Pädagogische Hochschule Steiermark

FutureDEAL
Digital Education And Learning

TU Graz

Quality Evaluation of AI-Generated Recommendations for Improving Worksheets in Self-Regulated Learning

Gerald Geier, Ognjen Ilic, Martin Ebner

17th International Conference on Intelligent Multimodal Communication and Learning Technologies
Artificial Intelligence in Education, Learning and Research
12–13 November 2026, UNIC Athens Campus in Elliniko, Athens, Greece



Licensed unter CC BY SA 4.0 International
Except Icons from freepik.com

Pädagogische Hochschule Steiermark

FutureDEAL
Digital Education And Learning

TU Graz

Slides available unter https://www.doi.org/10.3217/ntppd-9h124

Licensed unter CC BY SA 4.0 International
Except Icons from freepik.com

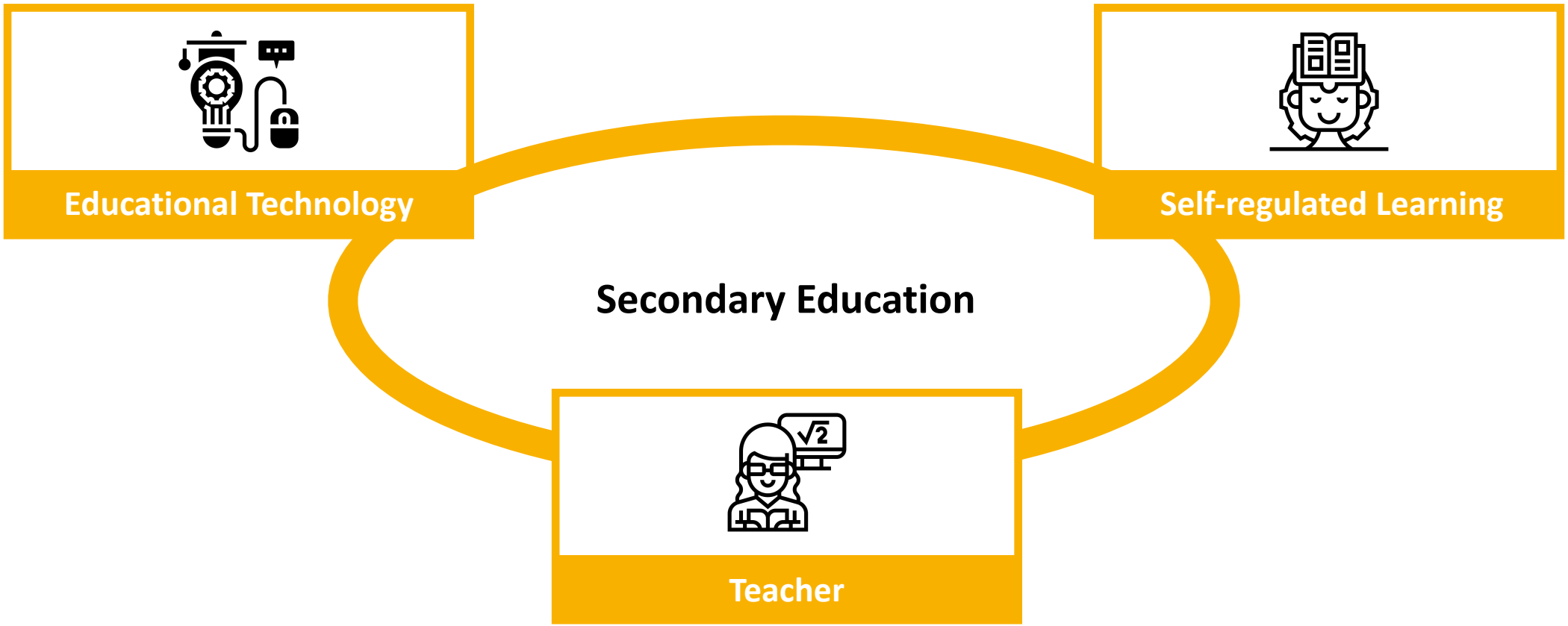


Figure 1: Our research field

Icons by Eucalyp and Artisan on freepik.com

Motivation

- Impact of GenAI on education widely learner-oriented (Jahić et al., 2024)
- Learning materials not designed to support SRL (Del Mario & Tran, 2024; Dignath & Veenman, 2021)
  - learning materials can enhance students SRL competencies (Leidinger & Perels, 2012)
- GenAI can support the development of educational materials (Choi et al., 2024)

SRLly

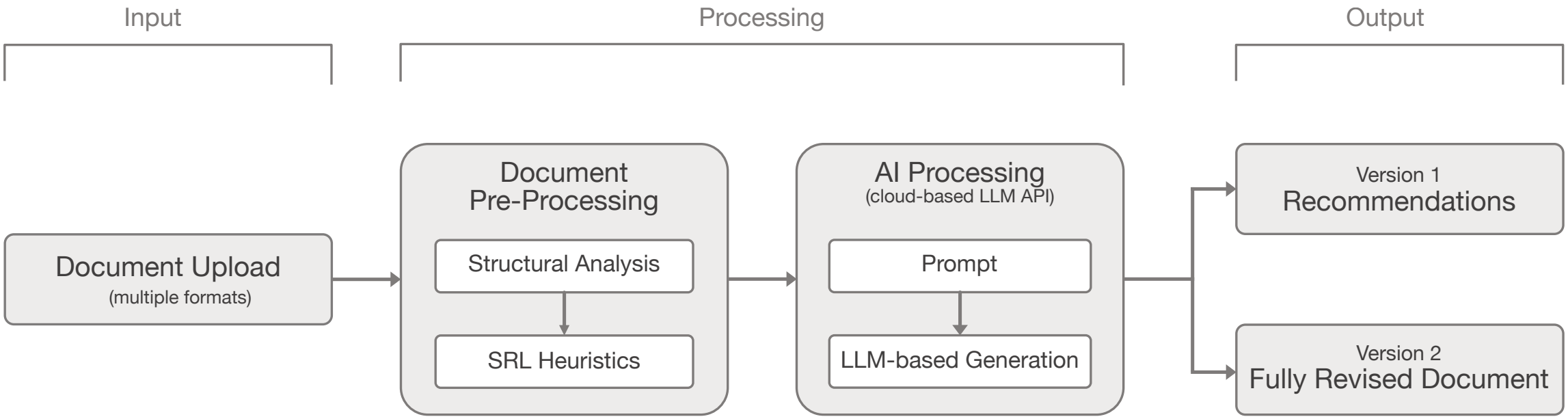


Figure 2: Schematic representation of the core components of SRLly (Geier et al., 2026)

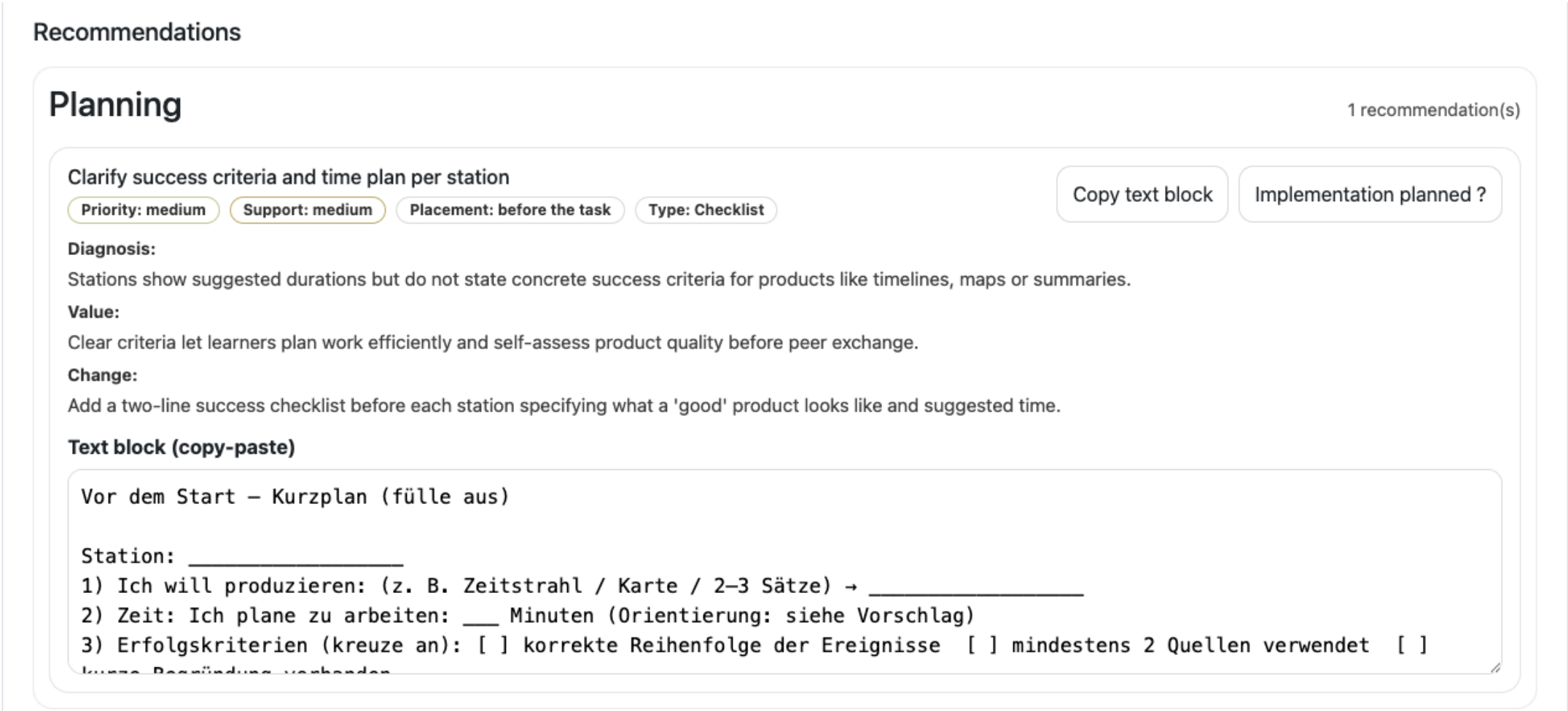


Figure 3: Analysis result view with a sample recommendation

Research Question

- To what extent do SRLly-generated recommendations indicate potential improvements in the SRL quality of worksheets compared to their original versions?

Methodology

- Rubric-based pre-post design

- Criteria-based set of SRL dimensions (rubrics)
  - planning/forethought
  - strategy use/active processing
  - monitoring/self-checking
  - reflection/self-evaluation/feedback
  - motivation/task value/self-efficacy
  - scaffolding/worksheet structure/resource support
  - *collaboration/co-regulation*

Methodology

- Annotation guidelines

Dimension	Score 0	Score 1	Score 2	Score 3	Evidence Examples to look for	Zimmerman phase
1. Planning / Forethought	No planning support is visible. The worksheet only presents content or tasks.	Planning is only implicit, for example a topic title, task instruction, or learning objective written by the teacher without student involvement.	The worksheet includes **one explicit planning element**, such as a learning goal, prior knowledge prompt, task-planning question, or preparation task.	The worksheet includes **several connected planning elements**, for example goal setting + prior knowledge activation + task planning, or guides students through a clear preparation process before starting.	Learning goals; "What do you already know?"; "Which steps will you follow?"; preview/preparation task; time/effort planning.	**Forethought phase**
2. Strategy Use / Active Processing	No active or strategic processing is required. Students only read, copy, or complete very basic tasks.	The worksheet requires some activity, but mainly mechanical work, such as matching, copying, underlining, or filling gaps without deeper processing.	The worksheet includes **one clear active-processing or strategy element**, such as explaining in own words, applying an example, organizing ideas, or choosing a strategy.	The worksheet guides students through **deeper or repeated active processing**, for example selecting a strategy, applying it, explaining reasoning, comparing information, and/or organizing knowledge meaningfully.	"Explain in your own words"; "Which strategy will you use?"; "Apply the example"; "Find and write the correct information"; concept maps; own examples.	**Performance phase**

Table 1: Scoring descriptions from the annotation guidelines

Methodology

- 16 worksheets from different teachers and schools
  - reflect a range of authentic classroom materials
  - most frequent subject → English as a foreign language
- Two raters
  - quadratic weighted Cohen's Kappa
  - low agreement values were resolved through consensus



Methodology

- Mean difference between pre- and post-ratings
  - paired wilcoxon signed-rank tests (asymptotic, two-sided, N = 16)

- Post-Rating
  - not all worksheets were fully revised
  - assumption that all of the SRLly recommendations were implemented
  - expected SRL quality

Present continuous

Die present continuous (auch present progressive genannt) wird verwendet um auszudrücken, was **JETZT GERADE** passiert, z.B.:

Listen! John is playing the guitar!

So wird die Present continuous gebildet: subject + form of *to be* + verb + -ing

I am dancing.	We are dancing.
You are dancing.	You are dancing.
He / She / It is dancing.	They are dancing.



Question (Frage): Bei der Frage werden subject und form of *to be* vertauscht.

Am I dancing?	Are we dancing?
Are you dancing?	Are you dancing?
Is he / she / it dancing?	Are they dancing?

Negation (Verneinung): Bei der Verneinung wird zwischen die form of *to be* und das verb ein „not" eingefügt.

I am not dancing.	We are not dancing.
You are not dancing.	You are not dancing.
He / She / It is not dancing.	They are not dancing.

Spelling rules

→ Konsonant nach kurzem, betontem Vokal am Wortende wird verdoppelt: sit – sitting, put – putting, cut – cutting, stop – stopping, swim – swimming

→ stummes „e" am Wortende: „e" fällt weg: write – writing, dance – dancing, take – taking

→ Verben, die auf „ie" enden: „ie" wird zu y: lie – lying, die – dying, tie – tying

Figure 4: Example of an original worksheet

Pädagogische Hochschule Steiermark

FutureDEAL
Digital Education And Learning

TU Graz

Present continuous

Lernziel: Ich kann Sätze im Present Continuous bilden und anwenden.
Erfolgskriterien (kreuze an):
- ☐ Ich bilde 3 Aussagen korrekt.
- ☐ Ich bilde 2 Fragen korrekt.
- ☐ Ich bilde 2 Verneinungen korrekt.

Wenn du alle Kästchen ankreuzt, hast du das Lernziel erreicht.

Die present continuous (auch present progressive genannt) wird verwendet um auszudrücken, was **JETZT GERADE** passiert, z.B.:

Listen! John is playing the guitar!

So wird die Present continuous gebildet: subject + form of *to be* + verb + -ing

I am dancing.	We are dancing.
You are dancing.	You are dancing.
He / She / It is dancing.	They are dancing.



Question (Frage): Bei der Frage werden subject und form of *to be* vertauscht.

Am I dancing?	Are we dancing?
Are you dancing?	Are you dancing?
Is he / she / it dancing?	Are they dancing?

Negation (Verneinung): Bei der Verneinung wird zwischen die form of *to be* und das verb ein „not" eingefügt.

I am not dancing.	We are not dancing.
You are not dancing.	You are not dancing.
He / She / It is not dancing.	They are not dancing.

So bilde ich Sätze (3 Schritte):
1) Aussage: Subject + richtige Form von 'to be' + Verb+-ing (z. B. "She is dancing").
2) Frage: Tausche Subject und 'to be' (z. B. "Is she dancing?").
3) Verneinung: Setze 'not' zwischen 'to be' und Verb (z. B. "She is not dancing").

Übung: Schreibe je 1 Satz pro Typ zum Thema 'Musik'.

Spelling rules

→ Konsonant nach kurzem, betontem Vokal am Wortende wird verdoppelt: sit – sitting, put – putting, cut – cutting, stop – stopping, swim – swimming

→ stummes „e" am Wortende: „e" fällt weg: write – writing, dance – dancing, take – taking

→ Verben, die auf „ie" enden: „ie" wird zu y: lie – lying, die – dying, tie – tying

Wähle ein Thema (kreuze an) und schreibe 3 Sätze im Present Continuous dazu:
- ☐ Musik (z. B. "John is playing the guitar")
- ☐ Sport (z. B. "They are swimming")
- ☐ Schule/Alltag (z. B. "She is doing her homework")

Wähle eines und schreibe deine Sätze unterhalb.

Selbstkontrolle (prüfe nach jeder Aufgabe):
1. Ist die richtige Form von 'to be' verwendet? (am/are/is)
2. Steht das Verb in der -ing-Form?
3. Bei Fragen: Wurde die Form von 'to be' vor das Subject gestellt?
4. Bei Verneinungen: Steht 'not' zwischen 'to be' und dem Verb?

Korrigiere jeden Satz, bei dem du eine Frage mit 'nein' beantworten würdest.

Mini-Übung (3 Schritte):
1) Lücke füllen: "He ___ (dance)" → Schreibe die richtige Form.
2) Umwandeln: Wandeln Sie den Satz vom Simple Present ins Present Continuous um.
3) Eigenes Beispiel: Schreibe einen eigenen Satz zum Alltagsthema (mind. 1 Satz).
Markiere nach Schritt 1–3 jeweils mit ✓, wenn korrekt.

Figure 5: Manually revised version of the original worksheet

Interrater reliability

Dimension	Pre Weighted Cohen's κ	Agreement (%)	Post Weighted Cohen's κ	Agreement (%)
Planning	0.89	87.5	0.88	93.8
Strategy	0.88	81.2	**0.64**	93.8
Monitoring	**0.65**	62.5	1.00	100.0
Reflection	0.94	87.5	0.87	87.5
Motivation	0.94	87.5	0.83	87.5
Scaffolding	0.85	81.2	-	100.0
Collaboration	0.83	81.2	0.92	81.2

Table 2: Weighted Cohen's Kappa and agreement values for all dimensions in pre- and post-test

Overall SRL score

ID (W)	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16
Pre	5	5	3	4	5	4	8	4	13	14	8	7	6	15	8	17
Post	17	14	15	11	16	15	15	16	17	18	16	15	15	18	18	18
Δ	+12	+9	+12	+7	+11	+11	+7	+12	+4	+4	+8	+8	+9	+3	+10	+1

Table 3: Overall SRL scores by worksheet

Overall SRL score

- Paired wilcoxon signed-rank tests
  - $p < .001$
  - $r = .88$
- Substantial potential improvement in overall SRL quality
  - assumption that the SRLly recommendations are implemented

Changes across SRL dimensions

Dimension	Pre M	Post M	Δ M	Post > Pre	p	p_{holm}	r
Planning	1.38	2.44	+1.06	14/16	0.001	0.004	0.895
Strategy	2.13	2.88	+0.75	10/16	0.004	0.008	0.779
Monitoring	0.69	2.88	+2.19	15/16	0.001	0.004	0.887
Reflection	0.56	2.19	+1.63	12/16	0.002	0.007	0.829
Motivation	1.25	2.50	+1.25	13/16	0.001	0.006	0.857
Scaffolding	1.88	3.00	+1.12	12/16	0.002	0.007	0.833
Collaboration	0.50	0.69	+0.19	3/16	0.149	0.149	0.433

Table 4: Results of pre- and post-tests

Discussion

- SRL-supportive elements are rarely included in worksheets
- Elements to foster monitoring and reflection received the lowest scores
  - central components of metacognitive regulation
- Higher overall scores
  - highest in monitoring and reflection dimension
  - possible ceiling effect

Discussion

- GenAI has considerable potential to support teachers
- Time savings as an additional benefit
  - easy to use for every teacher
- Supportive tool to integrate SRL-promoting elements

Limitations

- General indication rather than a subject-specific evidence
- Limited coverage of the fine-grained SRL challenges
- No blinded raters

SRLly

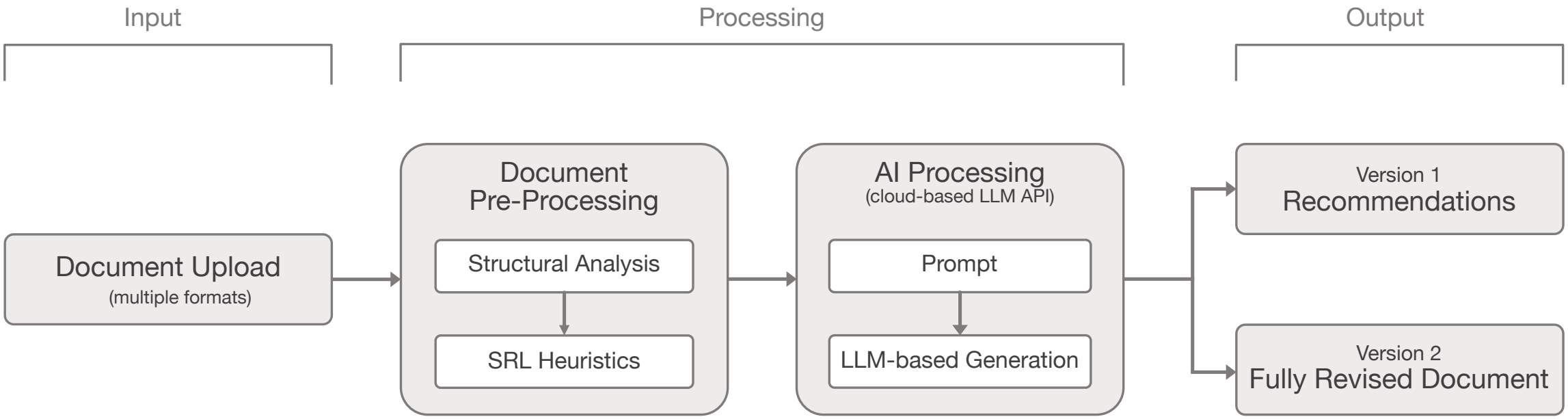


Figure 2: Schematic representation of the core components of SRLly (Geier et al., 2026)



Test SRLly Beta

https://beta.srlly.at/

Pädagogische Hochschule Steiermark

FutureDEAL
Digital Education And Learning

TU Graz

Feedback

SRLly



https://t1p.de/nip1t

References

- Choi, G.W., Kim, S.H., Lee, D., Moon, J.: Utilizing Generative AI for Instructional Design: Exploring Strengths, Weaknesses, Opportunities, and Threats. TechTrends 68(4), 832–844 (2024). https://doi.org/10.1007/s11528-024-00967-w
- Del Mario, J., Tran, H.: A literature review: analyzing barriers hindering the implementation of self-regulated learning in science classrooms. Frontiers in Education 9:1330449 (2024). https://doi.org/10.3389/feduc.2024.1330449
- Dignath, C., Veenman, M.: The Role of Direct Strategy Instruction and Indirect Activation of Self-Regulated Learning - Evidence from Classroom Observation Studies. Educational Psychology Review 33 (2021). https://doi.org/10.1007/s10648-020-09534-0
- Geier, G., Ebner, M., Burgsteiner, H.: Development and Evaluation of an AI-Supported Teacher Assistant Tool (AI-TAT) to Foster Self-Regulated Learning. In: Smith, B.K., Borge, M. (eds.) Learning and Collaboration Technologies. pp. 199–211. Springer Nature Switzerland, Cham (2026). https://doi.org/10.1007/978-3-032-30784-2_12
- Jahić, I., Ebner, M., Schön, S., Edelsbrunner, S.: Exploring the Use of Generative AI in Education: Broadening the Scope. In: Zaphiris, P., Ioannou, A. (eds.) Learning and Collaboration Technologies. pp. 283–304. Springer Nature Switzerland, Cham (2024). https://doi.org/10.1007/978-3-031-61691-4_19
- Leidinger, M., Perels, F.: Training Self-Regulated Learning in the Classroom: Development and Evaluation of Learning Materials to Train Self-Regulated Learning during Regular Mathematics Lessons at Primary School. Education Research International 2012 (2012). https://doi.org/10.1155/2012/735790

Acknowledgements

- This research was done as part of the "FutureDEAL - Future of Digital Education and Learning" initiative within the doctoral program "**Bildungsinnovation braucht Bildungsforschung**", which is supported and partially funded by the Austrian Federal Ministry of Education, Science, and Research.

Pädagogische Hochschule Steiermark

FutureDEAL
Digital Education And Learning

TU Graz

Questions?

Contact

Gerald Geier
gerald.geier@phst.at
ORCID: 0009-0006-6096-7688

Slides available unter https://www.doi.org/10.3217/ntppd-9h124