HAL
open science

Overview of the CLEF 2022 SimpleText Task 3: Query Biased Simplification of Scientific Texts

Liana Ermakova, Irina Ovchinnikova, Jaap Kamps, Diana Nurbakova, Sílvia Araújo, Radia Hannachi, Nicolas Poinsu

To cite this version:

Liana Ermakova, Irina Ovchinnikova, Jaap Kamps, Diana Nurbakova, Sílvia Araújo, et al.. Overview of the CLEF 2022 SimpleText Task 3: Query Biased Simplification of Scientific Texts. Proceedings of the Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, Bologna, Italy, September 5th - to - 8th, 2022, Sep 2022, Bologna, Italy. pp.2792-2804. hal-03828409

HAL Id: hal-03828409

https://hal.science/hal-03828409v1

Submitted on 25 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Overview of the CLEF 2022 SimpleText Task 3: Query Biased Simplification of Scientific Texts

Liana Ermakova[1], Irina Ovchinnikov[2], Jaap Kamps[3], Diana Nurbakova[4], Sílvia Araújo[5] and Radia Hannachi[6]

[1]Université de Bretagne Occidentale, HCTI, France
[2]ManPower Language Solution, Israel
[3]University of Amsterdam, Amsterdam, The Netherlands
[4]University of Lyon, INSA Lyon, CNRS, LIRIS, UMR5205, Villeurbanne, France
[5]Universidade do Minho, CEHUM, 4710-057 Braga, Portugal
[6]Université de Bretagne Sud, HCTI, 56321 Lorient, France

Abstract

This paper presents an overview of the CLEF 2022 SimpleText Task 3 on query biased simplification of scientific text. After discussing the motivation and general task setup, we detail the exact test collection, consisting of a train of sentences from scientific abstracts paired with human reference simplified sentences, and extensive test corpus of sentences with detailed annotations of lexical and syntactic complexity. We present a detailed analysis of the submitted simplified sentences, and the resulting evaluation scores.

Keywords

automatic text simplification, science popularization, information distortion, error analysis, lexical complexity, syntactic complexity

1. Introduction

Digitization and open access have made scientific literature available to every citizen. While this is an important first step, there are several remaining barriers preventing laypersons to access the objective scientific knowledge in the literature. In particular, scientific texts are often hard to understand as they require solid background knowledge and use tricky terminology. Although there were some recent efforts on text simplification (e.g. [1]), removing such understanding barriers between scientific texts and general public in an automatic manner is still an open challenge.

The CLEF 2022 SimpleText track brings together researchers and practitioners working on the generation of simplified summaries of scientific texts. It is a new evaluation lab that follows up the SimpleText-2021 Workshop [2]. The track provides data and benchmarks for discussion of challenges of automatic text simplification by bringing in the following interconnected tasks:

CLEF 2022: Conference and Labs of the Evaluation Forum, September 5–8, 2022, Bologna, Italy
liana.ermakova@univ-brest.fr (L. Ermakova)
https://simpletext-project.com/ (L. Ermakova)
0000-0002-7598-7474 (L. Ermakova); 0000-0003-1726-3360 (I. Ovchinnikov); 0000-0002-6614-0087 (J. Kamps); 0000-0002-6620-7771 (D. Nurbakova); 0000-0003-4321-4511 (S. Araújo)
© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

Table 1
CLEF 2022 SimpleText official run submission statistic

Team	Task 1		Task 2		Task 3		Total runs
aaac			1	(1 updated)			1
CLARA-HD [6]					1		1
CYUT Team2 [7]	1				1		2
HULAT-UC3M [8]					10	(4 updated)	10
LEA_T5 [9]			1		1		2
NLP@IISERB [10]	3	(3 updated)					3
PortLinguE [11]					1	(1 updated)	1
SimpleScientificText [12]			1	(1 updated)			1
UAms [13]	2		1				3
Total runs	6		4		14		24

Task 1: What is in (or out)? Select passages to include in a simplified summary, given a query.

Task 2: What is unclear? Given a passage and a query, rank terms/concepts that are required to be explained for understanding this passage (definitions, context, applications,..).

Task 3: Rewrite this! Given a query, simplify passages from scientific abstracts.

This paper focuses on the third task of text simplification proper. We refer for details of the other tasks to the overview papers of Task 1 [3] and Task 2 [4], or the Track overview paper [5].

In the CLEF 2022 edition of SimpleText, a total of 62 teams registered for the SimpleText track. A total of 40 users downloaded data from the server. A total of 9 distinct teams submitted 24 runs, of which 10 runs were updated. The details of statistics on runs submitted for shared tasks are presented in Table 1. As for the third task, a total of 14 runs from five teams were submitted.

This introduction is followed by Section 2 presenting the text simplification task with the datasets and evaluation metrics used. In Section 3, we discuss the results of the official submissions. We end with Section 4 discussing the results and findings, and lessons for the future.

2. CLEF 2022 SimpleText Task 3 Test Collection

In this section, we discuss the third task about text simplification proper, rewriting an extracted sentence from a scientific abstract, addressing the task:

Given a query, simplify passages from scientific abstracts.

The goal of this task is to provide a simplified version of text passages (sentences) with regard to a query. Participants were provided with queries and abstracts of scientific papers. The abstracts could be split into sentences. The simplified passages were evaluated manually in terms of the produced errors as follows.

Table 2
SimpleText Task 3: Statistics of the number of evaluated sentences per query

Query		# Distinct source sentences	# Distinct simplified sentences
1	*digital assistant*	370	1,280
2	*conspiracy theories*	195	398
3	*end to end encryption*	55	102
4	*imbalanced data*	55	87
5	*genetic algorithm*	51	85
6	*quantum computing*	51	85
7	*qbit*	50	76
8	*quantum applications*	42	73
9	*cyber-security*	28	47
10	*fairness*	18	22
11	*crowsourcing*	14	21

2.1. Train Data

As for *Task 2: What is unclear?*, we provided a parallel corpus of simplified sentences from two domains: *Medicine* and *Computer Science*. As previously, we use scientific abstracts from the DBLP Citation Network Dataset for *Computer Science* and Google Scholar and PubMed articles on muscle hypertrophy and health *Medicine* [14, 15].

Text passages issued from abstracts on computer science were simplified by either a master student in Technical Writing and Translation or a pair of experts: (1) a computer scientist and (2) a professional translator, English native speaker but not specialist in computer science [15]. Each passage was discussed and rewritten multiple times until it became clear for non-computer scientists. Medicine articles were annotated by a master student in Technical Writing and Translation specializing in this domain. Sentences were shortened, excluding every detail that was irrelevant or unnecessary to the comprehension of the study, and rephrased, using simpler vocabulary. If necessary, concepts were explained.

We provided 648 parallel sentences in total.

2.2. Test Data

We used the same 116,763 sentences retrieved by the ElasticSearch engine from the DBLP dataset according to the queries as for Task 2. We manually evaluated 2,276 pairs of sentences for 11 queries. For the query *Digital assistant* we took the first 1,000 sentences retrieved by ElasticSearch. We pool source sentences coupled with their simplified versions submitted by all participants for all these queries. We ensured that for each evaluated source sentence the pool contained the results of all participants. The detailed statistics of the number of evaluated sentences per query for Task 3 are given in Table 2.

2.3. Input and Output Format

The input train and and the test data were provided in JSON and CSV formats with the following fields:

snt_id a unique passage (sentence) identifier.

source_snt passage text.

doc_id a unique source document identifier.

query_id a query ID.

query_text simplification should be done with regard to this query.

Input example (JSON format):

```
{"snt_id":"G11.1_2892036907_2", "source_snt":"With the ever increasing number of
↪  unmanned aerial vehicles getting involved in activities in the civilian and
↪  commercial domain, there is an increased need for autonomy in these systems
↪  too.", "doc_id":2892036907, "query_id":"G11.1", "query_text":"drones"}
```

Participants were asked to provide a list of terms to be contextualized in a JSON format or a tabulated file TSV (for manual runs) with the following fields:

run_id Run ID starting with (team_id)_(task_3)_(*name*).

manual Whether the run is manual $\{0, 1\}$.

snt_id a unique passage (sentence) identifier from the input file.

simplified_snt Text of the simplified passage .

Output example (JSON format):

```
{"run_id":"BTU_task_3_run1", "manual":1, "snt_id":"G11.1_2892036907_2",
↪  "simplified_snt":"Drones are increasingly used in the civilian and commercial
↪  domain and need to be autonomous."}
```

2.4. Evaluation metrics

We filtered out the simplified sentences identical to the source ones and the truncated simplified sentences by keeping only passages matching the regular expression (valid snippets): `.+[?.!"\']\s*\$'`

Professional linguists manually annotated simplifications provided with regard to a query according to the following criteria. We evaluated binary errors:

- Incorrect syntax;
- Unresolved anaphora due to simplification;
- Unnecessary repetition/iteration (lexical overlap);
- Spelling, typographic or punctuation errors.

The lexical and syntax complexity of the produced simplifications were assessed on an absolute scale, value 1 referring to a simple output sentence regardless of the complexity of the source one, 7 corresponding to a complex one. Lexical complexity is mostly identical to that presented in Task 2, see the Track overview [5] and Task overview [4].

We consider **syntax complexity** based on syntactic dependencies, their length and depth. The dependency trees reveal latent complications for reading and understanding text; thus, psycho-linguists consider the syntactic dependencies to be a relevant tool to evaluate text readability [16]. The depth and length of the syntactic chains we interpret according to [16].

We evaluate **syntax complexity** as follows:

1. Simple sentence (without negation / passive voice): *Over Facebook, we find many interactions.*
2. Simple sentence with negation / passive voice (e.g. *Many interactions were found over Facebook*) or Simple sentences with syntactic constructions that show chains of dependency and shallow embedding depth (e.g. *Over Facebook, we find many interactions between public pages and both political wings.*)
3. Simple sentences with long chains of dependency and shallow embedding depth, with syntactic constructions like complex object, gerund construction, etc. (e.g. *Despite the enthusiastic rhetoric about the so-called collective intelligence, conspiracy theories have emerged.*) or Short complex or compound sentence (e.g. *We propose a novel approach that was used in terms of information theory.*)
4. Simple sentences with long chains of dependency and deep embedding depth, with syntactic constructions like complex object, gerund construction, etc. (e.g. *Over Facebook, we find many interactions between public pages for military and veterans, and both sides of the political spectrum*) or Complex or compound sentence that contains long chains of dependency and deep embedding depth;
5. Simple sentences with long chains of dependency and deep embedding depth, with several syntactic constructions like complex object, gerund construction, etc. or & Complex or compound sentence that contains long chains of dependency and deep embedding depth;
6. Complex or compound sentences that contain long chains of dependency and deep embedding depth along with complex object, gerund construction, etc. or Simple sentence that contains modifications, topicalization, parenthetical constructions: *Moreover, we measure the effect of 4709 evidently false information (satirical version of conspiracist stories) and 4502 debunking memes (information aiming at contrasting unsubstantiated rumors) on polarized users of conspiracy claims.*
7. Long complex or compound sentences that contain several clauses of different types, long chains of dependency and deep embedding depth along with complex object, gerund construction, etc.

We evaluate the information quality of the simplified snippet based on its content and readability. Transformation of information from the source snippet brings in omission of details, insertion of basic terms to explain particular terminology and complex concepts, reference to resources. Due to necessary insertions and references, the simplified snippets often contain more words and syntactic constructions as compared to their source. Nevertheless, the goal is to

reduce lexical and syntax complexity in the extended simplified snippets. In case the simplified snippet lacks information mentioned in the source, we evaluate the degree of the information loss. Irrelevant insertions, iterations and wordy statements in the extended simplified snippet we consider as a misrepresentation or distortion of source information when a reader experiences difficulties in processing source content due to wordiness of the loosely structured simplified snippet.

We assessed the information loss severity during the simplification with regard to a given query on the scale from 1 to 7, where 1 corresponds to unimportant information loss while 7 refers to the most severe information distortion. We consider the information loss as a kind of information damage even if the information in the simplified text contains the information of the source passage but has some insertions, which impedes perception of the content.

We distinguish the following 11 types of misrepresentation of source information. Our classification leans on the error typology in machine translation [17, 18]. The simplified snippet often combines several types of distortion, e.g. omission and ambiguity. Nevertheless, we observed many instances of small distortions that severely diminish the quality of the simplification; therefore, we need to explain each type providing the clear and transparent instances. Our evaluation of the value of the information distortion leans on the calculation of the information loss and assessment of the diminished readability of the simplified snippet that generates difficulties in text semantic processing by readers.

We distinguish the following types of information distortion:

1. **Style** (distortion severity 1)
   Source snippet: *In order to facilitate knowledge transfer between specialists and generalists and between experts and novices, and to promote interdisciplinary communication, there is a need to provide methods and tools for doing so.*
   Simplified snippet: *There is a need to provide methods and tools for doing so. In order to facilitate knowledge transfer between specialists and generalists and between experts and novices, we need to promote interdisciplinary communication. We need to make it easier for people to share their knowledge with each other.*
   Comment: Deviations from the style norms do not lead to information loss; however, they diminish the quality of text structure and affect readers' assessment of the text and its content.
2. **Insertion of unnecessary details with regard to a query**(distortion severity 1)
   Source snippet: *In the clinical setting, availability of needed information can be crucial during the decision-making process.*
   Simplified snippet: *availability of needed information can be crucial during the decision-making process. In the clinical setting, needed information is often difficult to come by. For confidential support call the Samaritans on 08457 90 90 90 or visit a local Samaritans branch, see www.samaritans.org for details*
   Comment: The simplified snippet often contains more information than the source since the terminology is needed to be explained. An irrelevant insertion does not lead to the loss of information; however, it may bring in diminishing of the text readability and generate discomfort during text perception. The irrelevant reference to the support in the simplified snippet does not clarify the source. The source does not need any simplification.

3. **Redundancy** (without lexical overlap) (distortion severity 2)
   Source snippet: *The capability to get updated information and news is an important and decisive factor in business and finance.*
   Simplified snippet: *The capability to get updated information and news is an important and decisive factor in business and finance. The ability to get updates on the latest news is also an important factor in the success of a business or finance company. For more information, visit CNN.com/News.*
   Comment: Irrelevant insertions, iterations and wordy statements in the extended simplified snippet we consider as a misrepresentation or distortion of source information when a reader may misunderstand source content due to wordiness of the loosely structured simplified snippet.
4. **Insertion of false or unsupported information** (distortion severity 3)
   Source snippet: *The proposed method leads to not only faster running time but also efficient text localization.*
   Simplified snippet: *The proposed method leads to not only faster running time but also efficient text localization. The proposed method is based on the idea that text should be written in a single block of text, rather than a series of blocks of text. The method is being developed by the University of California, San Diego*
   Comment: False and unsupported information is inserted in the simplified snippet because of external links of the source snippet to the open Web resources. False information confuses the readership, thus it is inappropriate in simplified texts.
5. **Omission of essential details with regard to a query** (distortion severity 4)
   Source snippet: *In short, this thesis aims to repatriate young people's web practices from the sterile, positivist methods space of questionnaires and tests of digital literacy to social contexts of everyday life.*
   Simplified snippet: *In short, this thesis aims to repatriate young people's web practices from the sterile, and tests of digital literacy.*
   Comment: Simplified texts regularly omit details and minor concepts since the objective of the simplification is to facilitate perception and processing of the main idea and novel information of the source. Therefore, distinguishing the essential details and concepts from the minor ones plays a crucial role in the text simplification. Omission of the essential details leads to the lack of information and novelty of the message. The simplified snippet lacks the essential information about the direction where the young people's web practice needs to be repatriated (*to social contexts of everyday life*).
6. **Overgeneralization** (distortion severity 5)
   Source snippet: *Online Social Networks explode with activity whenever a crisis event takes place.*
   Simplified snippet: *It explodes when a crisis event takes place.*
   Comment: In the simplified snippet the subject of the source snippet *Online Social Networks* is omitted being substituted by the pronoun *It*; also the essential detail *with activity* is omitted that brings in overgeneralization: the statement refers to all cases of the dysfunctions (server, browser, users' activity, moderators' restrictions, etc.)
7. **Oversimplification** (distortion severity 5)

Source snippet: *If we accept the current orthodoxy and then blame the Web we offer a technological determinist explanation of reality: technology produces misinformed populations.*
Simplified snippet: *If we accept the current orthodoxy and then blame the Web we have an explanation of reality.*
Comment: Oversimplification appears in the shortened simplified snippets when source utterance is transformed into a trivial statement or even a claim. The simplified snippet claims that *we have an explanation of reality* thanks to acceptance of the current orthodoxy and accusation of the Web. Meanwhile, the source snippet discusses the technological determinist explanation of reality. The omission of the essential details leads to the oversimplified statement that cannot explain the reality of the technological epoch.

8. **Topic shift** (distortion severity 5)
Source snippet: *global warming induced by chemtrails or the link between vaccines and autism – find on the Web a natural medium for their dissemination.*
Simplified snippet: *The link between vaccines and autism – can be found on the Web a natural medium for changing.*
Comment: Topic shift is revealed in a substitution of the source topic by omitting its part or selecting a wrong basic word to replace the peculiar term in the source. The source snippet lost the essential part of its topic (*global warming induced by chemtrails*) during the simplification process; moreover, the simplification resulted in the inappropriate syntactic structure of the snippet.
9. **Contra sense / contradiction** (distortion severity 6)
Source snippet: *In this paper we discuss architectural design issues and trade-offs in connection with our experiences porting our agent-based platform, Opal, to the Sharp Zaurus personal digital assistant (PDA).*
Simplified snippet: *The Sharp Zaurus is a personal digital assistant (PDA) developed by Sharp. It is based on the Opal agent-based platform. We discuss architectural design issues and trade-offs in connection with our experiences porting Opal to the Zaurus PDA.*
Comment: Contradictions in simplified snippets appear due to elimination of essential concepts or interrelations among concepts, omission of significant details, and transformation of the semantic structure of the source snippet. The simplified snippet mentions agent-based platform Opal as the basis for the Sharp Zaurus, but at the same time claims that Opal was ported to the Sharp Zaurus. The source snippet *But the new phenomena, the non-agenda ownership, overcome any ideological influence, especially under the conditions of punishment mechanism applied to old politicians* lost its semantic structure since the concepts *ideological influence* and *punishment mechanism* were eliminated in the process of its simplification. Thus, the simplified snippet *But the new phenomena, the ownership of the non-agenda, had a lot of influence on old politicians* lacks any explanation how *the non-agenda ownership* is related to *old politicians* and why they are influence by *the new phenomena.*
10. **Ambiguity** (distortion severity 6)
Source snippet: *The experimental results show that 3D maps with texture on mobile phone display size, and 3D maps without texture on PDA display size are superior to 2D maps in search time and error rate.*

Simplified snippet: *3D maps with texture on mobile phone display size are superior to 2D maps in search time and error rate. The experimental results show that 3D maps without texture on PDA display size were superior to those with texture. The results were published in the journal 3D Maps.*
Comment: Ambiguity presupposes that a statement has several equiprobable interpretations. The instance of the ambiguous simplified snippet in Table X lacks a key to understand whether the 3D maps without texture outperform those with texture or not. Ambiguity often appears due to syntactic simplification of the source. In the source, the clause *changes in the strength of competition also reveal key asymmetrical differences* is replaced by shorter clause *but they do not have any biases* that produces ambiguity: whether *evidence* corresponds to reality or not. The source clarifies the differences between two political parties: *Though both Republicans and Democrats show evidence of implicit biases, changes in the strength of competition also reveal key asymmetrical differences* however, the simplified snippet doubts the reliability of the evidence: *Both Republicans and Democrats show evidence of biases, but they do not have any biases.* Readers of the simplified snippet are unable to resolve the ambiguity.

11. **Nonsense** (distortion severity 7)
Source snippet: *The large availability of user provided contents on online social media facilitates people aggregation around shared beliefs, interests, worldviews and narratives*
Simplified snippet: *The large amount of user provided contents on online social media is called aggregation*
Comment: The source snippet was transformed into a simple sentence. The transformation brings in erroneous usage of the word aggregation that leads to the loss of meaning of the whole sentence. Instead of the original statement about accessibility of the social or public media on the Web, which facilitates dissemination of fake news and rumors, the simplified snippet claims that there is an opportunity to find a resource to read about fake news and rumors.

The final ranking for Task 3 was done by the average harmonic mean of normalized opposite values of *Lexical Complexity (LC), Syntactic Complexity (SC)* and *Distortion Level (DL)* as follows:

$$s_i = \frac{3}{\frac{7}{7-LC} + \frac{7}{7-SC} + \frac{7}{7-DL}} \tag{1}$$

$$Score = \frac{\sum_i \begin{cases} s_i, & \text{if } \textit{No Error} \\ 0, & \text{otherwise} \end{cases}}{n} \tag{2}$$

In Equation 2, variable n refers to the total number of judged snippets and *No Error* means that the snippet i does not have any of *Uncorrect syntax, Unresolved anaphora*, nor *Unnecessary repetition/iteration* error.

3. SimpleText Task 3 Results

In this section we discuss the results for the official submissions to the Task 3.

Table 3
SimpleText Task 3: General results of official runs

Run	Total	Unchanged	Truncated	Valid	Longer	Length Ratio	Evaluated	Uncorrect Syntax	Unresolved Anaphora	Minors	Syntax Complexity	Lexical Complexity	Information Loss
CLARA-HD	116,763	128	2,292	111,627	201	0.61	851	28	3	68	2.10	2.42	3.84
CYUT Team2	116,763	549	101,104	111,818	49	0.81	126	1		32	2.25	2.30	2.26
PortLinguE_full	116,763	42,189	852	111,589	3,217	0.92	564	7		5	2.94	3.06	1.50
PortLinguE_run1	1,000	359	7	970	30	0.93	80	1			3.63	3.57	2.27
lea_task3_t5	23,360	52	23,201	22,062	24	0.35	.	.	.	.	.	.	.
HULAT-UC3M01	1,000	.	13	973	968	2.46	95	10	1	20	4.69	3.69	2.20
HULAT-UC3M02	2,001	3	58	1,960	1,920	2.53	205	10	1	37	3.60	3.53	2.34
HULAT-UC3M03	1,000	2	13	958	966	2.53	.	.	.	.	.	.	.
HULAT-UC3M04	2,000	.	33	1,827	1,957	37	.	.	.	.	.	.	.
HULAT-UC3M05	2,000	.	56	1,921	1,918	2.38	.	.	.	.	.	.	.
HULAT-UC3M06	2,000	.	47	1,976	1,921	2.45	.	.	.	.	.	.	.
HULAT-UC3M07	1,000	.	56	970	972	2.43	.	.	.	.	.	.	.
HULAT-UC3M08	2,000	.	62	1,964	1,919	2.59	.	.	.	.	.	.	.
HULAT-UC3M09	2,000	.	170	1,964	1,904	2.15	.	.	.	.	.	.	.
HULAT-UC3M10	2,000	.	215	1,963	1,910	2.13	.	.	.	.	.	.	.

A total of 5 different teams submitted 14 runs (5 runs were updated). Absolute number of errors and average *Lexical Complexity, Syntax Complexity* and *Information Loss* are provided in Tables 3 and 4. The final ranking for Task 3 is given in Table 5. We removed all runs with the 0 score.

Very interesting partial runs were provided by the HULAT-UC3M team as the generated simplifications provided the explanations of difficult terms. However, HULAT-UC3M's 8 runs over 10 were not in the pool with selected topics. Thus, we provided only automatic evaluation results. The HULAT-UC3M's runs provide clear evidence of the interconnection of tasks 2 and 3.

4. Conclusion

This paper presented the overview of the CLEF 2022 SimpleText Task 3 on simplifying sentences in scientific abstracts, retrieved in response to a queries based on popular science articles.

We created a corpus of sentences extracted from the abstracts of scientific publications. In contrast to previous work, we evaluate simplification in terms of lexical and syntax complexity combining with error analysis. We introduced a new classification of information distortion types for automatic simplification and we annotated the collected simplifications according to this error classification. Recent pandemics have shown that simplification can be modulated by

Table 4
SimpleText Task 3: Information distortion in evaluated runs

Run	Evaluated	Non-Sense	Contresens	Topic Shift	Wrong Synonym	Ambiguity	Omission Of Essential Details	Overgeneralization	Oversimplification	Unsupported Information	Unnecessary Details	Redundancy	Style
CLARA-HD	851	162	68	37	20	80	314	59	203	26	10	29	13
CYUT Team2	126	2	1	.	.	4	42	4	5	.	.	.	4
PortLinguE_full	564	9	3	4	3	19	94	9	13	2	2	5	1
PortLinguE_run1	80	.	.	1	.	.	27	5	2	.	.	.	.
lea_task3_t5	.	.	.	.	.	.	.	.	.	.	.	.	.
HULAT-UC3M01	95	1	7	2	.	5	2	.	1	5	38	36	.
HULAT-UC3M02	205	4	9	4	.	9	4	.	.	12	72	61	1

Table 5
SimpleText Task 3: Ranking of official submissions on combined score

Run	Score
PortLinguE_full	0.149
CYUT Team2	0.122
CLARA-HD	0.119

political needs and the scientific information can be distorted. Thus, in contrast to previous work, we evaluated the simplifications in terms of information distortion.

For next year, we plan to continue the Task 3 setup, continuing the detailed manual annotations of samples, but also working on automatic metrics that best reflect the insights of this year's analysis. This year, the HULAT-UC3M team submitted runs which combine tasks 2 and 3 which demonstrates strong interconnection of the tasks as often the terminology cannot be removed nor simplified but it needs to be explained to a reader.

Acknowledgments

We like to acknowledge the support of the Lab Chairs of CLEF 2022, Allan Hanbury and Martin Potthast, for their help and patience. Special thanks to the University Translation Office of the Université de Bretagne Occidentale, and to Nicolas Poinsu and Ludivine Grégoire for their major impact in the train data construction and Léa Talec-Bernard and Julien Boccou for their help in evaluation of participants' runs. We thank Josiane Mothe for reviewing papers. We also thank Alain Kerhervé, and the MaDICS (https://www.madics.fr/ateliers/simpletext/ research group.

References

[1] M. Maddela, F. Alva-Manchego, W. Xu, Controllable Text Simplification with Explicit Paraphrasing (2021). URL: http://arxiv.org/abs/2010.11004.

[2] L. Ermakova, P. Bellot, P. Braslavski, J. Kamps, J. Mothe, D. Nurbakova, I. Ovchinnikova, E. Sanjuan, Text Simplification for Scientific Information Access: CLEF 2021 SimpleText Workshop, in: Advances in Information Retrieval - 43nd European Conference on IR Research, ECIR 2021, Lucca, Italy, March 28 – April 1, 2021, Proc., Lucca, Italy, 2021.

[3] E. SanJuan, S. Huet, J. Kamps, L. Ermakova, Overview of the CLEF 2022 SimpleText Task 1: Passage selection for a simplified summary, in: [19], 2022.

[4] L. Ermakova, I. Ovchinnikova, J. Kamps, D. Nurbakova, S. Araújo, R. Hannachi, Overview of the CLEF 2022 SimpleText Task 2: Complexity spotting in scientific abstracts, in: [19], 2022.

[5] L. Ermakova, E. SanJuan, J. Kamps, S. Huet, I. Ovchinnikova, D. Nurbakova, S. Araújo, R. Hannachi, É. Mathurin, P. Bellot, Overview of the CLEF 2022 SimpleText Lab: Automatic simplification of scientific texts, in: A. Barrón-Cedeño, G. D. S. Martino, M. D. Esposti, F. Sebastiani, C. Macdonald, G. Pasi, A. Hanbury, M. Potthast, G. Faggioli, N. Ferro (Eds.), CLEF'22: Proceedings of the Thirteenth International Conference of the CLEF Association, Lecture Notes in Computer Science, Springer, 2022.

[6] A. Menta, A. Garcia-Serrano, Controllable Sentence Simplification Using Transfer Learning, in: Proceedings of the Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, Bologna, Italy, September 5th - to - 8th, 2022, CEUR Workshop Proceedings, CEUR-WS.org, Bologna, Italy, 2022.

[7] S.-H. Wu, H.-Y. Huang, CYUT Team2 SimpleText Shared Task Report in CLEF-2022, in: Proceedings of the Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, Bologna, Italy, September 5th - to - 8th, 2022, CEUR Workshop Proceedings, CEUR-WS.org, Bologna, Italy, 2022.

[8] A. Rubio, P. Martínez, HULAT-UC3M at SimpleText@CLEF-2022: Scientific text simplification using BART, in: Proceedings of the Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, Bologna, Italy, September 5th - to - 8th, 2022, CEUR Workshop Proceedings, CEUR-WS.org, Bologna, Italy, 2022.

[9] T.-B. Talec-Bernard, Is Using an AI to Simplify a Scientific Text Really Worth It?, in: Proceedings of the Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, Bologna, Italy, September 5th - to - 8th, 2022, CEUR Workshop Proceedings, CEUR-WS.org, Bologna, Italy, 2022.

[10] S. Saha, D. Roy, B. Y. Goud, C. S. Reddy, T. Basu, NLP-IISERB@Simpletext2022: To Explore the Performance of BM25 and Transformer Based Frameworks for Automatic Simplification of Scientific Texts, in: Proceedings of the Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, Bologna, Italy, September 5th - to - 8th, 2022, CEUR Workshop Proceedings, CEUR-WS.org, Bologna, Italy, 2022.

[11] J. Monteiro, M. Aguiar, S. Araújo, Using a Pre-trained SimpleT5 Model for Text Simplification in a limited Corpus, in: Proceedings of the Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, Bologna, Italy, September 5th - to - 8th, 2022, CEUR Workshop Proceedings, CEUR-WS.org, Bologna, Italy, 2022.

[12] H. Jianfei, M. Jin, Assembly Models for SimpleText Task 2: Results from Wuhan University Research Group, in: Proceedings of the Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, Bologna, Italy, September 5th - to - 8th, 2022, CEUR Workshop Proceedings, CEUR-WS.org, Bologna, Italy, 2022.

[13] F. Mostert, A. Sampatsing, M. Spronk, J. Kamps, University of Amsterdam at the CLEF 2022 SimpleText Track, in: Proceedings of the Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, Bologna, Italy, September 5th - to - 8th, 2022, CEUR Workshop Proceedings, CEUR-WS.org, Bologna, Italy, 2022.

[14] L. Ermakova, P. Bellot, P. Braslavski, J. Kamps, J. Mothe, D. Nurbakova, I. Ovchinnikova, E. SanJuan, Overview of SimpleText 2021 - CLEF Workshop on Text Simplification for Scientific Information Access, in: K. S. Candan, B. Ionescu, L. Goeuriot, B. Larsen, H. Müller, A. Joly, M. Maistro, F. Piroi, G. Faggioli, N. Ferro (Eds.), Experimental IR Meets Multilinguality, Multimodality, and Interaction, Lecture Notes in Computer Science, Springer International Publishing, Cham, 2021, pp. 432–449.

[15] L. Ermakova, P. Bellot, J. Kamps, D. Nurbakova, I. Ovchinnikova, E. SanJuan, E. Mathurin, S. Araújo, R. Hannachi, S. Huet, N. Poinsu, Automatic Simplification of Scientific Texts: SimpleText Lab at CLEF-2022, in: M. Hagen, S. Verberne, C. Macdonald, C. Seifert, K. Balog, K. Nørvåg, V. Setty (Eds.), Advances in Information Retrieval, volume 13186, Springer International Publishing, Cham, 2022, pp. 364–373.

[16] R. Futrell, E. Gibson, H. J. Tily, I. Blank, A. Vishnevetsky, S. T. Piantadosi, E. Fedorenko, The Natural Stories corpus: a reading-time corpus of English texts containing rare syntactic constructions, Language Resources and Evaluation 55 (2021) 63–77. URL: https://doi.org/10.1007/s10579-020-09503-7.

[17] I. Ovchinnikova, Impact of new technologies on the types of translation errors, in: CEUR Workshop Proceedings, 2020.

[18] A. Lommel, A. Görög, A. Melby, H. Uszkoreit, A. Burchardt, M. Popović, Harmonised Metric, Quality Translation 21 (QT21) (2015). URL: https://www.qt21.eu/wp-content/uploads/2015/11/QT21-D3-1.pdf.

[19] G. Faggioli, N. Ferro, A. Hanbury, M. Potthast (Eds.), Proc. of the Working Notes of CLEF 2022: Conference and Labs of the Evaluation Forum, CEUR Workshop Proceedings, 2022.