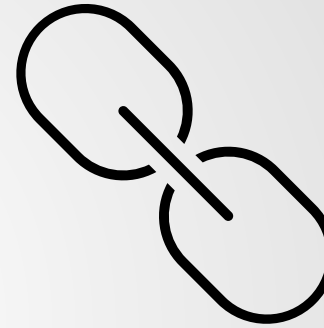


CHAIN prompt

**Evaluate proposed
methodology**

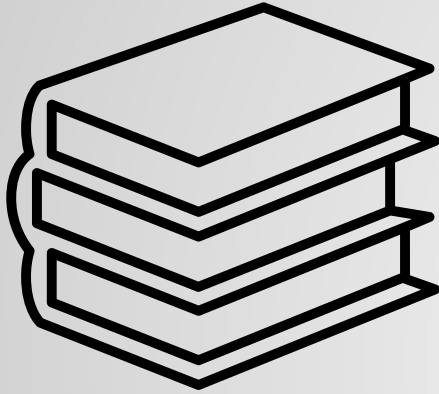


BASIC
prompting

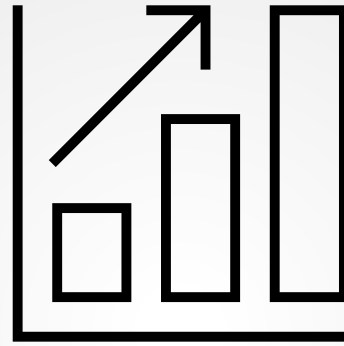


CHAIN
prompting

Compare BASIC versus CHAIN approach



**Regulations
Corpus**



**Metrics-based
evaluation**



**Jurist-based
evaluation**

Empirical study

RQ1

How effective is the CHAIN prompting in simplifying administrative documents compared to the BASIC prompting?

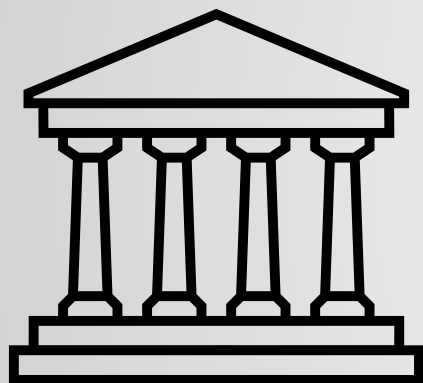
RQ2

How effective is the CHAIN prompting in maintaining juridical meaning compared to the BASIC prompting?

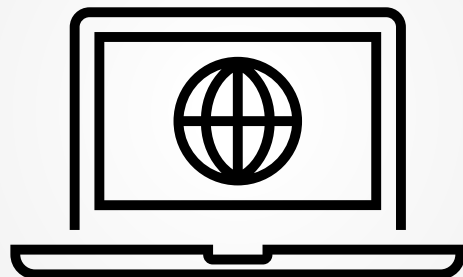
Research Questions

ItaRegol Corpus

The **ItaRegol** corpus was collected crawling regulations documents from several public administration **institutional website**. Each document was analyzed and **validated** by a **team of jurists**.



**Public
administration**



**Institutional
website**

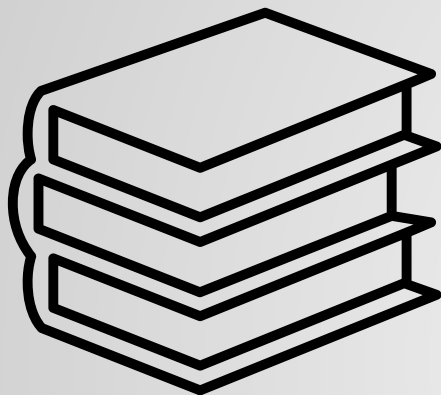


**Jurist
supervision**

ItaRegol Corpus

- Regulations of **garbage tax**
- Regulations of **urban green**
- Regulations of the **air quality**
- Regulations of **sport facilities**
- Regulations of **nursery schools**
- Regulations for **urban building**
- Regulations of **animal protection**
- Regulations for cargo and bus **driving license**





8
regulations

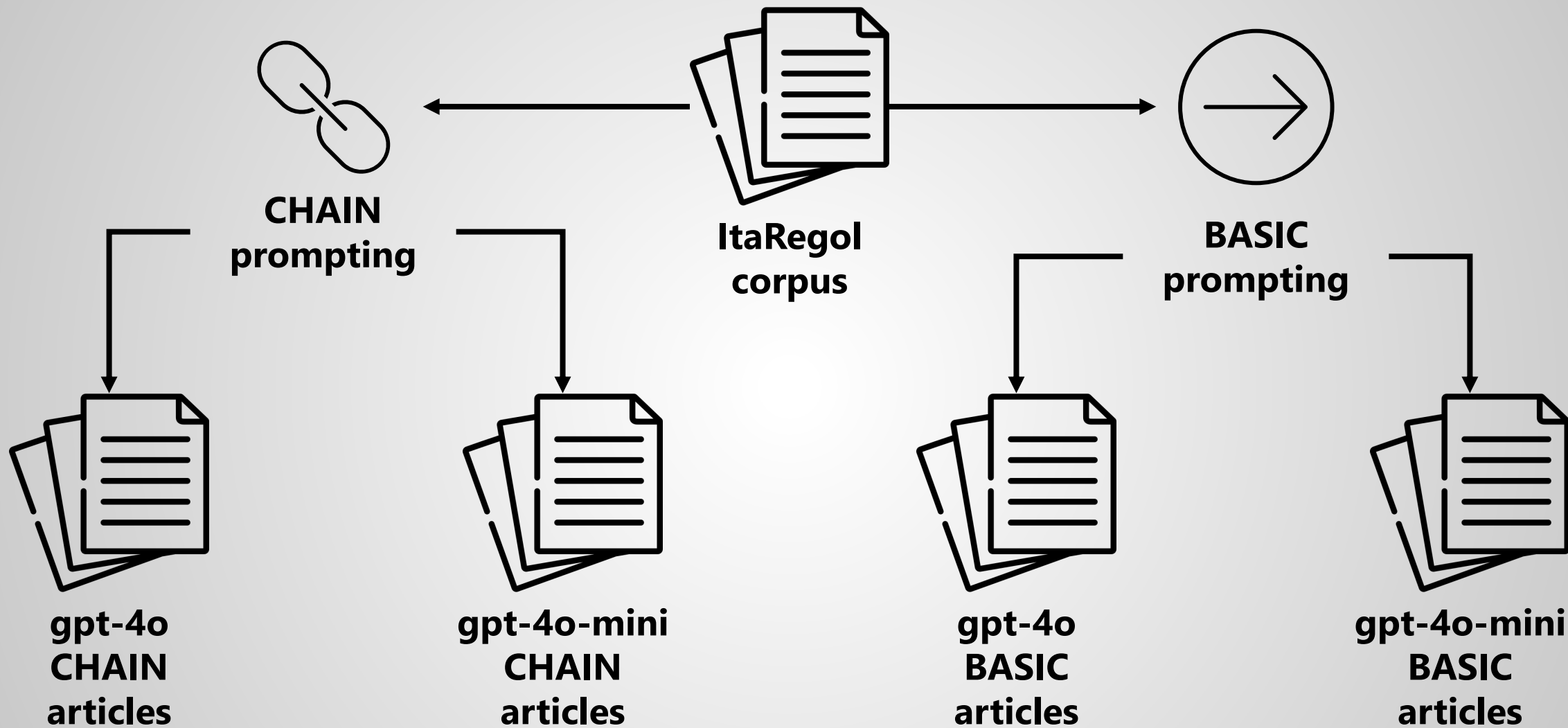


53
articles

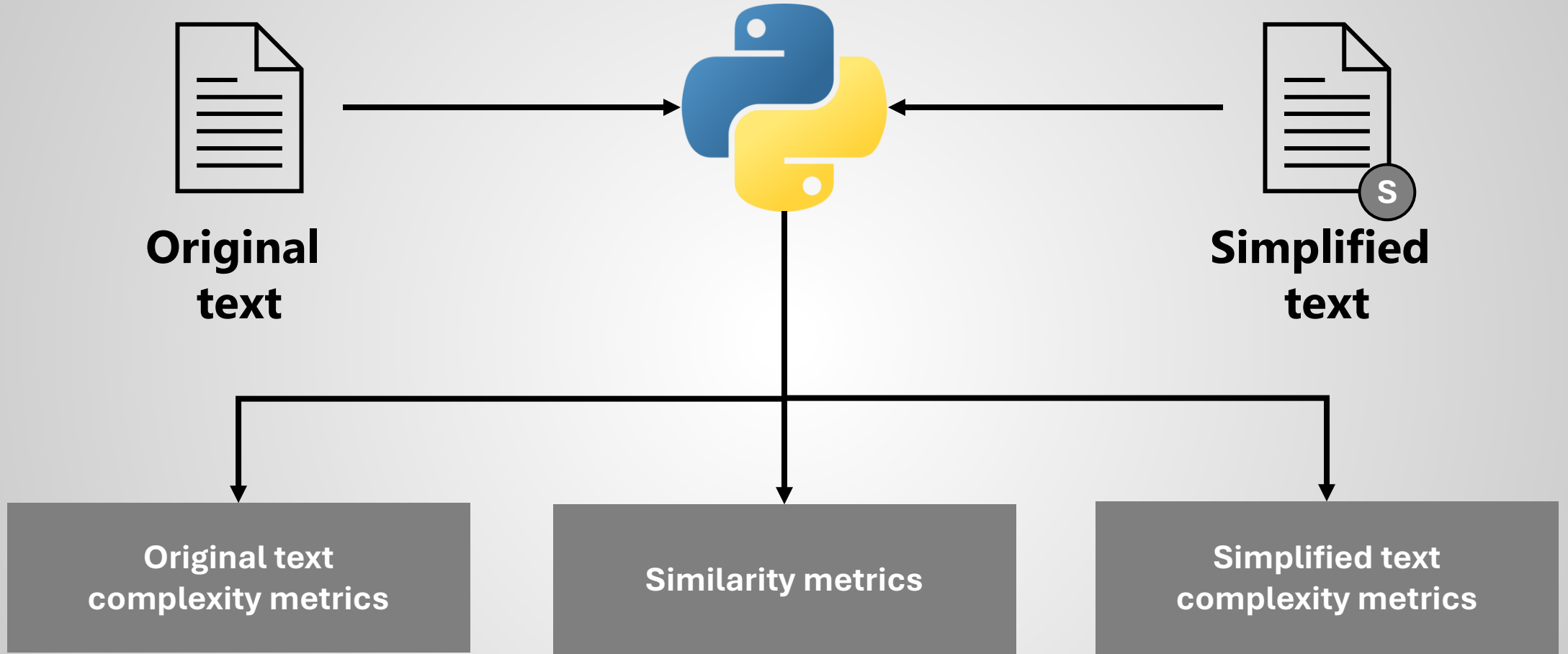


13 166
tokens

ItaRegol Corpus



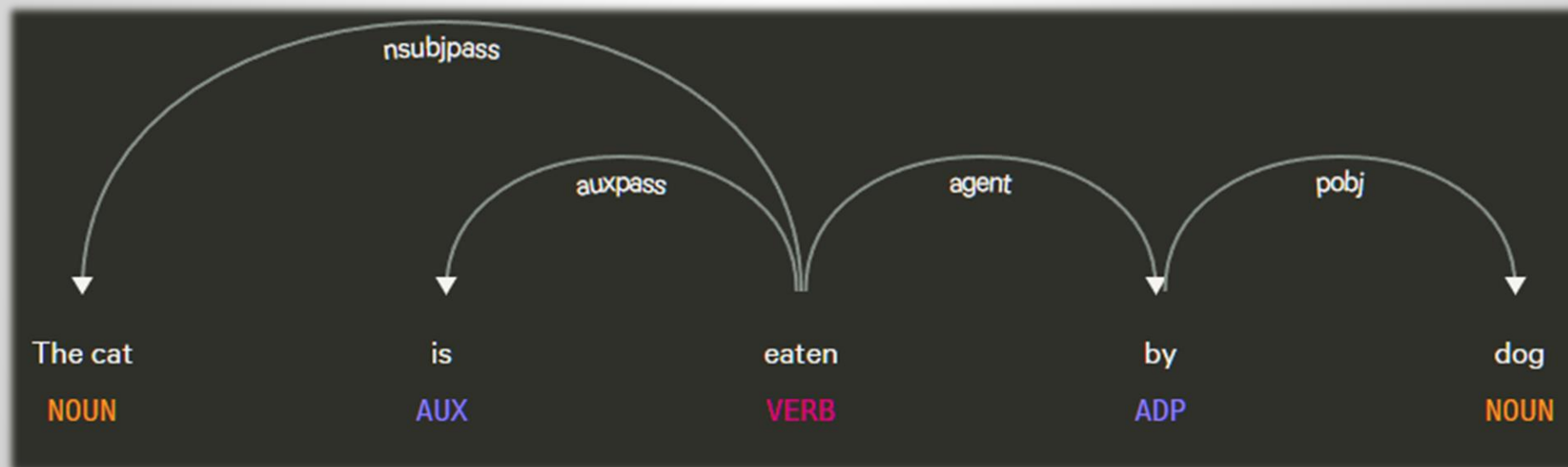
Pipeline



Pipeline

Complexity metrics

The distinction between active and passive verbs was achieved by analyzing the **Dependency Tree structure**, where specific annotations such as “**nsubj:pass**” (passive nominal subject) and “**aux:pass**” (passive auxiliary) provide crucial grammatical insights. This method allows to accurately identify passive verbs.



% passive verbs

This vocabulary contains the list of all **high-use** and **high-frequency** words of the modern Italian language.

Il nostro dizionario indica le parole fondamentali. Il dizionario online che parole di uso comune, il corsivo indica le parole di alta disponibilità.

23 dic 2016

Internazionale | dizionario.internazionale.it/nuevovocabolariodibase

Dizionario

Il Nuovo vocabolario di base della lingua italiana

A cura di Tullio De Mauro

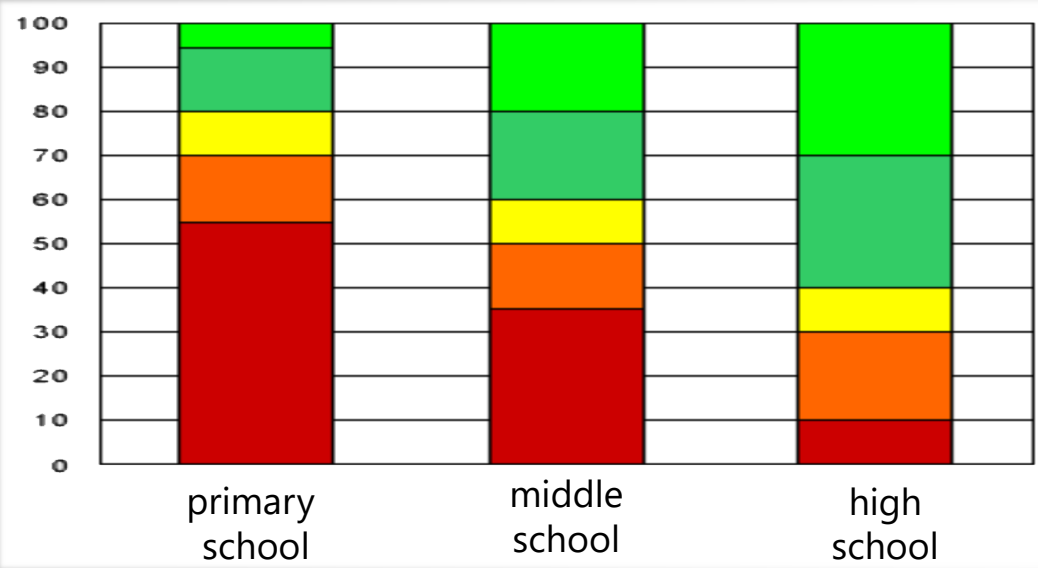
23 dicembre 2016

A

'a s.f. e m.inv., 'a prep., **abbagliante** p.pres., agg., s.m., **abbaiare** v.intr. e tr., **abbandonare** v.tr., abbandonato p.pass., agg., s.m., abbandono s.m., **abbassare** v.tr., **abbasso** avv., inter., **abbastanza** avv., abbattere v.tr., **abbrèverare** v.tr., abbreviato s.m., **abbinare** v.tr., abbonamento s.m., **abbonare** v.tr., abbondante p.pres., agg., **abbondare** v.intr., **abbottonare** v.tr., **abbracciare** v.tr., abbraccio s.m., **abbreviare** v.tr., **abbronzare** v.tr., **abete** s.m., abile agg., abilità s.infv., abisso s.m., abitante p.pres., agg., s.m., **abitare** v.intr. e tr., s.m., **abitazione** s.f., **abito** s.m., abituale agg., **abitare** v.tr., **abitudine** s.f., abolire v.tr., **abortire** v.intr., **aborto** s.m., **abruzzese** agg., s.m. e f., **abusare** v.intr., abuso s.m., **acca** s.f. e m.inv., accademia s.f., accademico agg., s.m., **accadere** v.intr., **accampamento** s.m., **accento** avv., **accapatoio** s.m., accarezza v.tr., **accattare** s.m., **accavallare** v.tr., **accedere** v.tr., intr., accedere v.intr., accelerare v.tr. e intr., **acceleratore** agg., s.m., accelerazione s.f., **accendere** v.tr., **accendino** s.m., accennare v.tr. e intr., **accenno** s.m., **accettare** v.tr., accertamento s.m., accertare v.tr., accesso p.pass., agg., **accesso** s.m., accessorio agg., s.m., **accetta** s.f., accettabile agg., **accettare** v.tr., **acchiappare** v.tr., **acciaio** s.m., acciaio s.m., accidente s.m., **acciaia** s.f., **accogliente** p.pres., agg., accoglienza s.f., **accogliere** v.tr., **accollare** v.tr., accomodare v.tr. e intr., **accompnare** v.tr., **acommentare** v.intr. e tr., accontentare v.tr., **accorciare** v.tr., v.intr., accordare v.tr., **accordo** s.m., **accorgersi** v.pronom.intr., **accorrere** v.intr., accostare v.tr. e intr., **accudire** v.intr. e tr., accumulare v.tr., **accumulatore** agg., s.m., accurato p.pass., agg., **accusa** s.f., **accusare** v.tr., acento s.m., **acerbo** agg., **aceto** s.m., acido agg., s.m., **acqua** s.f., **acquarello** s.m., **acquario** s.m., **acquasanta** s.f., acquisire v.intr., acquisizione s.f., **acquistare** v.tr. e intr., **acquisto** s.m., **acquolina** s.f., **acrobatia** s.f., acuto agg., s.m., adattare v.tr., **adattatore** s.m., **adatto** agg., **addetto** p-pass., agg., s.m., addio inter., s.m., **addirittura** avv., **addizione** s.f., **addobbare** v.tr., **addolcire** v.tr., **addomesticare** v.tr., addormentarsi v.pronom.intr., **addormentato** p-pass., agg., **addossare** v.tr., **addosso** avv., intr., **addrizzare** v.tr., adeguare v.tr., adeguato p-pass., agg., s.m., **aderente** p-pres., agg., s.m. e f., aderire v.intr., adesione s.f., **adesso** avv., **adolescente** agg., s.m. e f., adolescenza s.f., adoperare v.tr., v.intr., **adornare** v.tr., **adottare** v.tr., adozione s.f., **adriatico** agg., **adulto** agg., s.m., **aereo** agg., s.m., **aereo** s.m., **aeroplano** s.m., aeropuerto s.m., **afa** s.f., affacciare v.tr. e intr., **affamare** v.tr., v.intr., **affamato** p-pass., agg., s.m., **affannarsi** v.pronom.intr., **affannato** p-pass., agg., **affanno** s.m., **affare** s.m., affascinante p-pres., agg., **affascinare** v.tr., **affaticare** v.tr., **affatto** avv., **affermare** v.tr., affermazione s.f., affermare v.tr., v.intr., **affettare** v.tr., **affettato** p-pass., agg., s.m., **affetto** s.m., **affetto** avv., affettuoso agg., **affezionato** p-pass., agg., affiancare v.tr., affidamento s.m., **affidare** v.tr., **affilato** p-pass., agg., affinché cong., affittare v.tr., affitto s.m., **affogare** v.tr. e intr., **affollare** v.tr., affondare v.tr. e intr., affresco s.m., **affrontare** v.tr., **affumicare** v.tr., africa n.agg., s.m., agenda s.f., **agente** p-pres., agg., s.m. e f., **agenzia** s.f., agganziare v.tr., aggettivo s.m., aggiornamento s.m., aggiornare v.tr., aggirare v.tr., **aggiungere** v.tr., v.intr., aggiustare v.tr., aggrapparsi v.pronom.intr., **aggravare** v.tr., v.intr., aggredire v.tr., aggressione s.f., aggressivo agg., **agiato** p-pass., agg., **agile** agg., ago s.m., **agire** v.intr., agitare v.tr., agitazione s.f., **aglio** s.m., agnello s.m., ago s.m., **agonia** s.f., **agosto** s.m., agricolo agg., **agricoltore** s.m., agricoltura s.f., **agrimo** s.m., **aguazare** v.tr., **aguezo** agg., **aiutola** s.f., **aiutare** v.tr., **aiuto** s.m., inter., **'ala** s.f., **'alba** s.f., albanese agg., s.m. e f., **albergo** s.m., **'albero** s.m., **albicocca** s.f., s.m.inv., agg.infv., **album** s.m.inv., alcol s.m., **alcuno** agg.indef., pron.indef.m., **alfabeto** s.m., **alga** s.f., **algerino** agg., s.m., alieno agg., s.m., **alimentare** agg., s.m., **alimentare** v.tr., **alimentari** s.m.inv., alimentazione s.f., alimento s.m., **alito** s.m., **allacciare** v.tr., **allagare** v.tr., allargare v.tr. e intr., **allarmare** v.tr., allarme s.m., **allattare** v.tr., alleanza s.f., **allearsi** v.pronom.intr., alleato p-pass., agg., s.m., **allegato** p-pass., agg., s.m., **alleggerire** v.tr., allegria s.f., allegro agg., avv., s.m., allenamento s.m., **allenare** v.tr., allenatore agg., s.m., **allentare** v.tr., **allerigia** s.f., **allievare** v.tr., **allievo** s.m., **allineare** v.tr., alloggio s.m., **allontanare** v.tr., **allora** avv., cong., **alluce** s.m., alludere v.intr., **alluminio** s.m., **allungare** v.tr., **alluvione** s.f., **almeno** avv., alquanto agg.indef., pron.indef., avv., **altalenà** s.f., altamente avv., **altare** s.m., alterare v.tr., alternare v.tr. e intr., alternativa s.f., alternativo agg., **alterno** agg., **altezza** s.f., **'alto** agg., s.m., avv., **altocinese** agg., s.m., **altopiano** s.m., **altrettanto** agg.indef., pron.indef., avv., **altrimenti** avv., **altro** agg.indef., pron.indef.m., avv., altrove avv., altrui agg.poss.infv., pron.indef.infv., alunno s.m., **alzare** s.m., **alzare** v.tr., **'amante** p-pres., agg., s.m. e f., **amare** v.tr., amaro agg.,

Il Nuovo vocabolario di base della lingua italiana - Tullio De Mauro - Internazionale (dizionario.internazionale.it/nuevovocabolariodibase)

Gulpease Index



$$89 + \frac{300 \cdot \text{sentences} - 10 \cdot \text{characters}}{\text{tokens}}$$

Flash-Vacca Index

Flesch Ease Score	Reading Readability Level/ Category	Estimated Reading Grade
0-30	Very Difficult	College graduate
30-40	Difficult	13 th to 16 th grade
50-60	Fairly difficult	10 th to 12 th grade
60-70	Standard	8 th or 9 th grader
70-80	Fairly easy	7 th grader
80-90	Easy	6 th grader
90-100	Very Easy	5 th grader

$$206 - 65 \cdot \frac{\text{syllables}}{\text{tokens}} - \frac{\text{tokens}}{\text{sentences}}$$

Readability Index

READ-IT is a metric specifically designed for assessing **readability**, with a focus on **text simplification**.

Key Features considered in READ-IT:

- **Lexical features**
- **Morphographic features**
- **Syntactic features**

READ-IT provides a comprehensive framework for evaluating text readability.

READ-IT:
Assessing Readability of Italian Texts with a View to Text Simplification

Felice Dell'Orletta, Simonetta Montemagni and Giulia Venturi
Istituto di Linguistica Computazionale "Antonio Zampolli" (ILC-CNR)
via G. Moruzzi, 1 – Pisa (Italy)
{felice.dellorletta, simonetta.montemagni, giulia.venturi}@ilc.cnr.it

Abstract

In this paper, we propose a new approach to readability assessment with a specific view to the task of text simplification: the intended audience includes people with low literacy skills and/or with mild cognitive impairment. READ-IT represents the first advanced readability assessment tool for what concerns Italian, which combines traditional raw text features with lexical, morpho-syntactic and syntactic information. In READ-IT readability assessment is carried out with respect to both documents and sentences where the latter represents an important novelty of the proposed approach creating the prerequisites for aligning the readability assessment step with the text simplification process. READ-IT shows a high accuracy in the document classification task and promising results in the sentence classification scenario.

1 Introduction

Recently, there has been increasing interest in the exploitation of results from Natural Language Processing (NLP) for the development of assistive technologies. Here, we address this topic by reporting the first but promising results in the development of a software architecture for the Italian language aimed at assisting people with low literacy skills (both native and foreign speakers) or who have language disabilities in reading texts.

Within an information society, where everyone should be able to access all available information, improving access to written language is becoming more and more a central issue. This is the case, for

instance, of administrative and governmental information which should be accessible to all members of the society, including people who have reading difficulties for different reasons: because of a low education level or because of the fact that the language in question is not their mother tongue, or because of language disabilities. Health related information represents another crucial domain which should be accessible to a large and heterogeneous target group. Understandability in general and readability in particular is also an important issue for accessing information over the web as stated in the Web Content Accessibility Guidelines (WCAG) proposed by the Web Accessibility Initiative of the W3C.

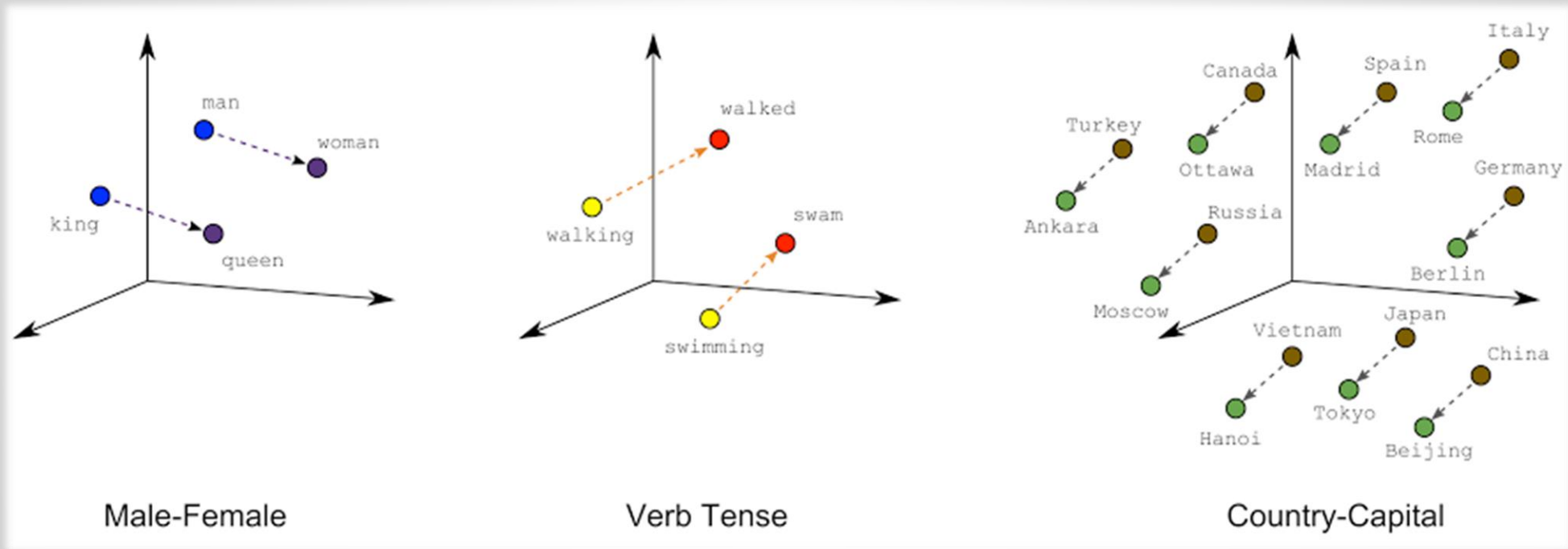
In this paper, we describe the approach we developed for automatically assessing the readability of newspaper texts with a view to the specific task of text simplification. The paper is organized as follows: Section 2 describes the background literature on the topic; Section 3 introduces the main features of our approach to readability assessment, with Section 4 illustrating its implementation in the READ-IT prototype; Sections 5 and 6 describe the experimental setting and discuss achieved results.

2 Background

Readability assessment has been a central research topic for the past 80 years which is still attracting considerable interest nowadays. Over the last ten years, within the NLP community the automatic assessment of readability has received increasing attention: if on the one hand the availability of sophisticated NLP technologies makes it possible to monitor a wide variety of factors affecting the readability

Similarity metrics

The **sentence embeddings** represent sentences as vectors in a multidimensional space, capturing their semantic meaning. By calculating the **cosine similarity** between two vectors (sentences), it is possible to quantify the semantic similarity between the texts.

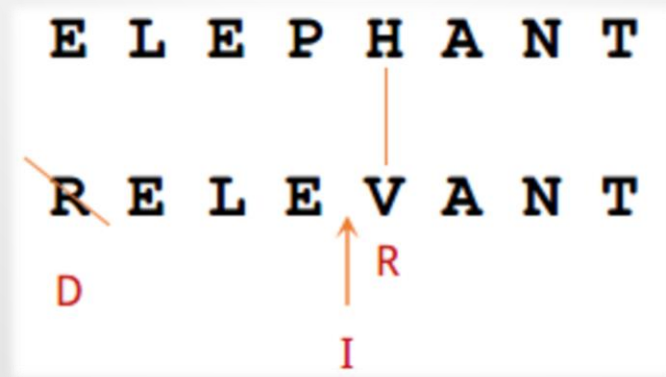


Semantic Text Similarity

Edit distance (or Levenshtein distance) is a metric used to quantify how **dissimilar two strings** are.

It is calculated by determining the minimum number of operations needed to transform one string into the other:

- **Insertion:** Adding a character.
- **Deletion:** Removing a character.
- **Replacement:** Substitute one character with another.



Edit distance

RQ1

How effective is the CHAIN prompting in simplifying administrative documents compared to the BASIC prompting?

Corpora	Tokens	Sentences	% Passive verbs	% Easy words
Original	13166	519	26 %	72 %
BASIC – gpt-4o	9171	552	13 %	80 %
BASIC – gpt-4o-mini	9851	600	20 %	79 %
CHAIN – gpt-4o	<u>13174</u>	631	8 %	75 %
CHAIN – gpt-4o-mini	<u>13381</u>	689	13 %	76 %

Complexity metrics

Corpora	Gulpease	Flesch-Vacca	READ-IT
Original	44	25	92 %
BASIC – gpt-4o	53	41	45 %
BASIC – gpt-4o-mini	52	39	47 %
CHAIN – gpt-4o	49	33	59 %
CHAIN – gpt-4o-mini	48	32	56 %

Complexity metrics – Readability indexes

Corpora	Semantic Similarity	Edit distance
Original vs BASIC – gpt4o	95 %	52 %
Original vs BASIC – gpt4o-mini	96 %	46 %
Original vs CHAIN – gpt4o	98 %	18 %
Original vs CHAIN – gpt4o-mini	98 %	25 %

Similarity metrics

RQ1

How effective is the CHAIN prompting in simplifying administrative documents compared to the BASIC prompting?

Complexity metrics

BASIC prompting shows better complexity metrics
(except for sentences and % passive verbs)

Similarity metrics

CHAIN prompting preserves information without altering too much the structure of the original document

The difference between the metrics are statistically significant

RQ1 outcome

RQ2

How effective is the CHAIN prompting in maintaining juridical meaning compared to the BASIC prompting?

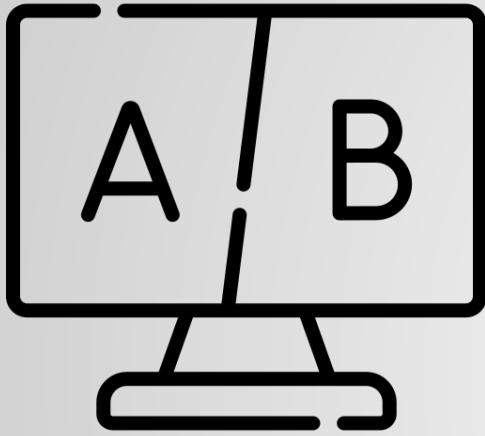


**Two
experts**



**Juridical
evaluation**

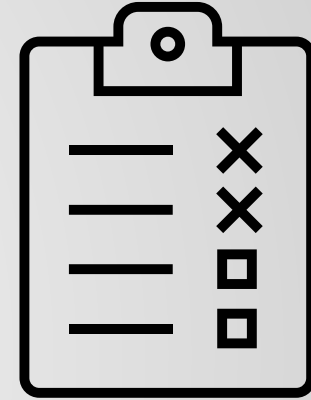
Comparison study



**A/B
test**



**Texts
pairs**



**Post-reading
questionnaire**

Comparison study

	ORIGINAL vs BASIC	ORIGINAL vs CHAIN
1	Jurist A	Jurist B
2	Jurist A	Jurist B
3	Jurist A	Jurist B
4	Jurist A	Jurist B
...
26	Jurist A	Jurist B
27	Jurist B	Jurist A
...
49	Jurist B	Jurist A
50	Jurist B	Jurist A
51	Jurist B	Jurist A
52	Jurist B	Jurist A

Comparison study



Regolamento degli Asili Nido

Art. 8 Assenza e decadenza

Testo A

Le assenze degli utenti, superiori a 5 giorni consecutivi, comportano, sempre, la presentazione del certificato medico al momento del rientro al nido.

Le assenze del bambino, superiori a 10 giorni consecutivi, non giustificate per iscritto, nonostante il sollecito, comportano la decadenza dal posto; la richiesta di giustificazione, con le predette modalità è prevista anche per assenze inferiori a 10 giorni se ripetute per un totale di 20 giorni in un trimestre. In questi casi la decadenza è proposta dal Comitato di gestione e disposta dal Direttore educativo con proprio atto, la copertura dei posti vacanti per rinuncia o per decadenza va predisposta da parte del Coordinatore educativo di concerto con il Direttore educativo nei dieci giorni successivi all'avvenuta vacanza e disposta dal Direttore educativo.

Testo B

Le assenze degli utenti, superiori a 5 giorni consecutivi, comportano sempre la presentazione del certificato medico al momento del rientro al nido.

Le assenze del bambino, superiori a 10 giorni consecutivi e non giustificate per iscritto, comportano la perdita del posto, nonostante il sollecito. La richiesta di giustificazione, con le stesse modalità, è prevista anche per assenze inferiori a 10 giorni se si ripetono per un totale di 20 giorni in un trimestre. In questi casi, il Comitato di gestione propone la perdita del posto e il Direttore educativo dispone la perdita del posto con proprio atto.

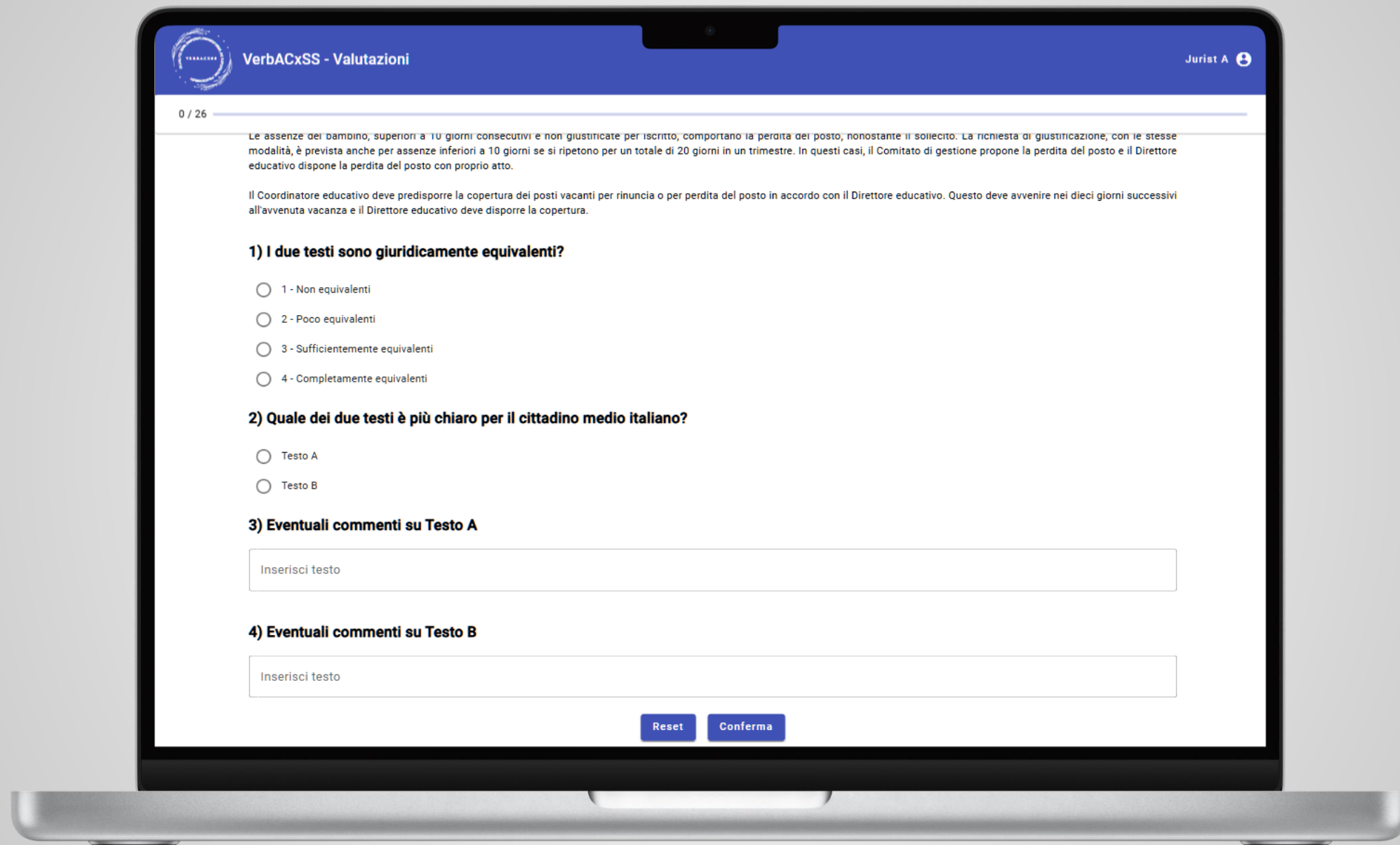
Il Coordinatore educativo deve predisporre la copertura dei posti vacanti per rinuncia o per perdita del posto in accordo con il Direttore educativo. Questo deve avvenire nei dieci giorni successivi all'avvenuta vacanza e il Direttore educativo deve disporre la copertura.

1) I due testi sono giuridicamente equivalenti?

- ☐ 1 - Non equivalenti
- ☐ 2 - Poco equivalenti
- ☐ 3 - Sufficientemente equivalenti
- ☐ 4 - Completamente equivalenti

2) Quale dei due testi è più chiaro per il cittadino medio italiano?

Comparison study – WebApp



Comparison study – WebApp

Are the two texts juridically equivalent? [4-point Likert scale]

1. Not equivalent
2. Slightly equivalent
3. Moderately equivalent
4. Fully equivalent

Which of the two texts is clearer for the average Italian citizen? [binary choice]

- Text A
- Text B

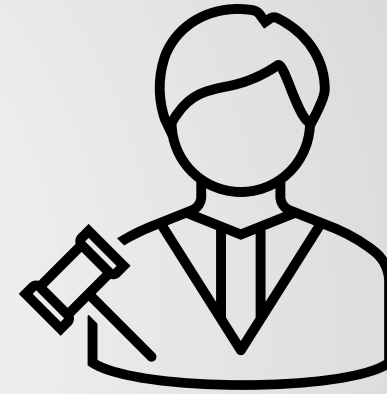
Additional comments on Text A [free text]

Additional comments on Text B [free text]

Comparison study – Questions



Jurist A
189 minutes



Jurist B
122 minutes

Comparison study – Evaluation time

~76 %

ORIGINAL-CHAIN pairs are "**Fully equivalent**"

~51 %

ORIGINAL-BASIC pairs are "**Moderately equivalent**".

*The difference between the juridical equivalence of ORIGINAL-BASIC and ORIGINAL-CHAIN pairs is **statistically** significant*

Comparison study – Juridical equivalence score

~ 77 %

preferred the “**simplified**” version among all the evaluated pairs

Comparison study – Citizen preference

13 comments

are related to the legal validity of the document

«Text B shows more appropriate legal language.»

"Il testo B presenta linguaggio giuridico più appropriato.»

Comments – ORIGINAL text

8 comments

are related to the juridical style

«Text A, although more understandable, has language and constructions not common in the legal context.»

"Il testo A, sebbene effettivamente più comprensibile, presenta linguaggio e costruzioni non comuni nel contesto giuridico (ad es. l'uso del plurale "vietiamo" invece dell'impersonale "è vietato").»

Comments – CHAIN text

27 comments

are related to the juridical style

«slightly imprecise || language too simple; law also has its own form.»

«leggermente impreciso || linguaggio troppo semplice; il diritto ha anche una sua forma.»

Comments – BASIC text

RQ2

How effective is the CHAIN prompting in maintaining juridical meaning compared to the BASIC prompting?

Juridical equivalence

CHAIN prompting is more effective at preserving juridical meaning compared to BASIC prompting.

Preference

Both BASIC and CHAIN simplified documents are preferred for citizens.

Comments

The style of BASIC document is not suitable for juridical documents.

RQ2 outcome

Conclusion

Open problems:

Two opposite needs:

- Citizens: improving metrics (increasing readability) and increasing user comprehension (reduce the relative complexity)
- Administration: preserving the legal value of (simplified) texts.

Possible solutions:

- Add to the chain - after the Nominalizations rule – a rule on the basic vocabulary to be increased to the detriment of uncommon, rare words // or Glossary
- Or Apply in sequence, the *chain prompt* followed by the *basic prompt*

Future prospects:

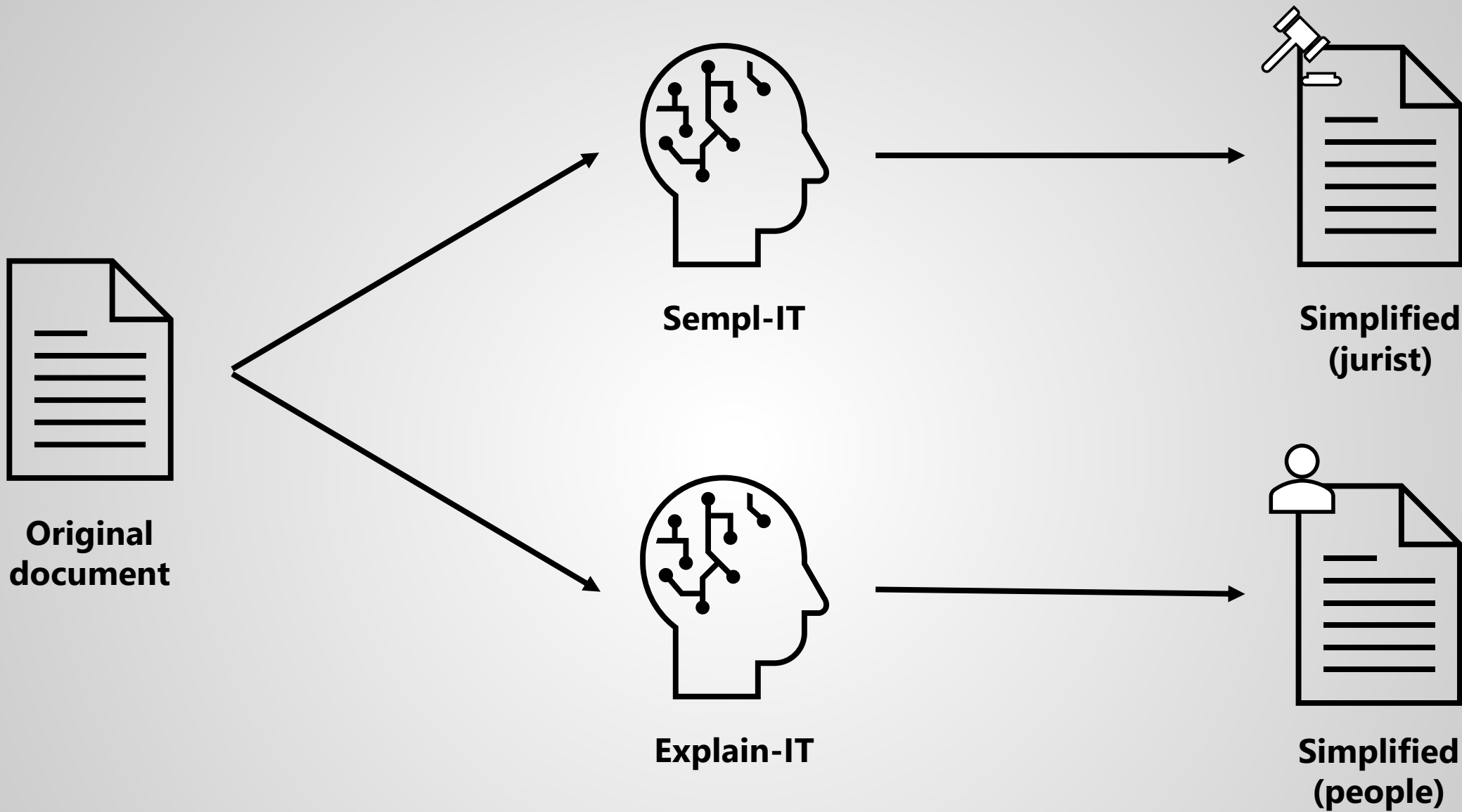
- apply the chain prompt to small Italst which has already been simplified with the basic prompt
- re-verify the metrics of all these operations
- re-do the comprehension experiment (to both corpora)
- compare the two prompts on Regulations and on Italst and evaluate the different characteristics of differently legally binding texts

Conclusion

Metrics	Original s-Italst	Simplified s-Italst	Original ItaRegol	Simplified ItaRegol
Gulpease Index	44	51	44	53
Flesch-Vacca Index	19	36	25	41
READ-IT	86 %	54 %	92 %	45 %
% Easy words	73 %	81 %	72 %	80 %
% Passives verbs	20 %	12 %	26 %	13 %
Semantic Similarity	-	96 %	-	95 %
Edit Distance	-	52 %	-	52 %

Outcomes of the Basic prompt show similar effects on the metrics independently from the different legally binding properties of the texts.

BASIC prompt comparison: outcomes



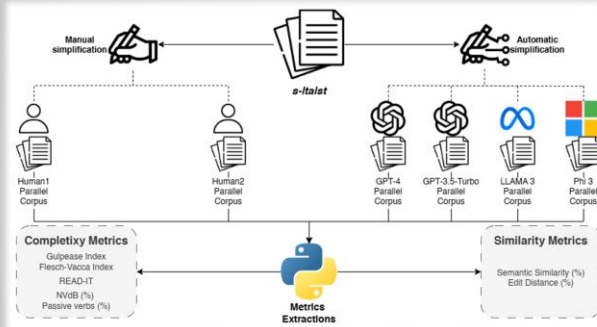
BASIC prompt

Thanks for your attention



Explore simplified parallel corpora

Summary



Experimental Phase 1 - Pipeline

1. Select administrative documents
2. Select a short excerpt from the document
3. Simplify the selected document
4. Formulate text comprehension questions
5. Administer the test

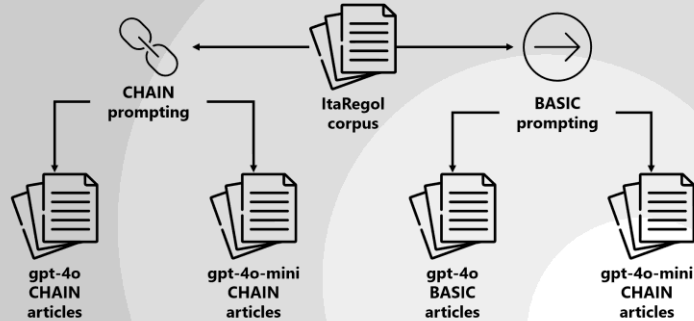


Experimental phase 2: Comprehension assessment

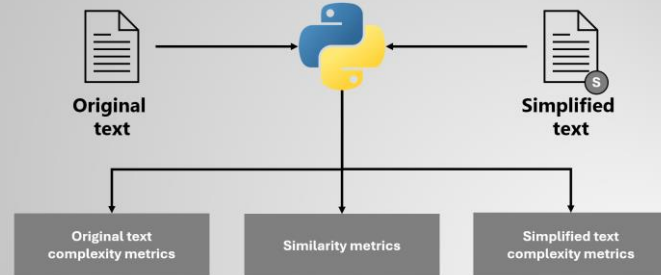
- Regulations of **garbage tax**
- Regulations of **urban green**
- Regulations of the **air quality**
- Regulations of **sport facilities**
- Regulations of **nursery schools**
- Regulations for **urban building**
- Regulations of **animal protection**
- Regulations for cargo and bus **driving license**



ItaRegol Corpus



Pipeline



Pipeline



Comparison study – WebApp



WEBSITE



CORPUS

