SILICON SURFER

iu INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

The Learning Ideas Conference

TIM SCHLIPPE & JÖRG SAWATZKI

AI-BASED MULTILINGUAL INTERACTIVE EXAM PREPARATION

New York, USA

June 16, 2021

AGENDA

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

MOTIVATION

MOTIVATION

SUSTAINABLE DEVELOPMENT GOALS

Sources: United Nations: Sustainable Development Goals: 17 Goals to Transform our World (2021); OpenClipart-Vectors/154119/Pixabay.

MOTIVATION

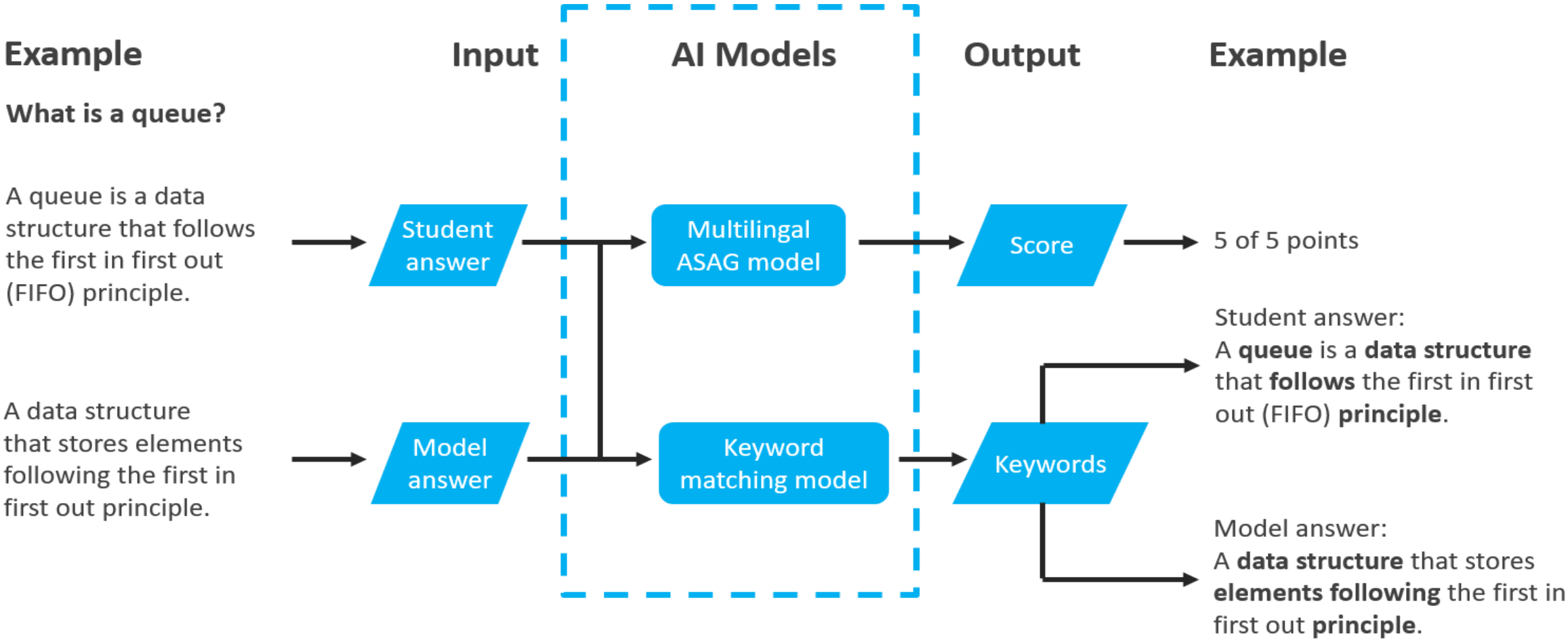
SUSTAINABLE DEVELOPMENT GOALS

Sources: United Nations: Sustainable Development Goals: 17 Goals to Transform our World (2021); OpenClipart-Vectors/154119/Pixabay; Statista: The Most Spoken Languages Worldwide in 2019 (2020).

MOTIVATION

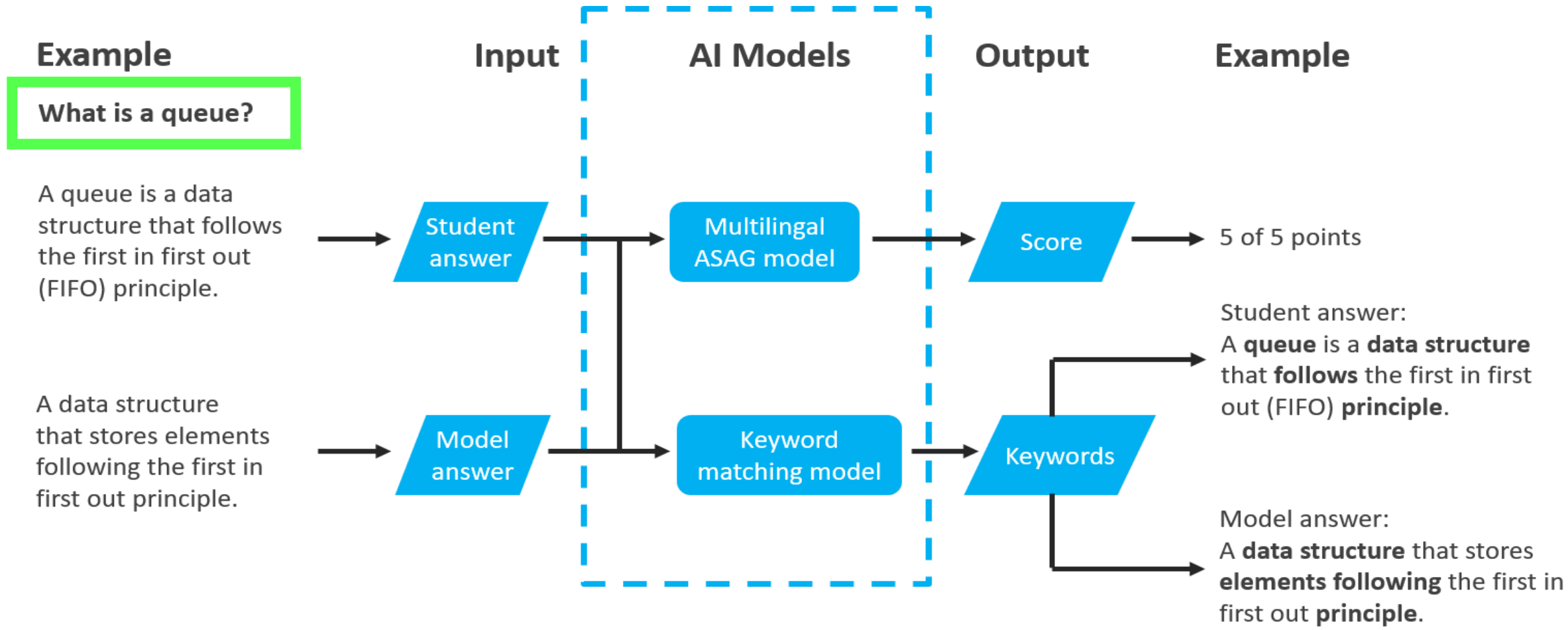
Graphic Source: Custom Depiction.

MOTIVATION



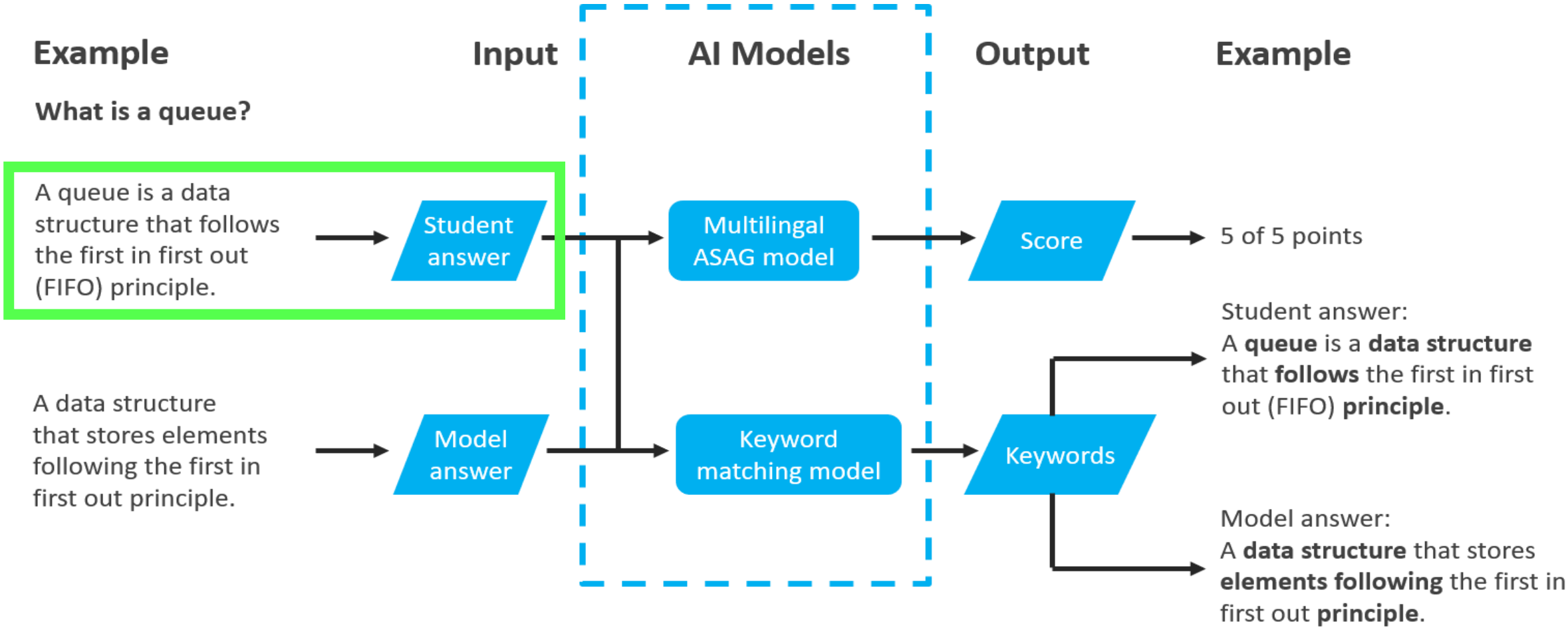
Graphic Source: Schlippe & Sawatzki (2021).

MOTIVATION



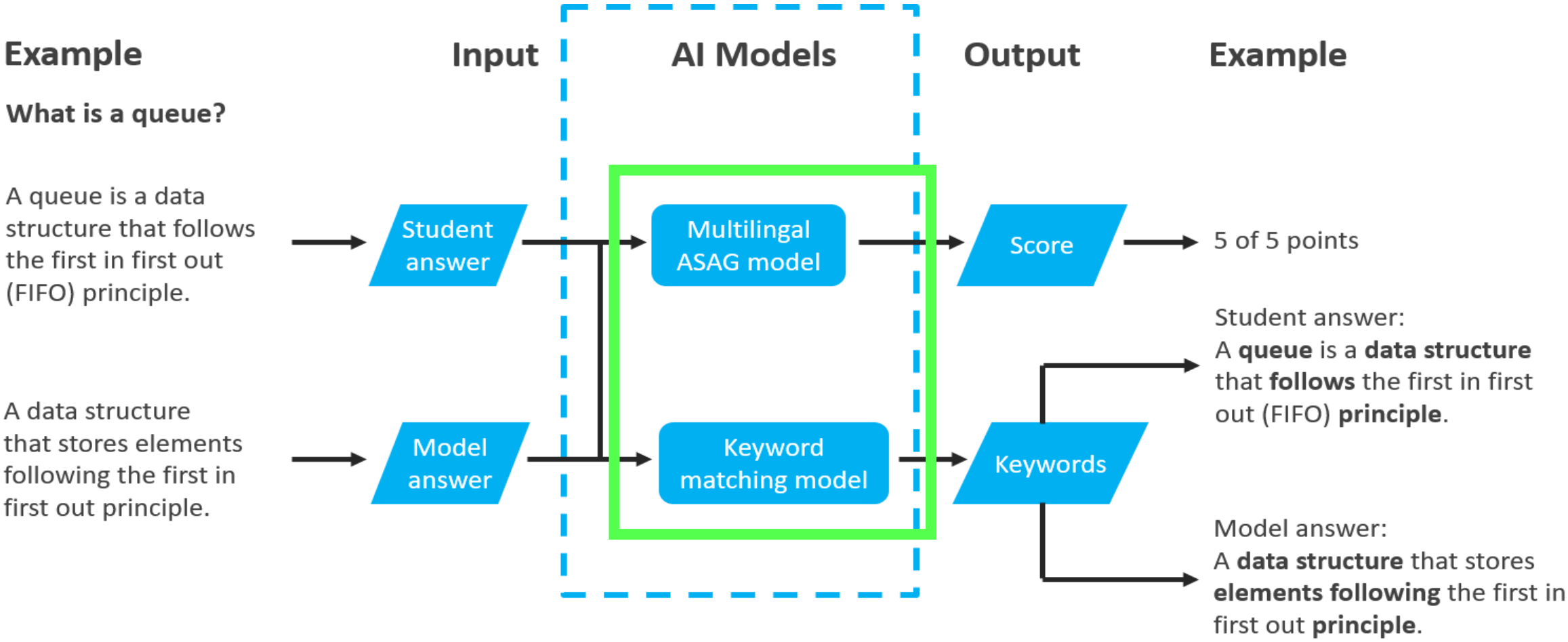
Graphic Source: Schlippe & Sawatzki (2021).

MOTIVATION



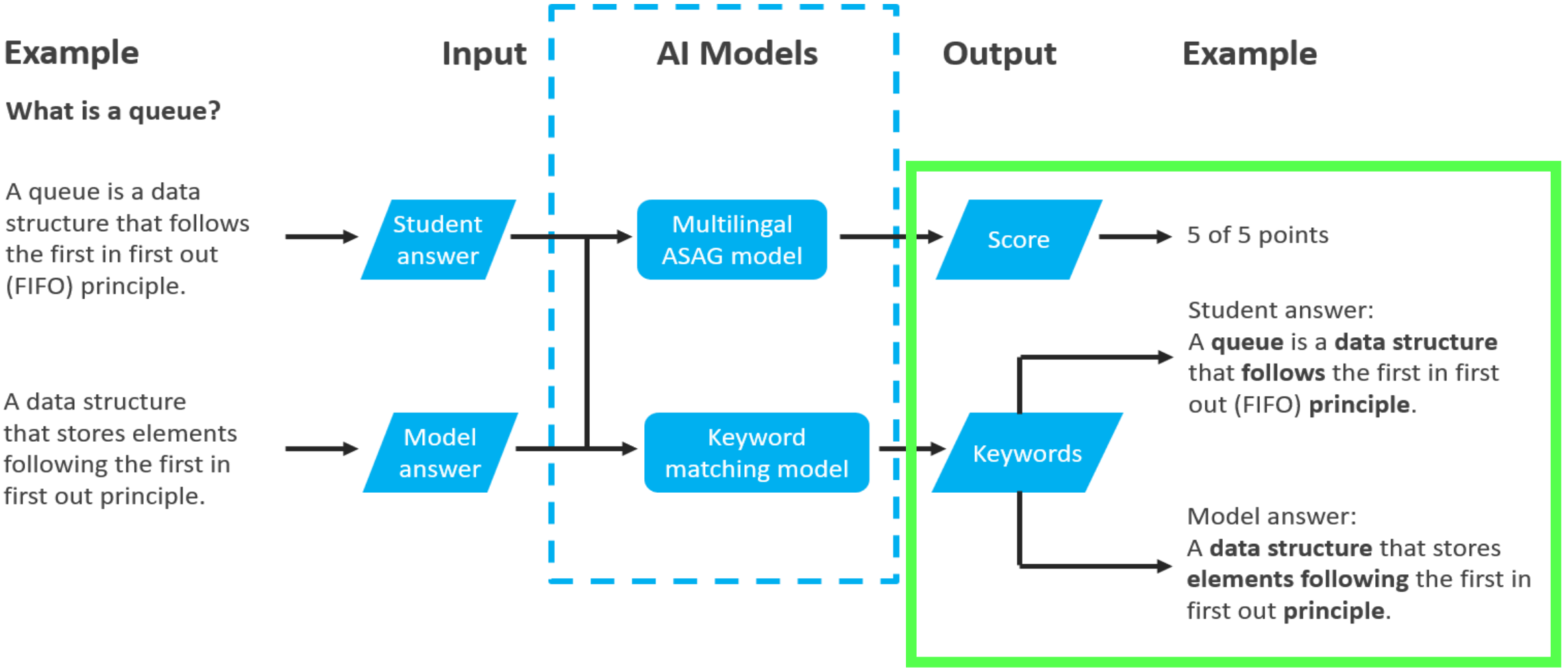
Graphic Source: Schlippe & Sawatzki (2021).

MOTIVATION



Graphic Source: Schlippe & Sawatzki (2021).

MOTIVATION



Graphic Source: Schlippe & Sawatzki (2021).

02

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

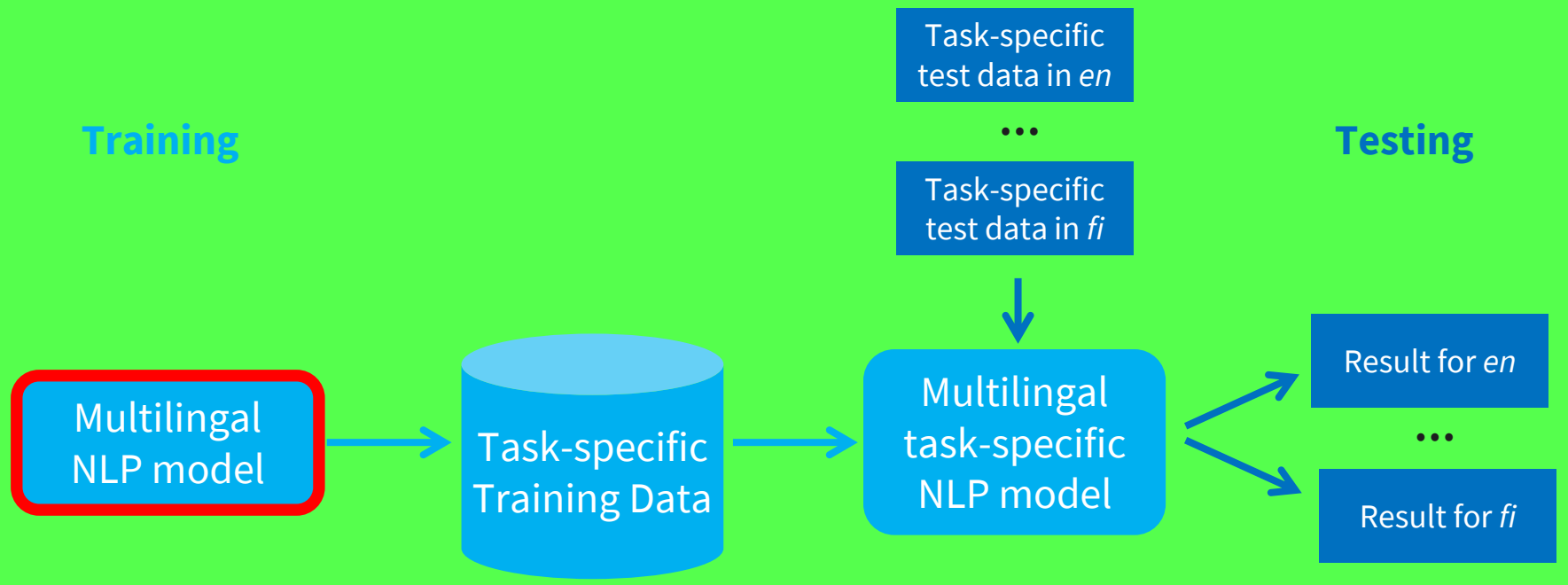
RELATED WORK

RELATED WORK

MULTILINGUAL NLP MODELS

Transfer learning
Cross-lingual transfer

e.g., Multilingual BERT (Devlin et al., 2019; Pires et al., 2019), RoBERTa (Liu et al., 2019), XLM-R (Conneau, 2018)



Graphic Source: Custom Depiction.

RELATED WORK

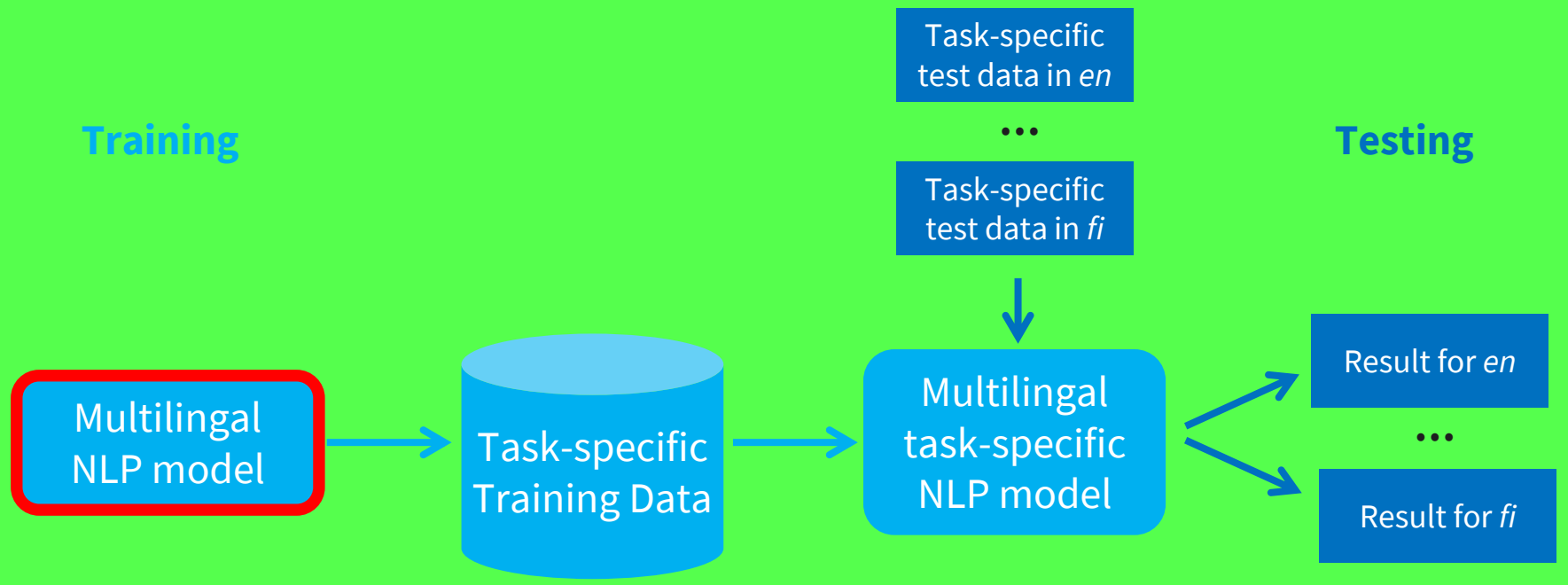
MULTILINGUAL NLP MODELS

Transfer learning
Cross-lingual transfer

e.g., Multilingual BERT

Multilingual BERT

Pires et al., 2019),
RoBERTa (Liu et al., 2019),
XLM-R (Conneau, 2018)



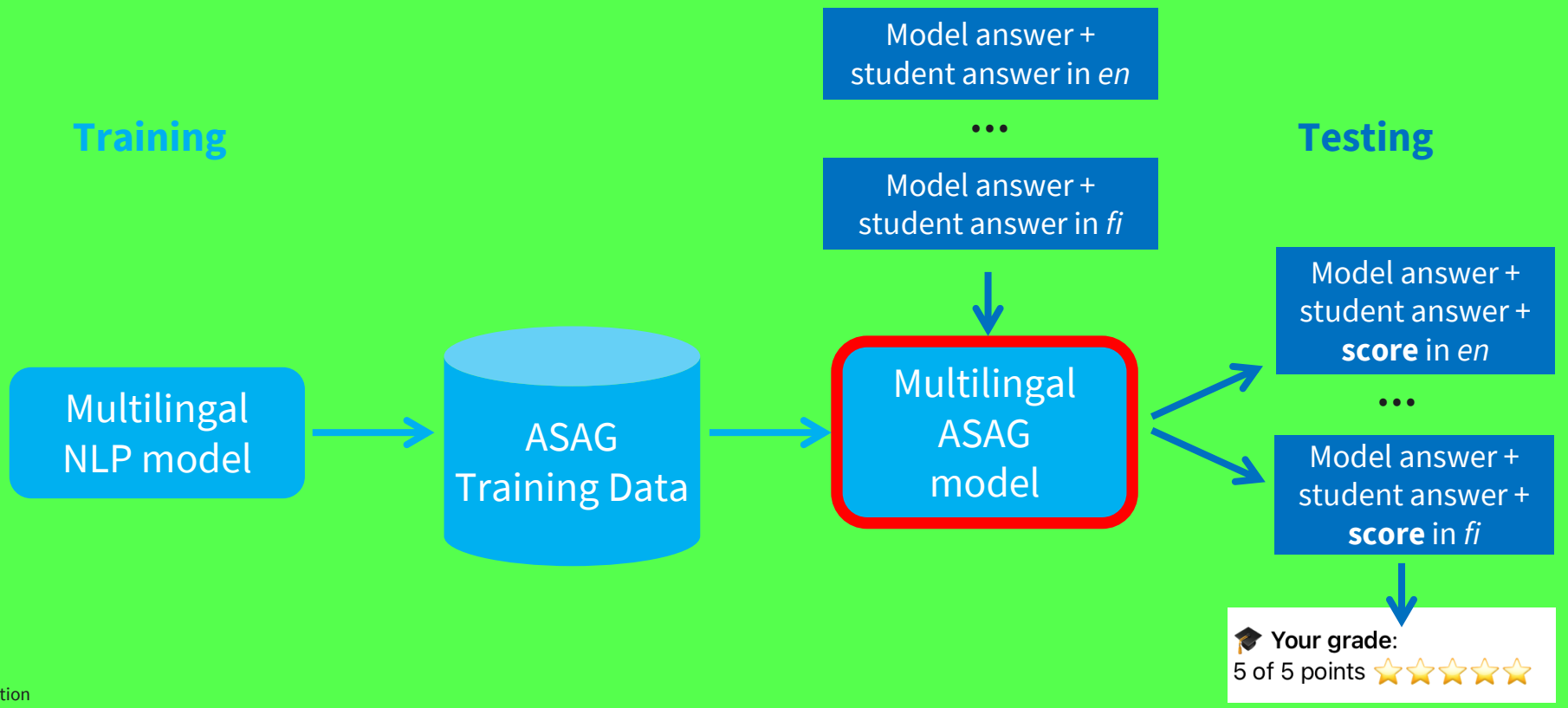
Graphic Source: Custom Depiction.

RELATED WORK

AUTOMATIC
SHORT ANSWER
GRADING

Deep learning

e.g., (Burrows et al., 2014; Camus & Filighera, 2020; Sawatzki et al., 2021; Schlippe & Sawatzki, 2021b)



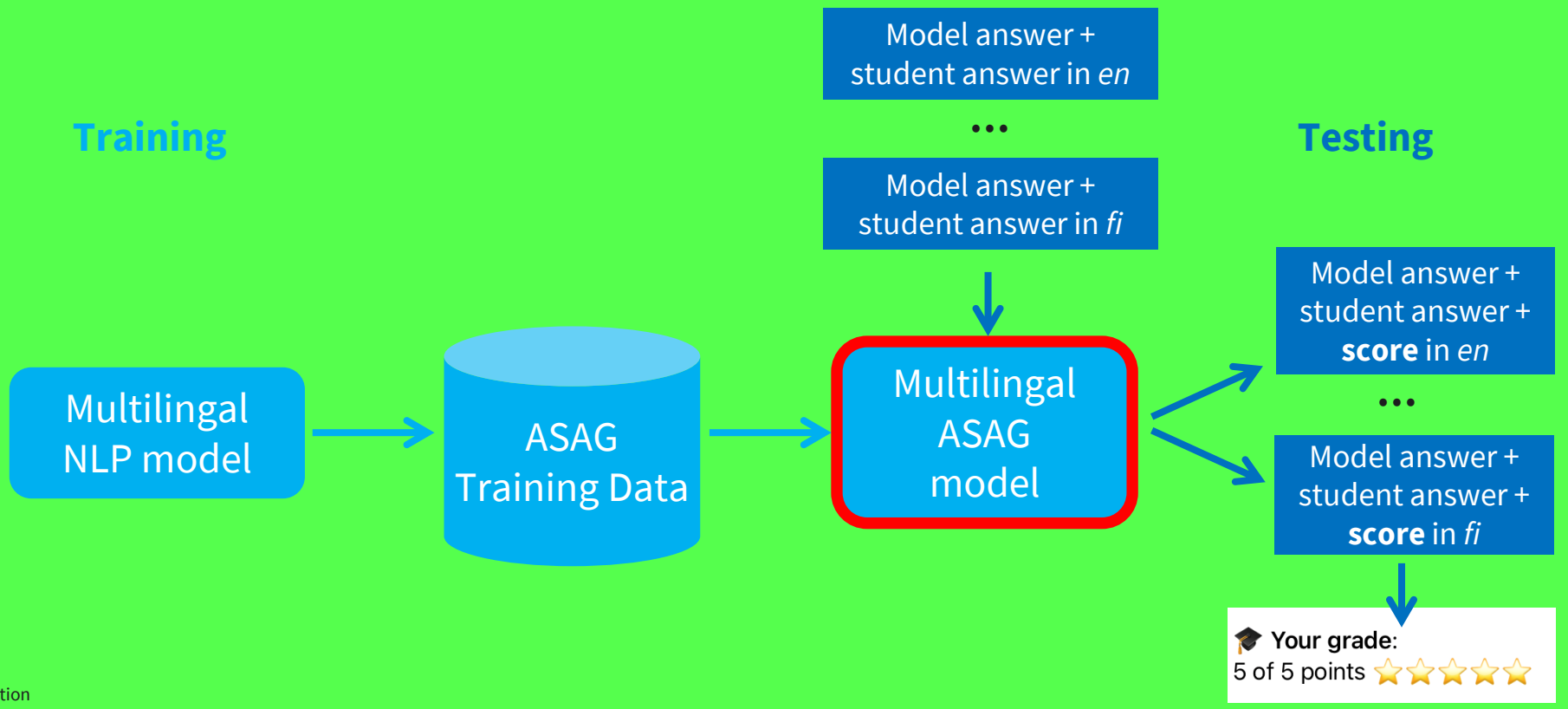
Graphic Source: Custom Depiction.

RELATED WORK

AUTOMATIC
SHORT ANSWER
GRADING

Deep learning

Cross-lingual Automatic Short Answer Grading



Graphic Source: Custom Depiction.

RELATED WORK

CONVERSATIONAL AI

Pedagogical Conversational Agents

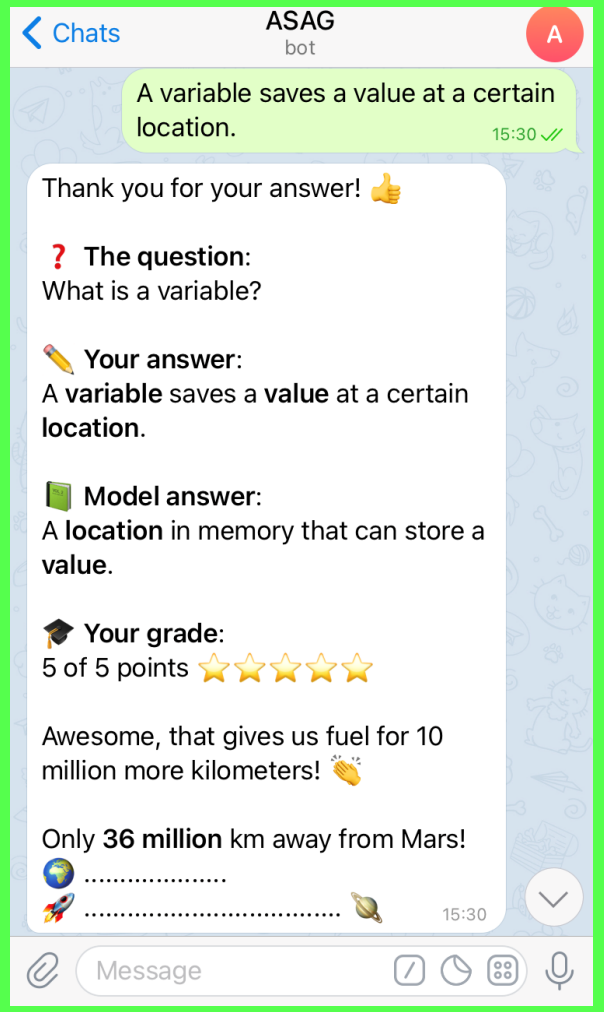
e.g. (Wölfel, 2021)

Frameworks

RASA (Bocklisch et al., 2019),

Dialogflow (Reyes et al., 2019),

Telegram (Setiaji & Paputungan, 2017),



Graphic Source: Custom Depiction.

RELATED WORK

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

CONVERSATIONAL AI

Pedagogical Conversational Agents

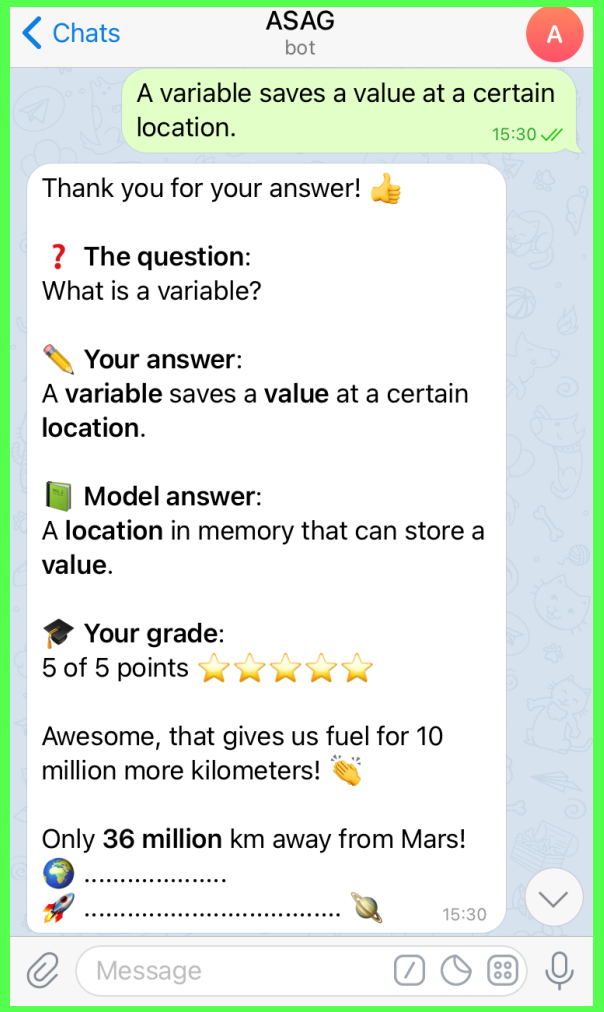
e.g. (Wölfel, 2021)

Frameworks

RASA (Bocklisch et al., 2019),

Dialogf**Telegram**., 2019),

Telegram (Setiaji & Paputungan, 2017),



Graphic Source: Custom Depiction.

RELATED WORK

KEYWORD EXTRACTION

SEMANTIC SIMILARITY

Keyword extraction

e.g., (Hasan, 2014; Merrouni et al., 2020)

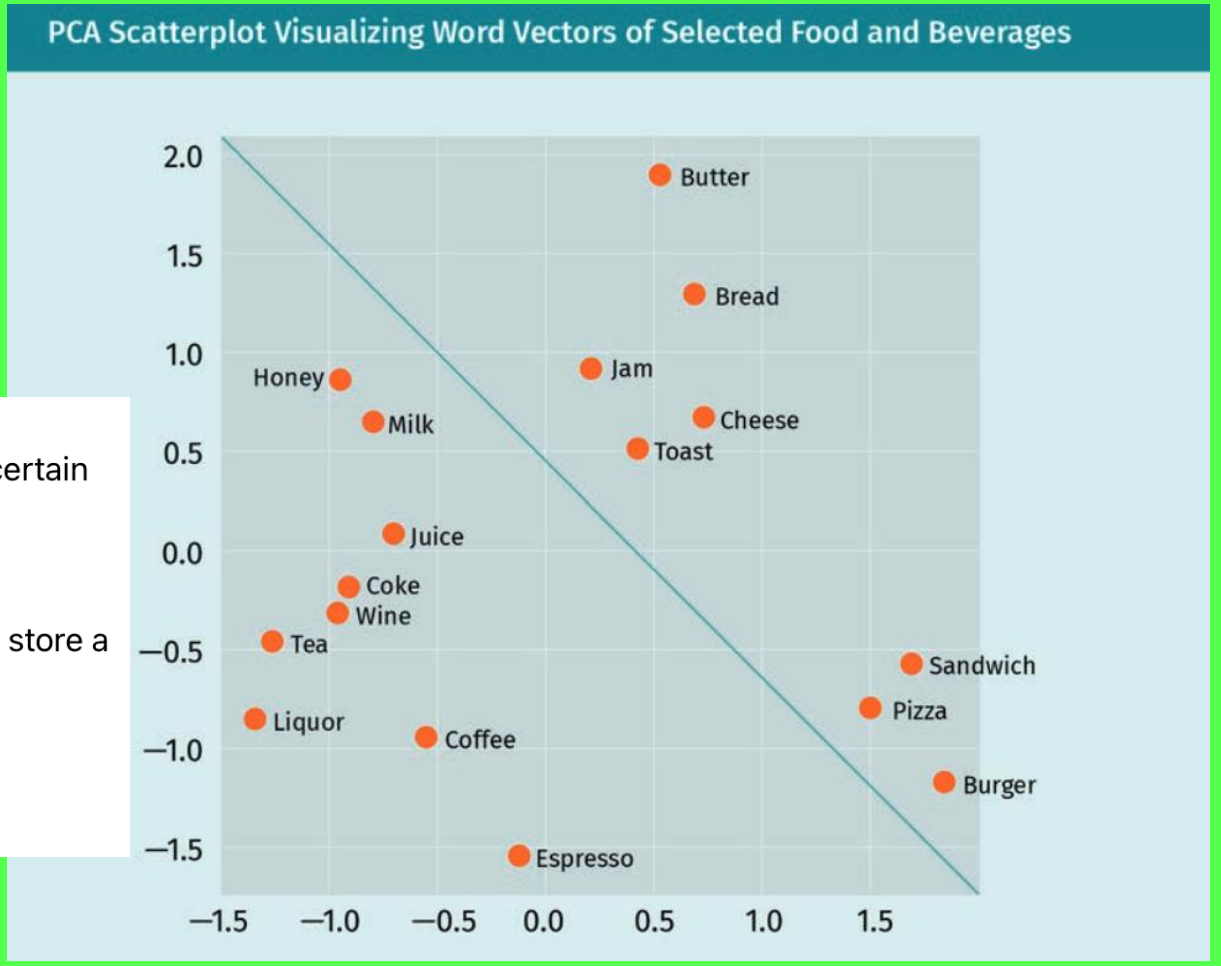
Semantic similarity

e.g., (Chandrasekaran, 2021)

✏️ **Your answer**:
A **variable** saves a **value** at a certain **location**.

📗 **Model answer**:
A **location** in memory that can store a **value**.

🎓 **Your grade**:
5 of 5 points ⭐⭐⭐⭐⭐



Graphic Sources: Custom Depiction; Graphic Source: Coursebook (NLP and Computer Vision), pg. 25.

RELATED WORK

KEYWORD EXTRACTION

SEMANTIC SIMILARITY

Ke
e.g
Me

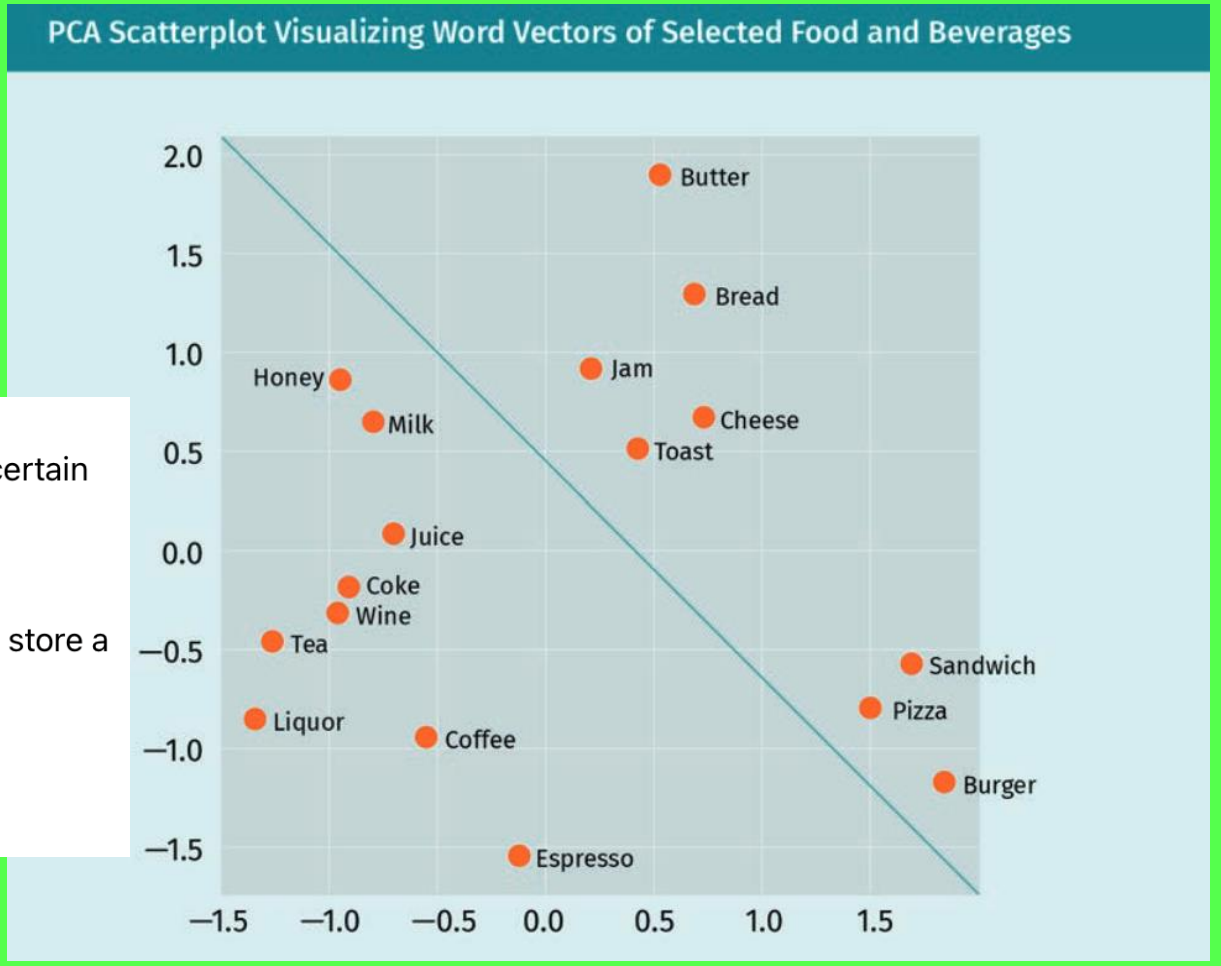
Se
e.g., (Chandrasekaran, 2021)

SpaCy (Honnibal & Montani, n.d.)

✏️ **Your answer**:
A **variable** saves a **value** at a certain **location**.

📗 **Model answer**:
A **location** in memory that can store a **value**.

🎓 **Your grade**:
5 of 5 points ⭐⭐⭐⭐⭐



Graphic Sources: Custom Depiction; Graphic Source: Coursebook (NLP and Computer Vision), pg. 25.

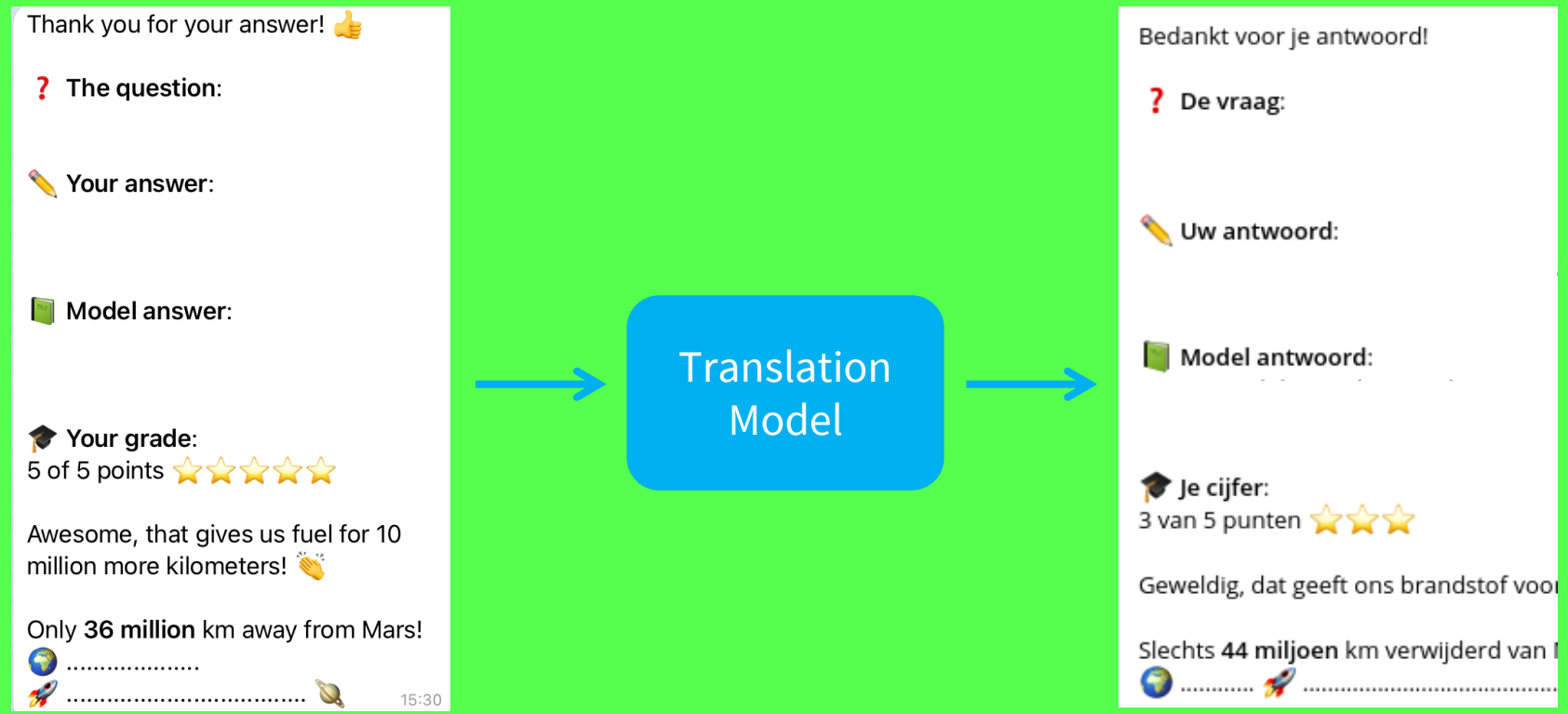
RELATED WORK

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

TRANSLATION MODELS

Google's Neural Machine Translation System

(Wu et al., 2016;
Aiken, 2019)



Graphic Source: Custom Depiction.

RELATED WORK

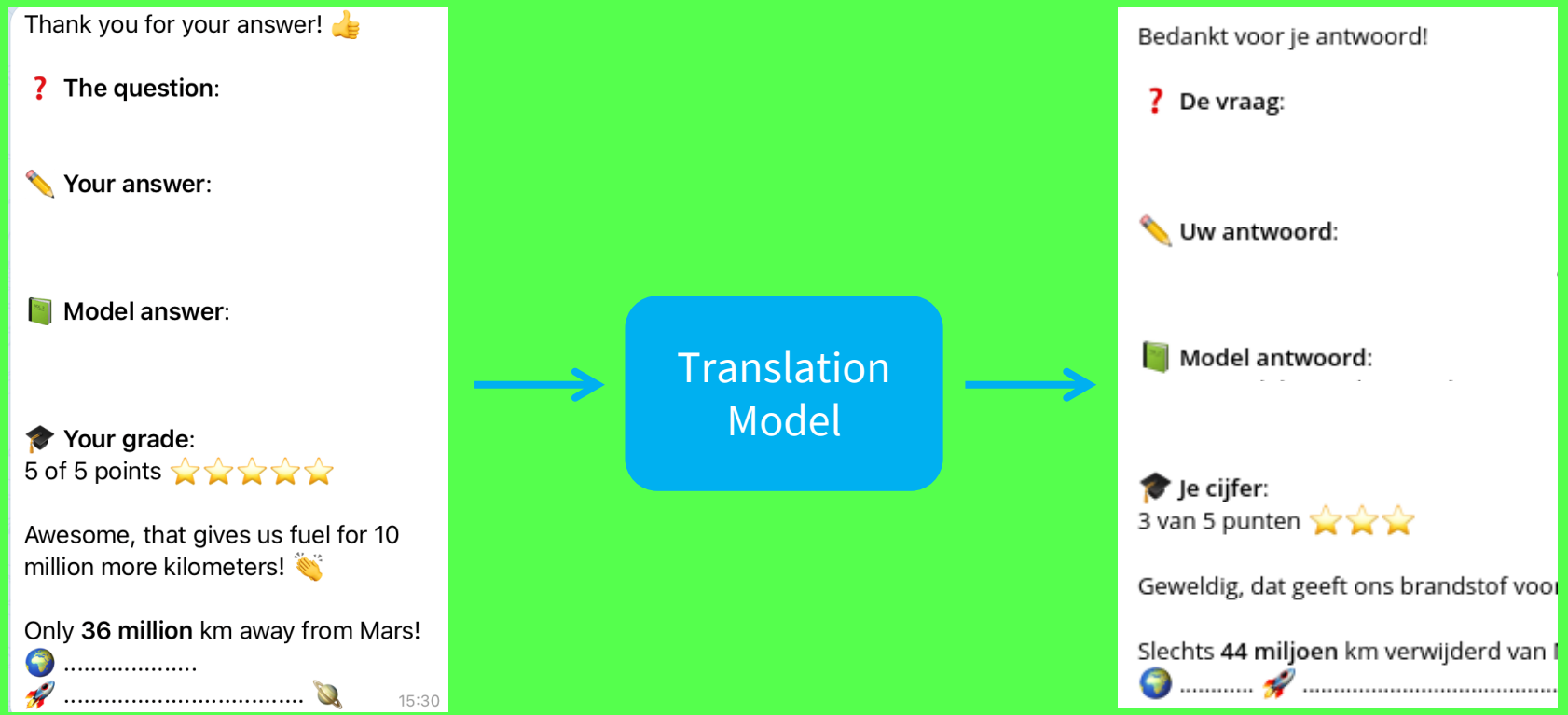
INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

TRANSLATION MODELS

Google's Neural Machine Translation System

(Wu et al., 2016;
Aiken, 2019)

translate.google.com



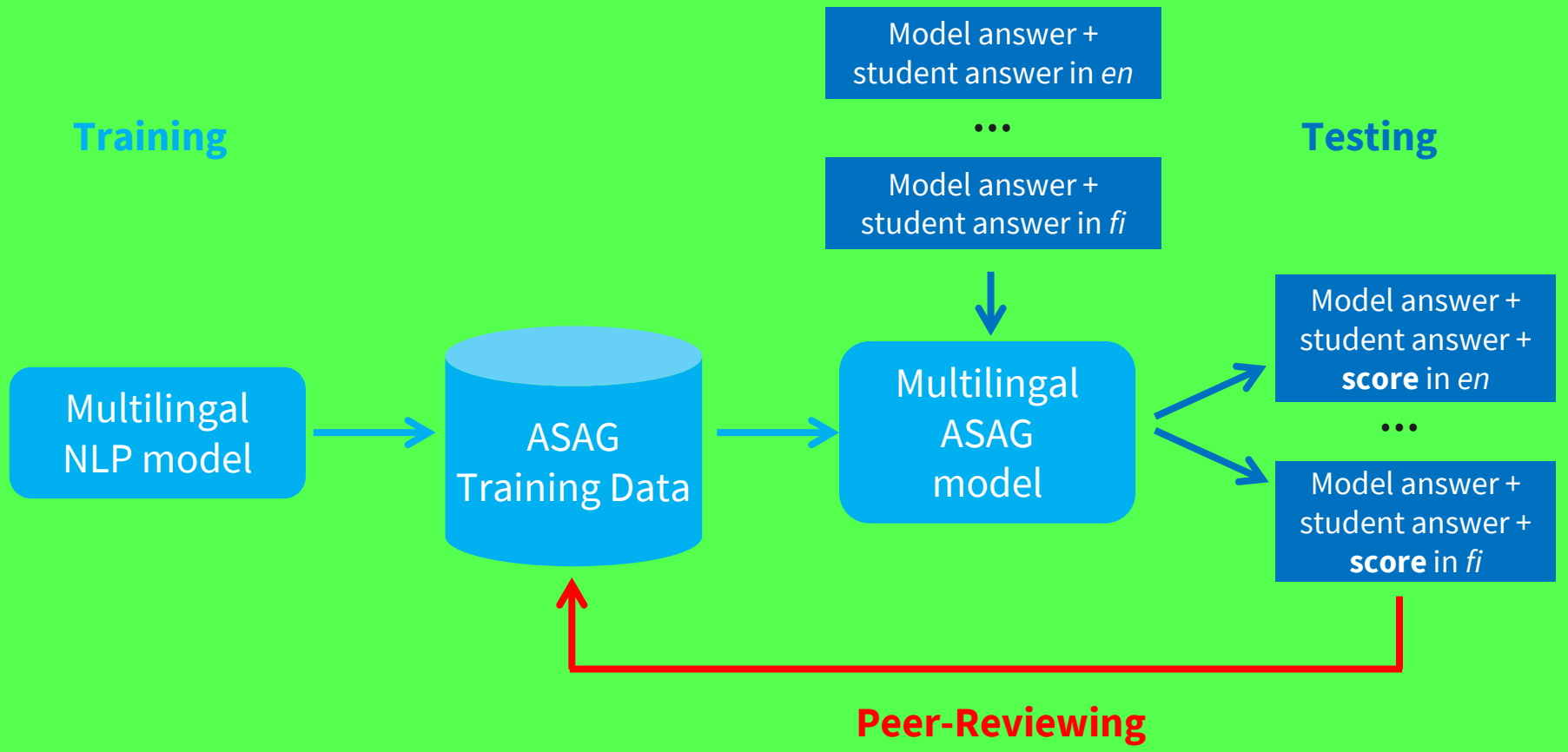
Graphic Source: Custom Depiction.

RELATED WORK

CROUDSOURCING
&
PEER-REVIEWING

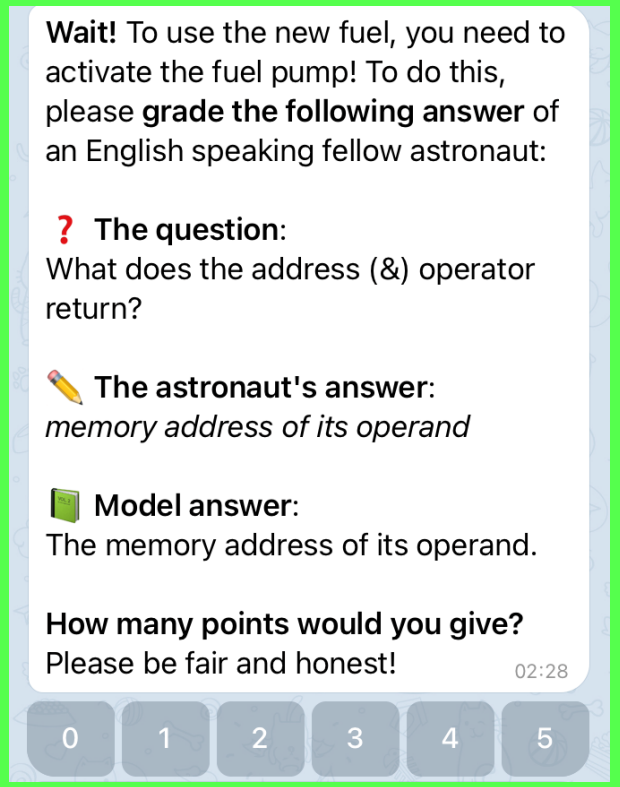
Scoring another answer

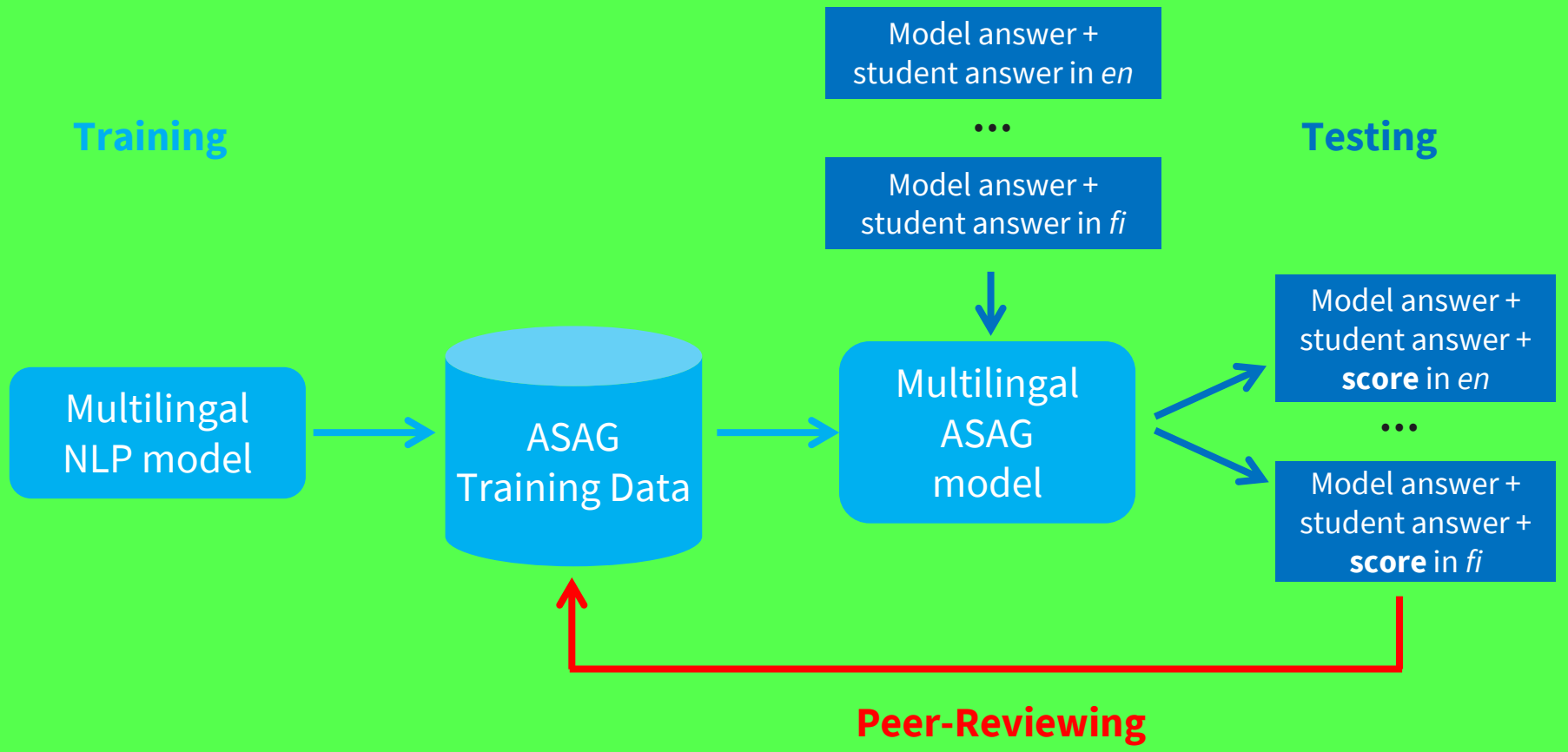
e.g., Peer-based proofreading (Luo et al., 2014)



Graphic Source: Custom Depiction.

RELATED WORK





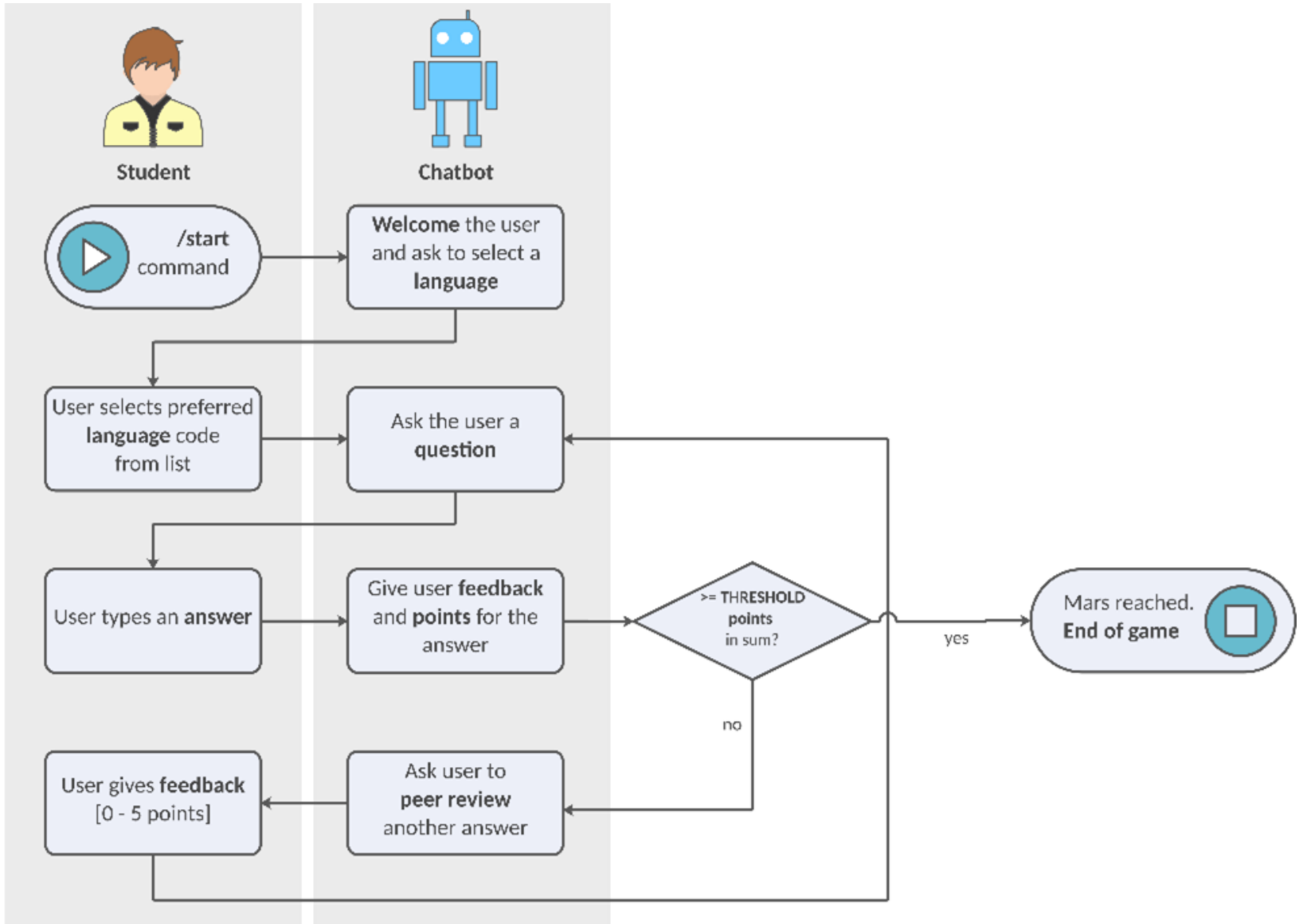
Graphic Source: Custom Depiction.

03

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

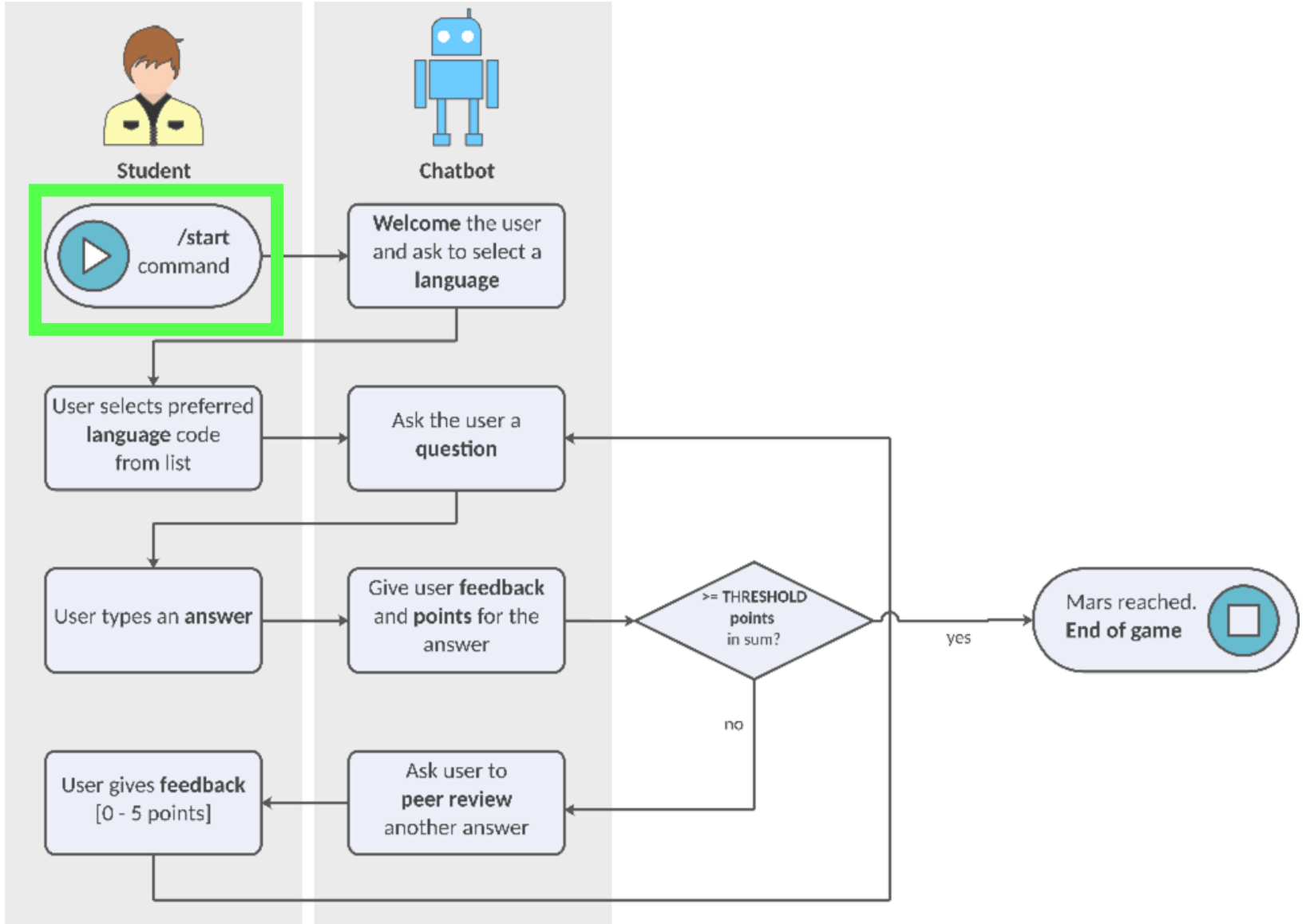
AI-BASED INTERACTIVE EXAM PREPARATION

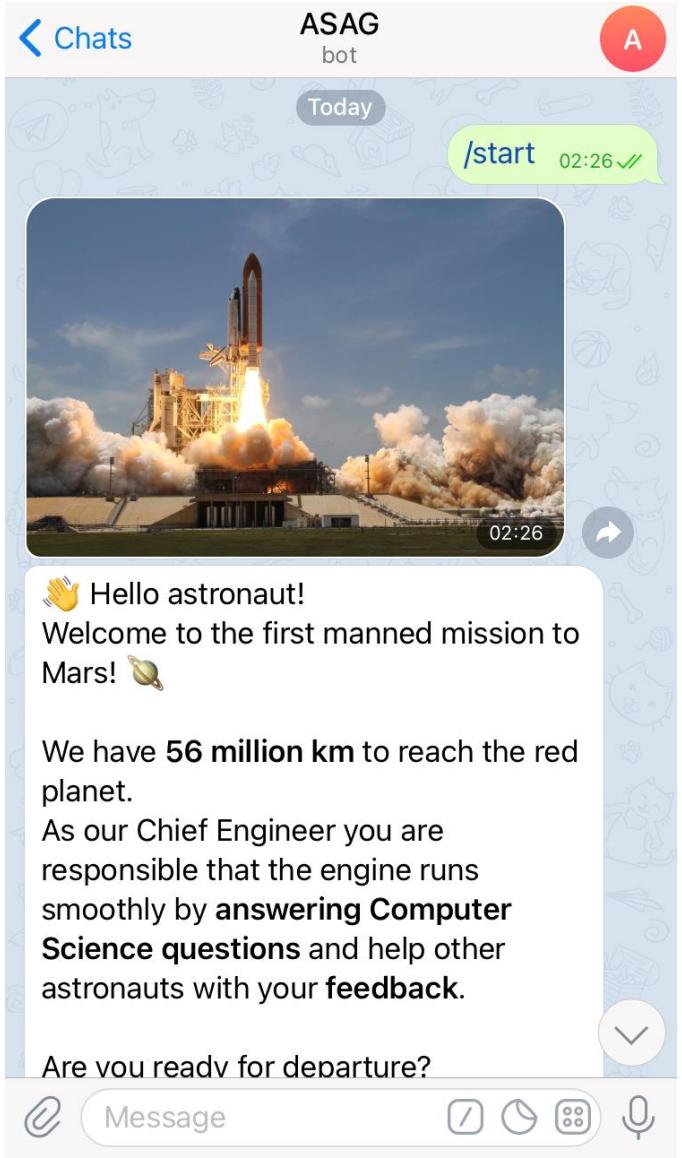
AI-BASED INTERACTIVE EXAM PREPARATION



Graphic Sources: Custom Depiction.

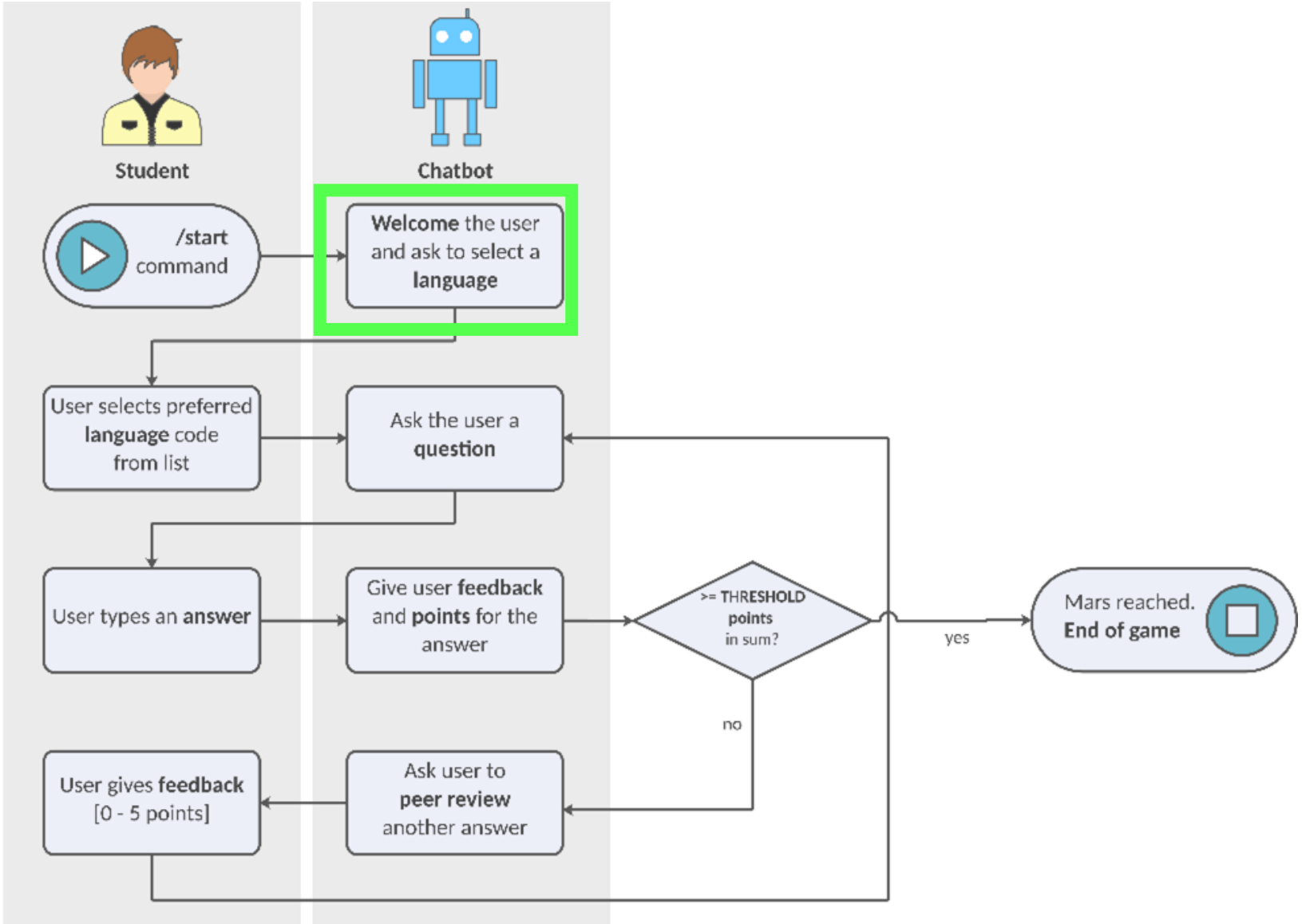
AI-BASED INTERACTIVE EXAM PREPARATION

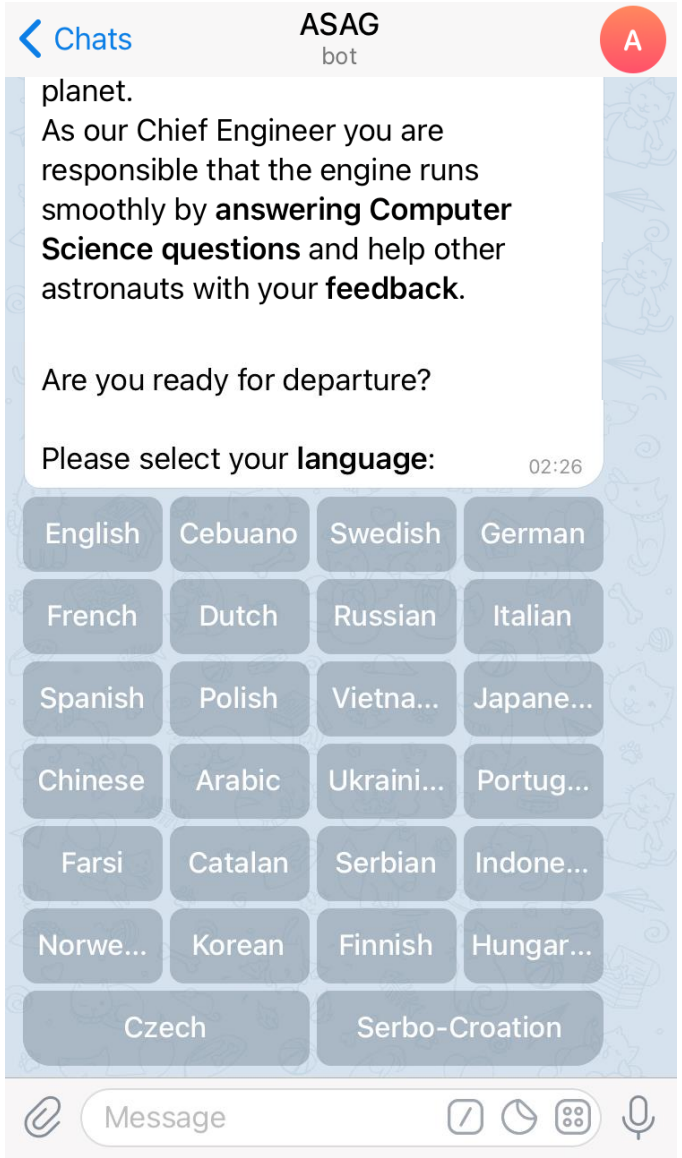




Graphic Sources: Custom Depiction.

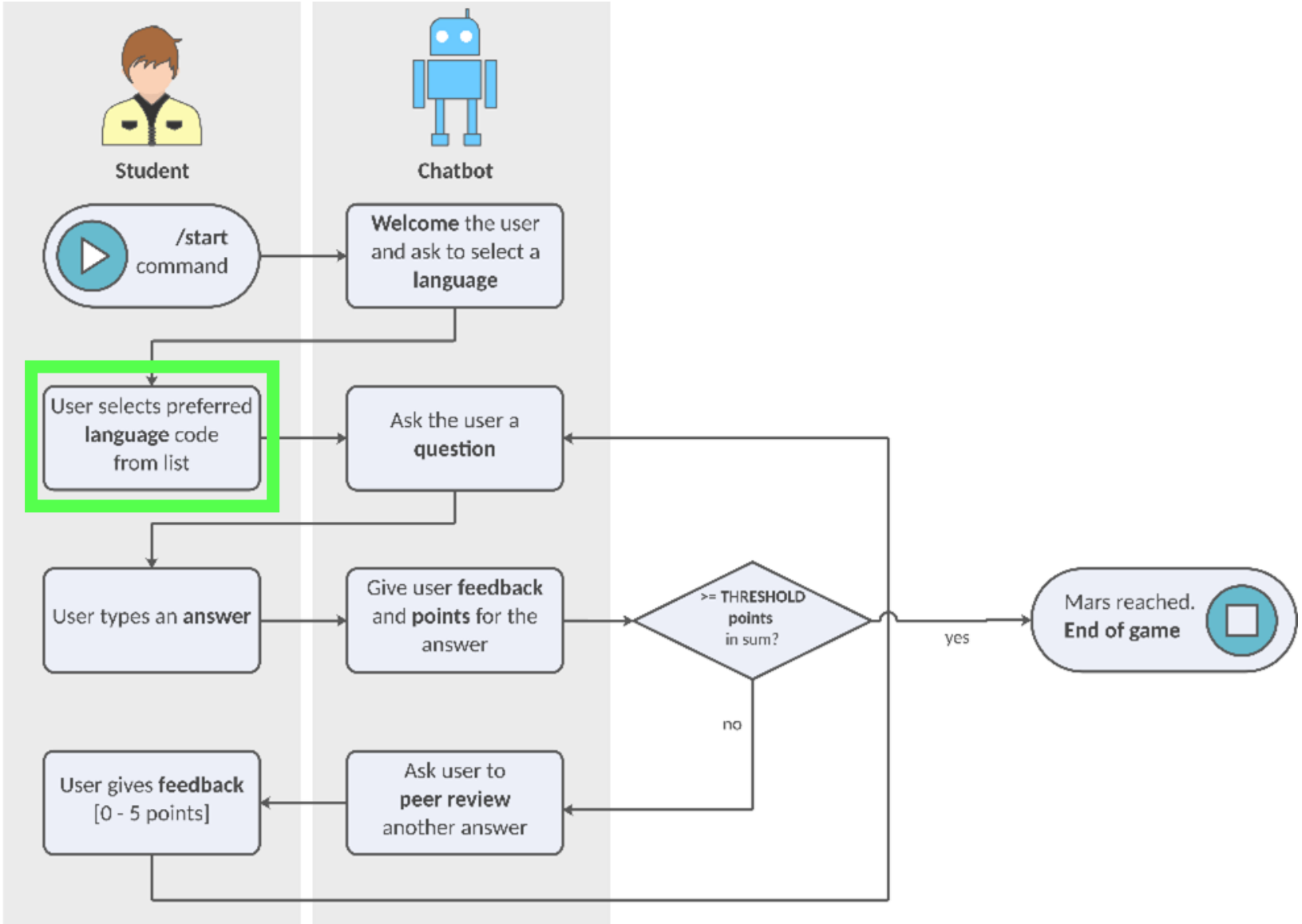
AI-BASED INTERACTIVE EXAM PREPARATION





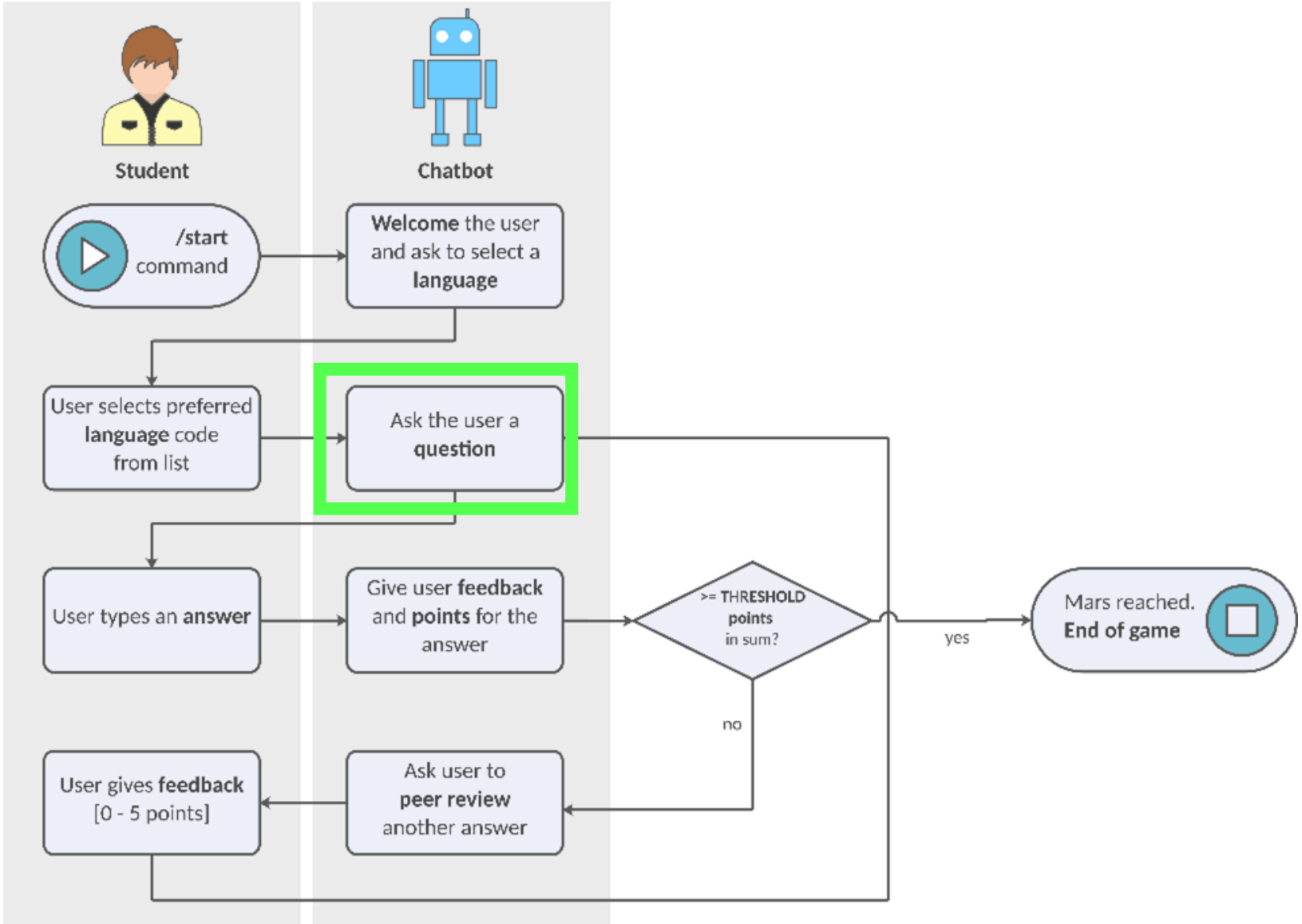
Graphic Sources: Custom Depiction.

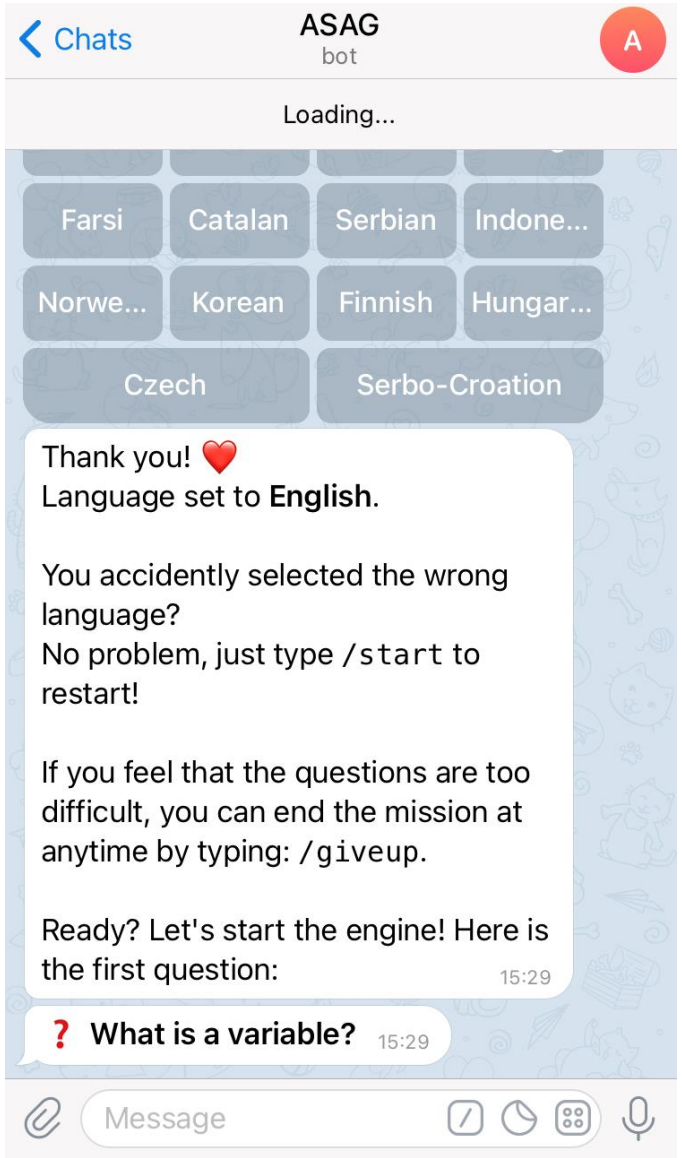
AI-BASED INTERACTIVE EXAM PREPARATION



Graphic Sources: Custom Depiction.

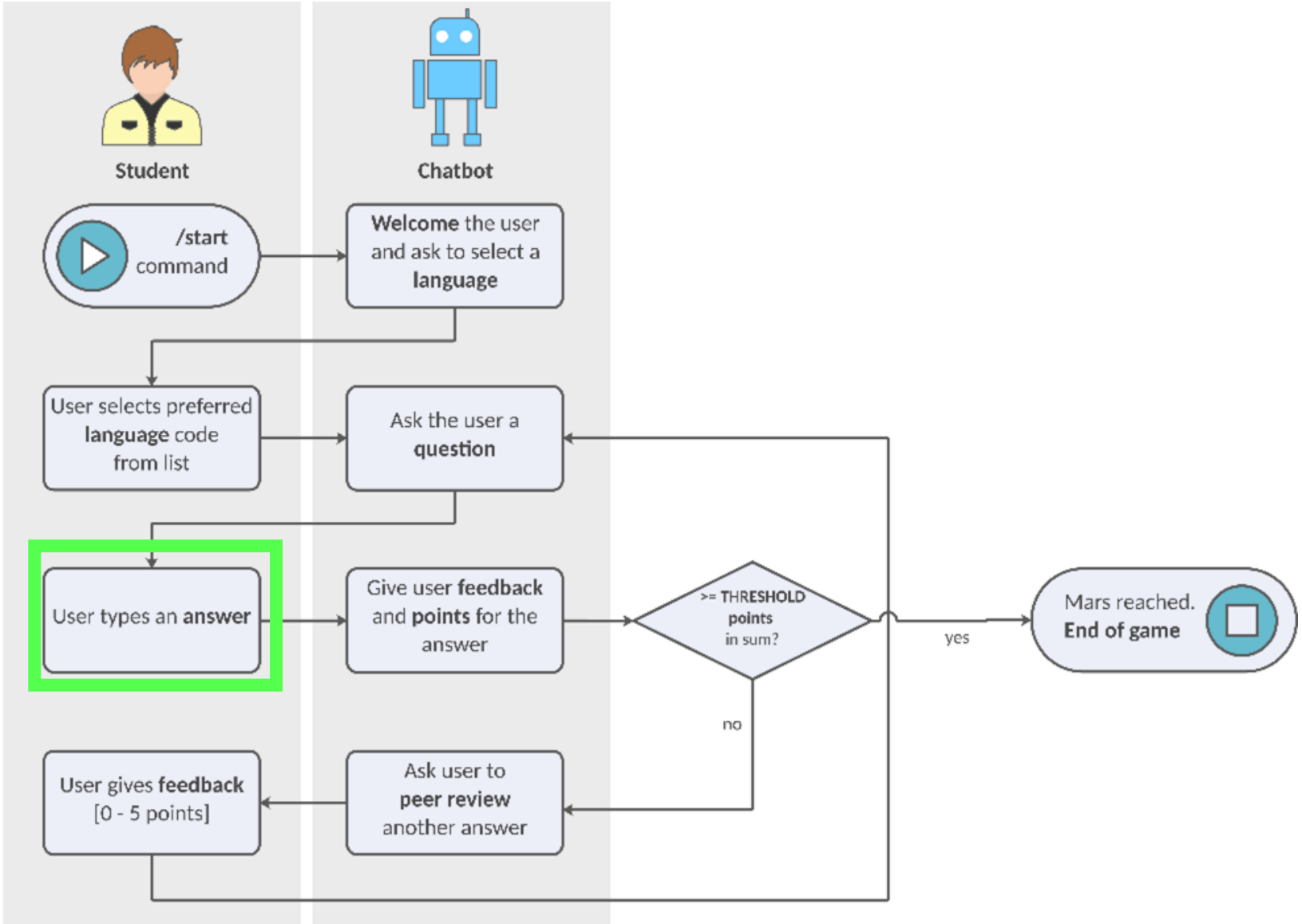
AI-BASED INTERACTIVE EXAM PREPARATION

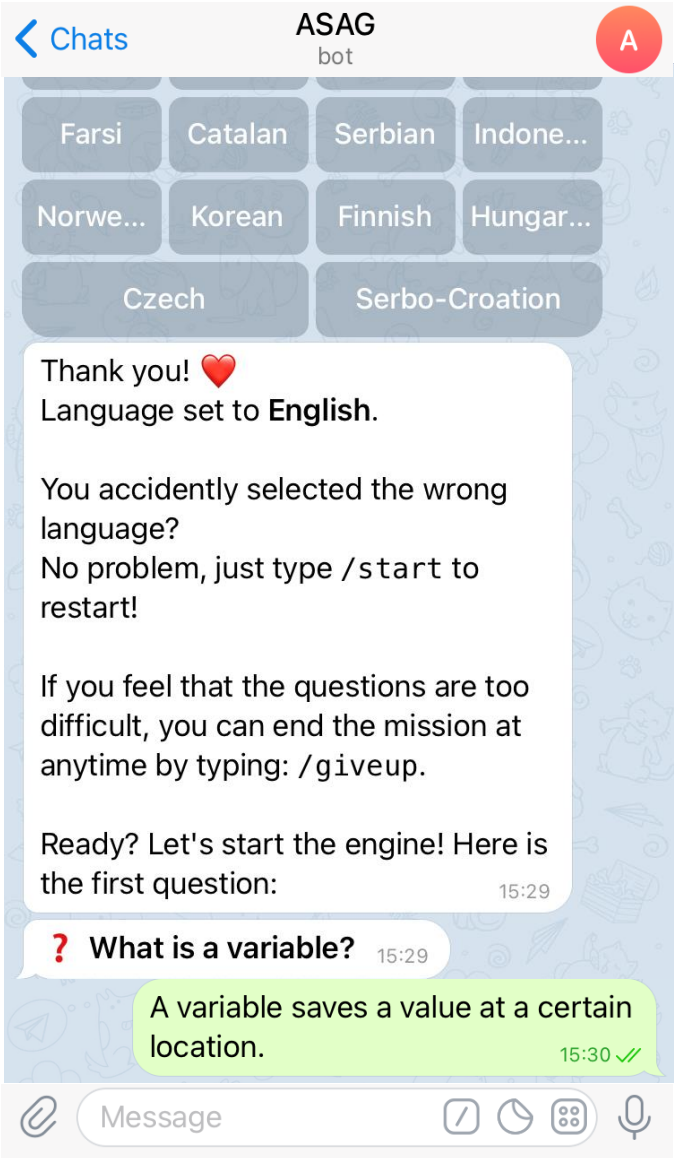




Graphic Sources: Custom Depiction.

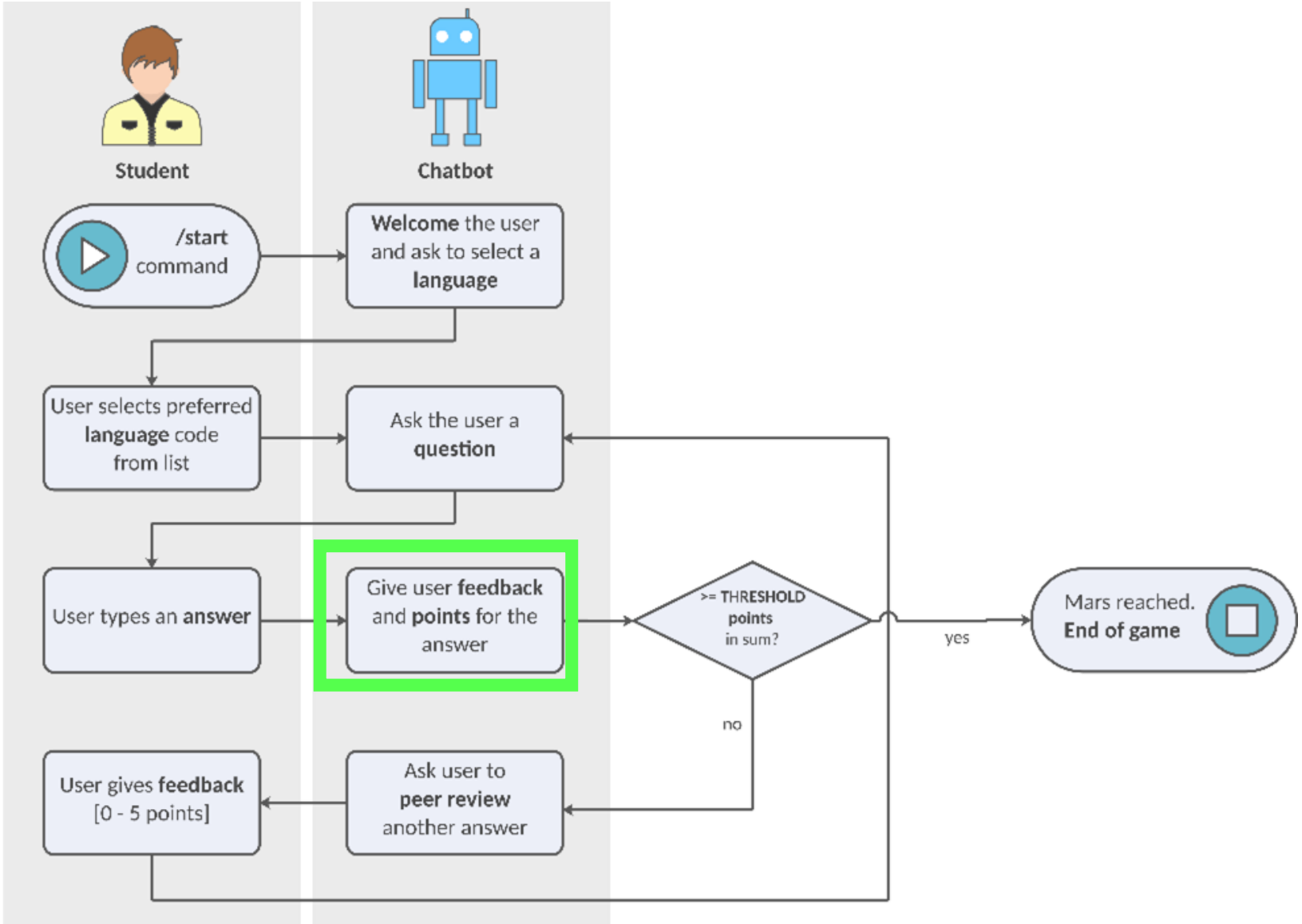
AI-BASED INTERACTIVE EXAM PREPARATION

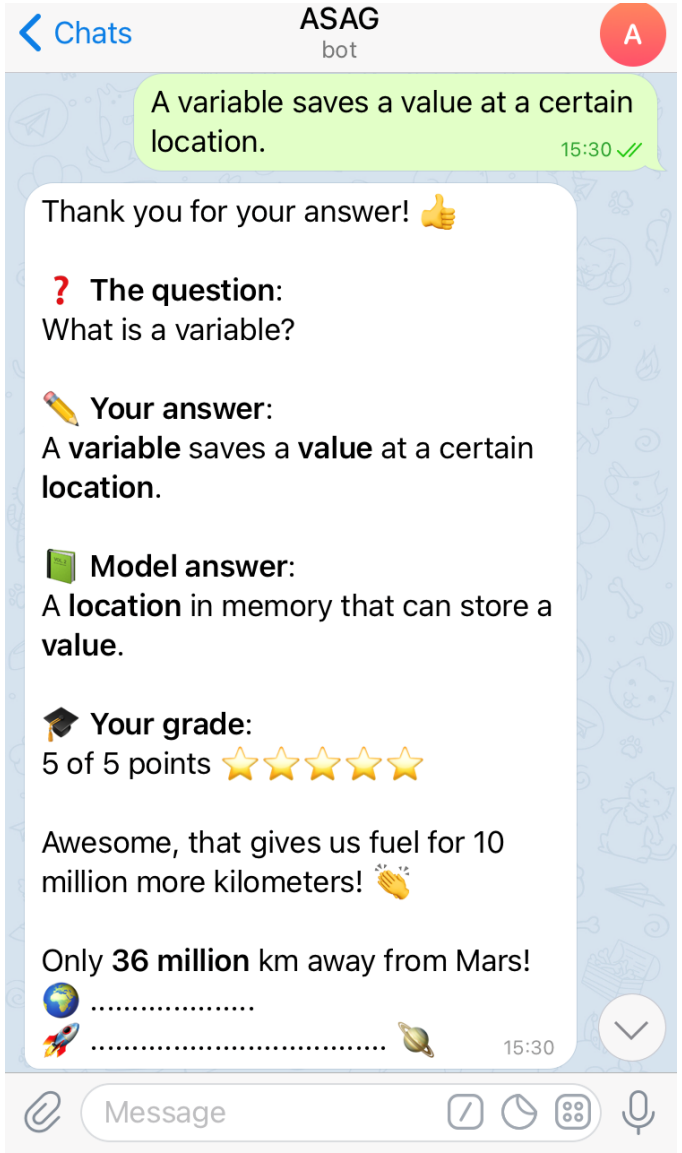




Graphic Sources: Custom Depiction.

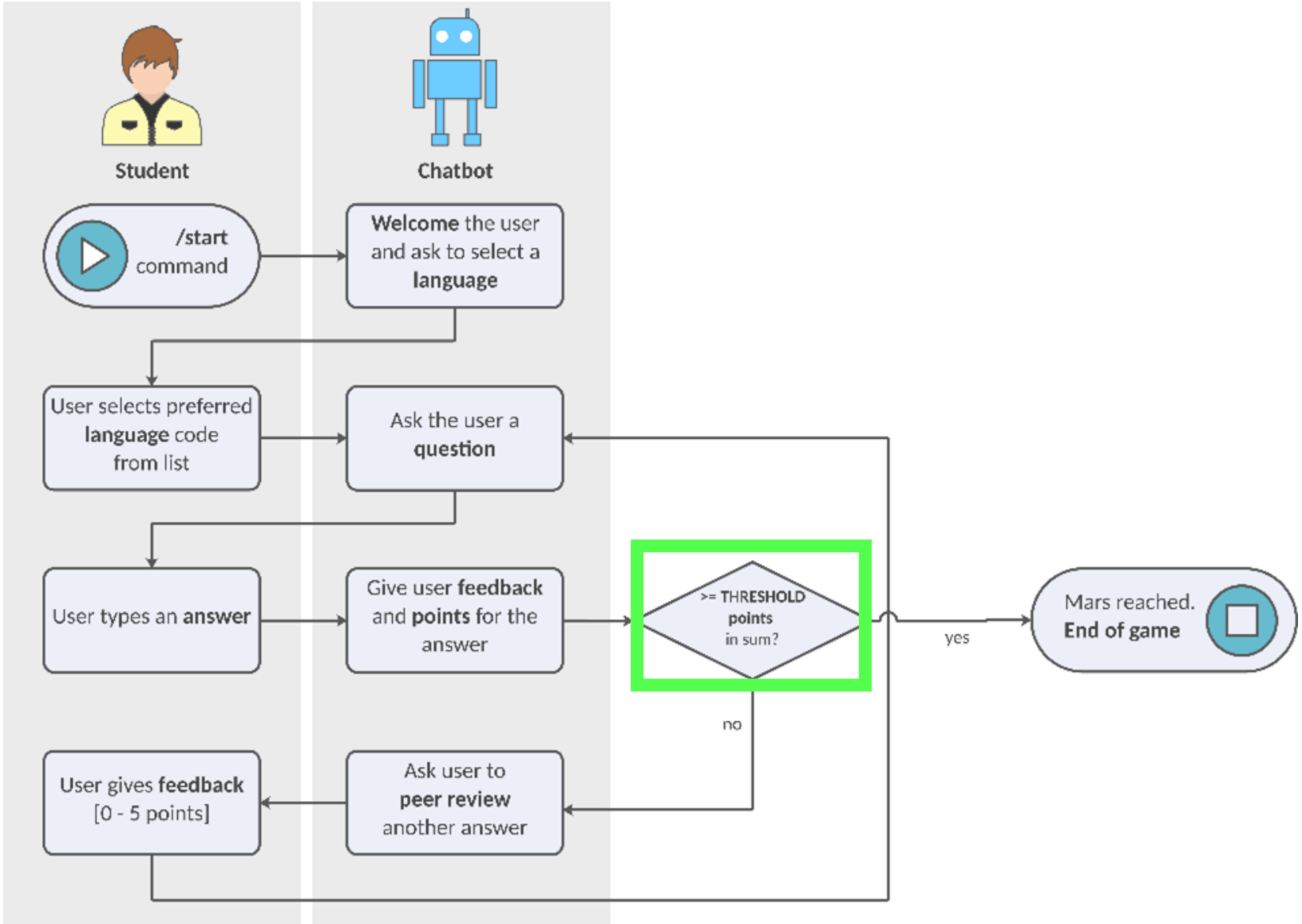
AI-BASED INTERACTIVE EXAM PREPARATION

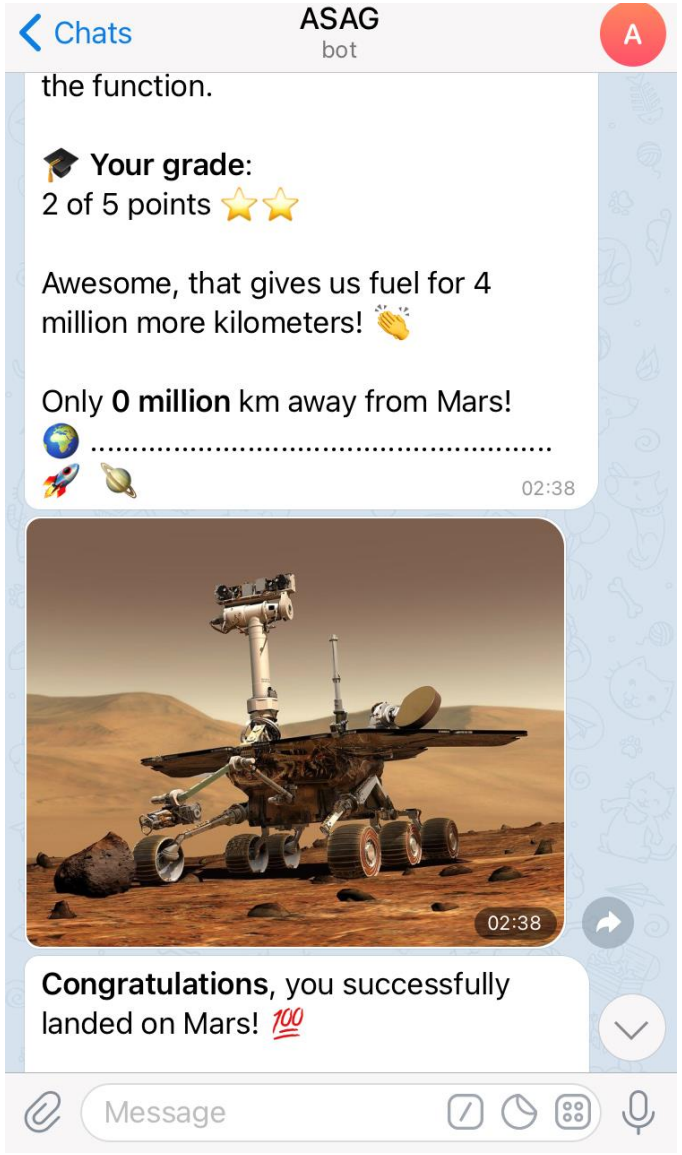




Graphic Sources: Custom Depiction.

AI-BASED INTERACTIVE EXAM PREPARATION

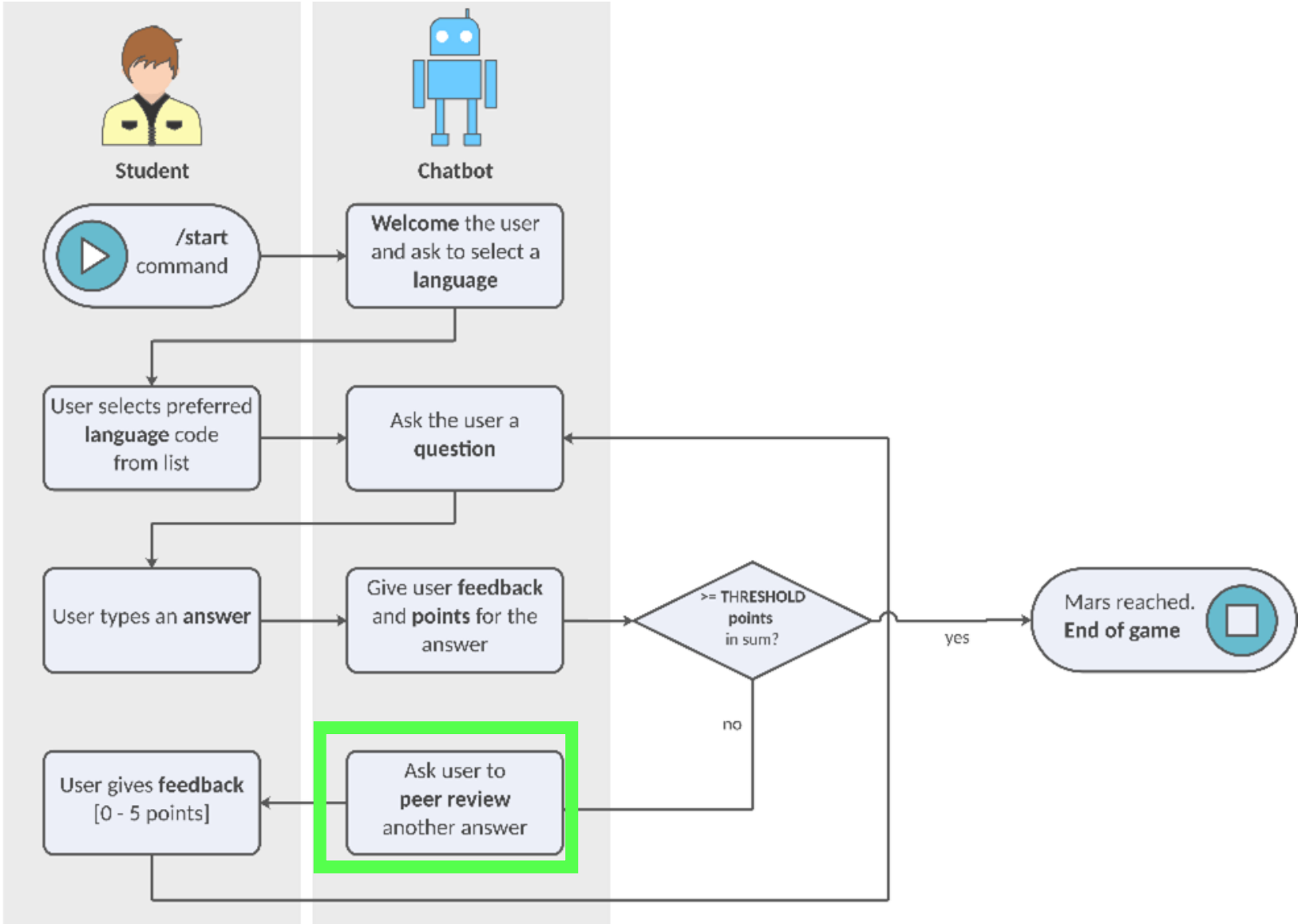


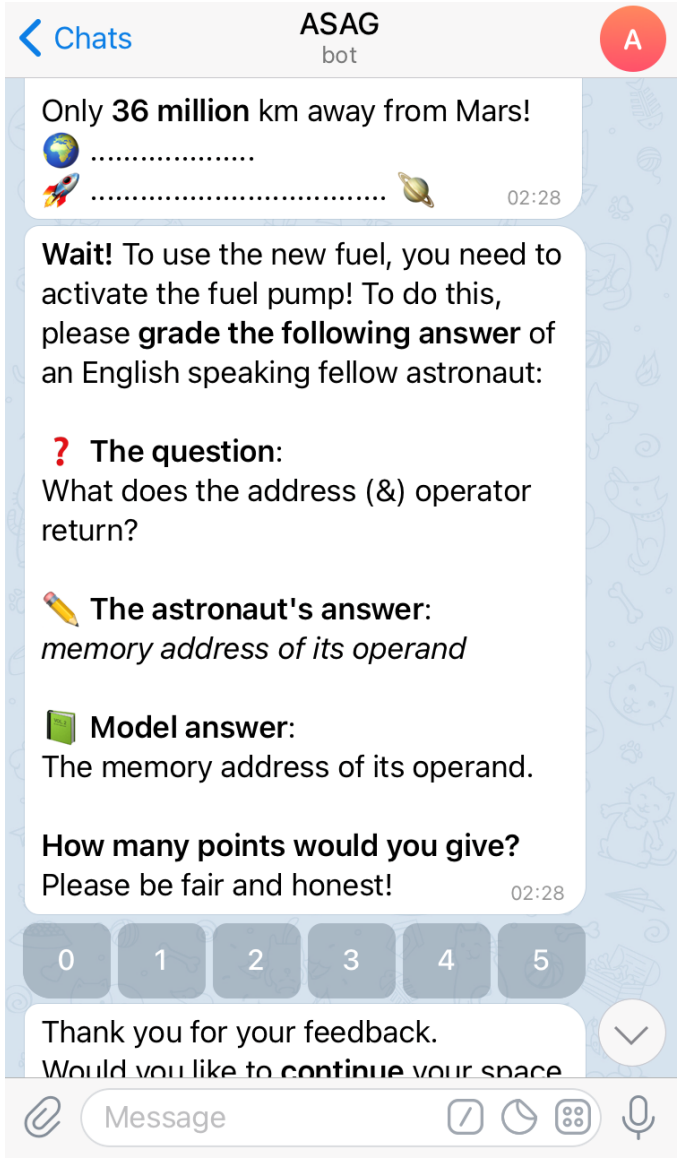


Graphic Sources: Custom Depiction.

AI-BASED INTERACTIVE EXAM PREPARATION

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES





Graphic Sources: Custom Depiction.

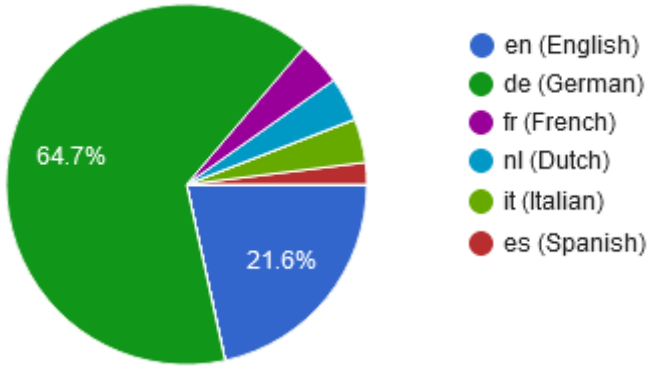
04

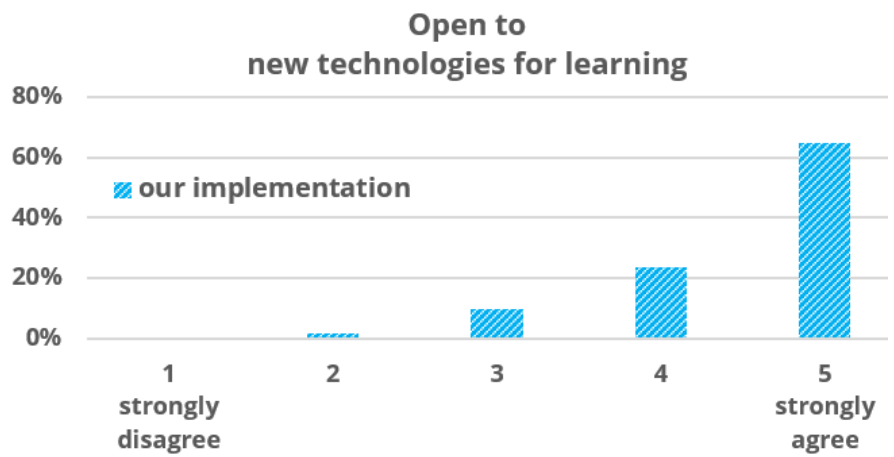
INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

EXPERIMENTS AND RESULTS

SURVEY

- 51 participants
- From 6 countries
- Most participants are students (most 18-44 years old)





Graphic Sources: Custom Depiction.

SURVEY

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

GAMIFICATION

- Story
- Theme

USER EXPERIENCE

- Easy to use?
- Fun?

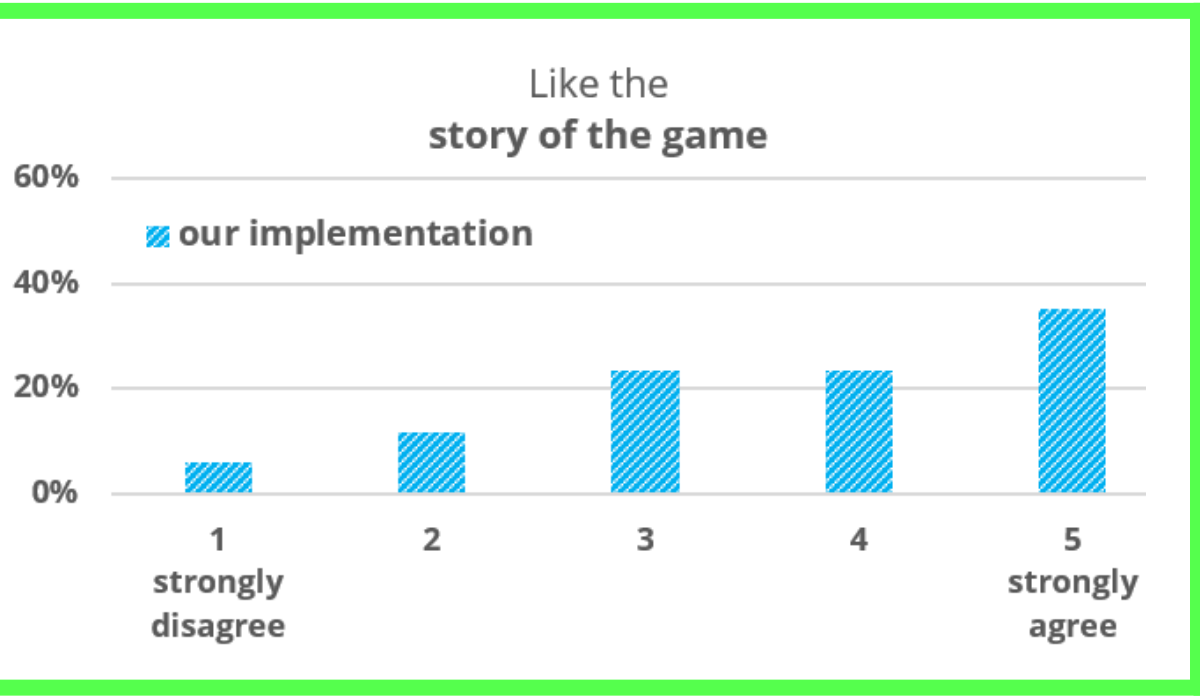
QUALITY OF THE NLP MODELS

- Translation quality?
- Scoring quality?
- Keyword highlighting?

LEARNING EXPERIENCE + MOTIVATION

- Improved experience?
- Makes sense?
- Usage?
- Fellow students?
- Acceleration?
- Scoring ?
- Languages?
- Elderly people?
- Motivation?

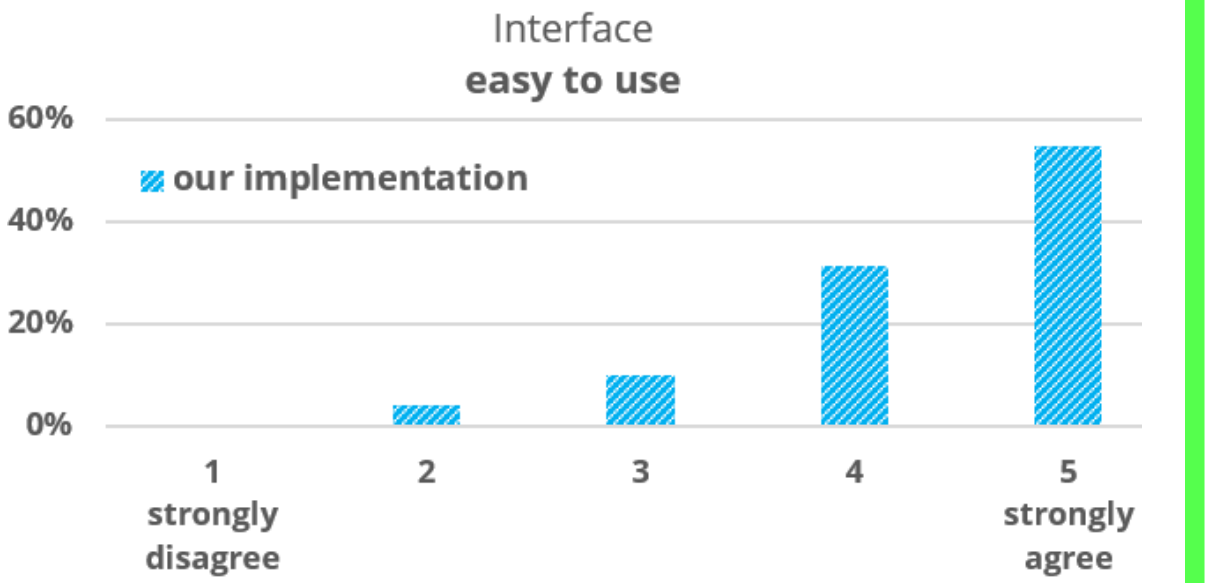
SURVEY – GAMIFICATION

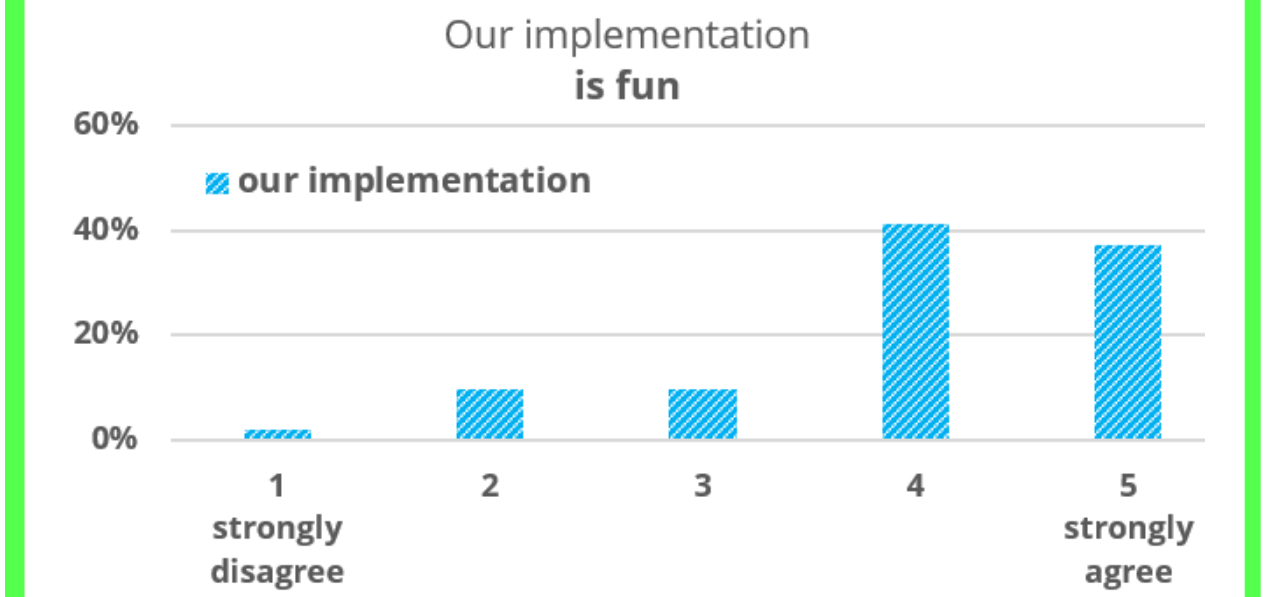


➔ Good gamification with a simple story and without special graphical features

Graphic Sources: Schlippe & Sawatzki (2021).

SURVEY – USER EXPERIENCE

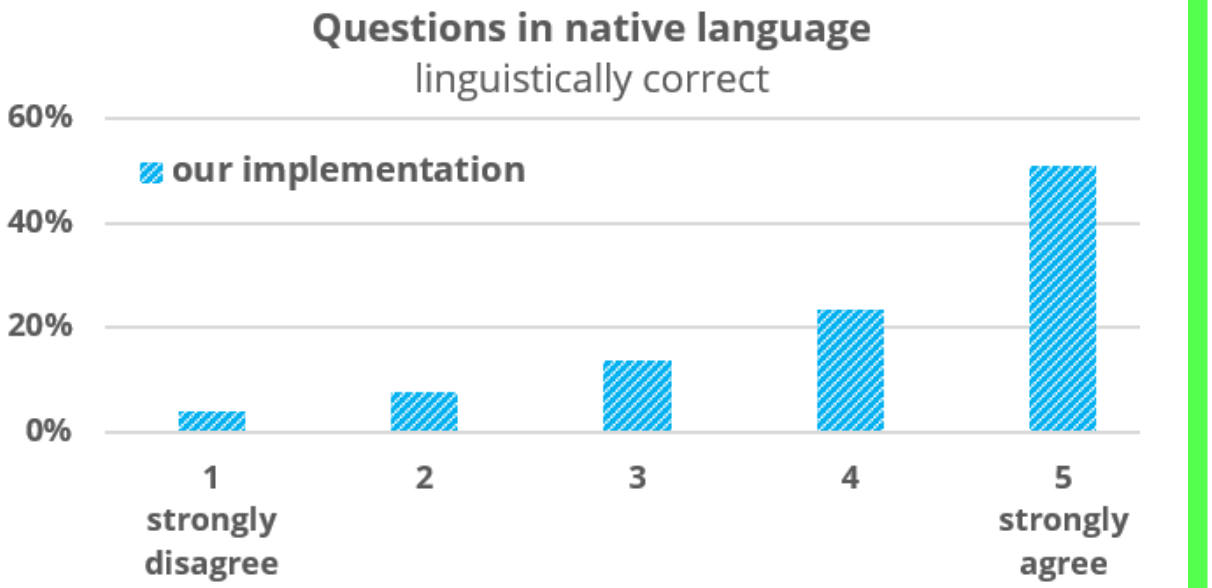


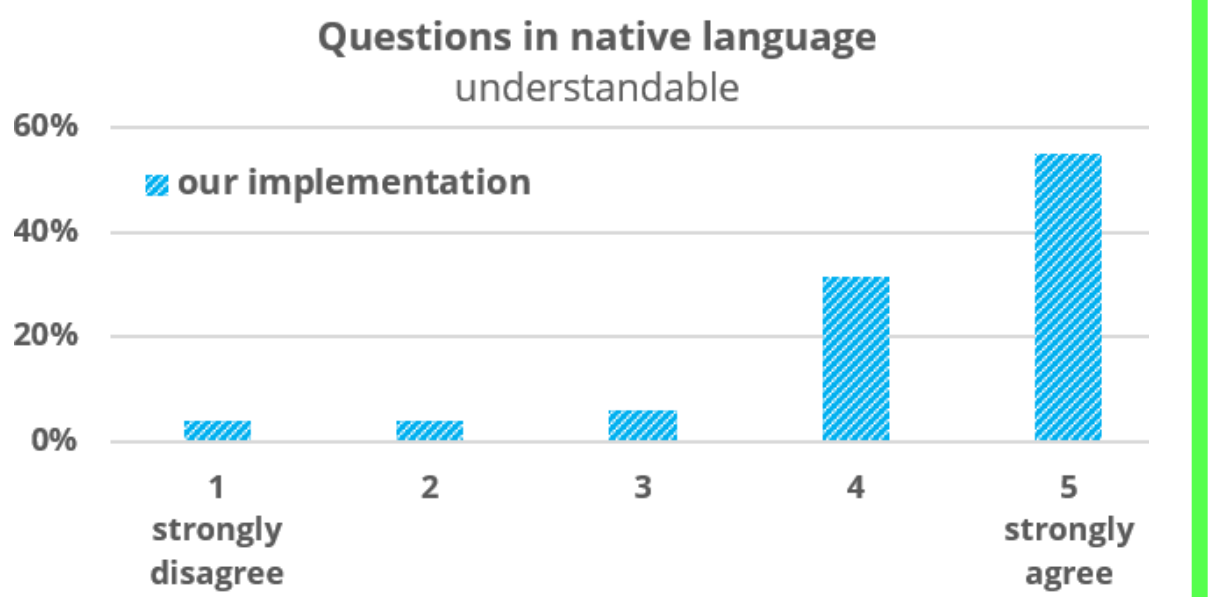


➔ Interface is easy to use and operating the exam trainer is fun

Graphic Sources: Schlippe & Sawatzki (2021).

SURVEY – NLP MODELS (1)

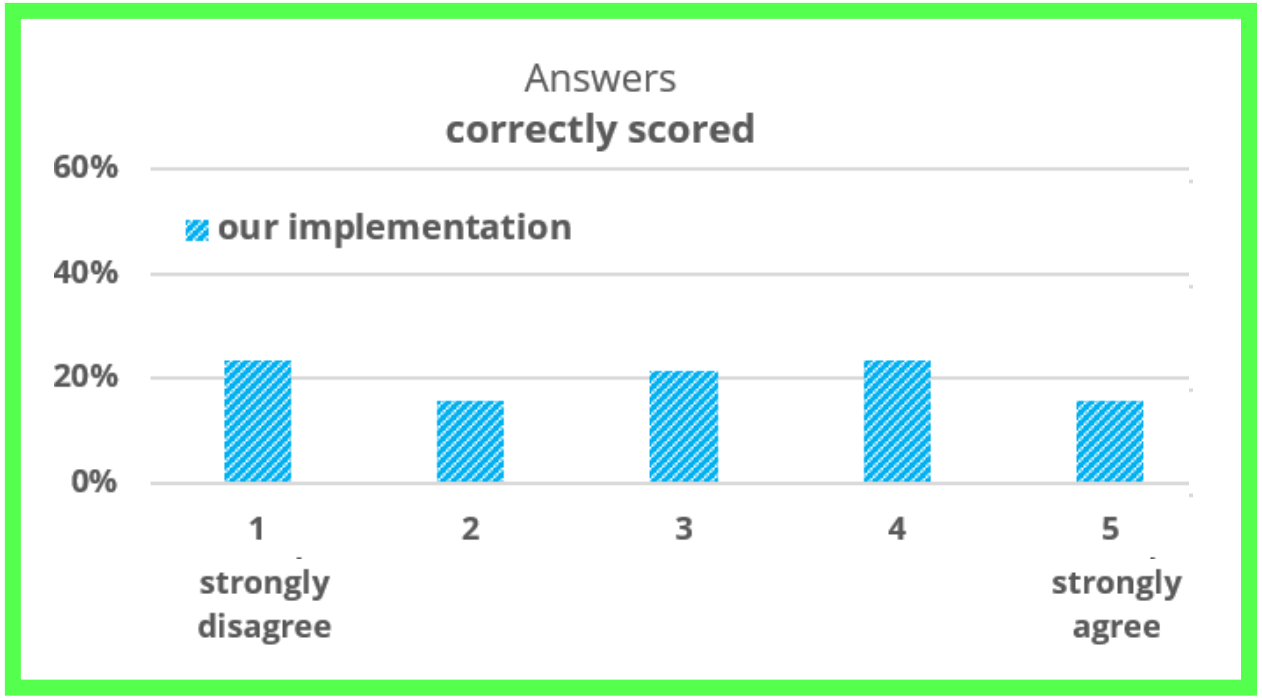




➔ Majority rates the machine-translated questions as linguistically correct and understandable

Graphic Sources: Schlippe & Sawatzki (2021).

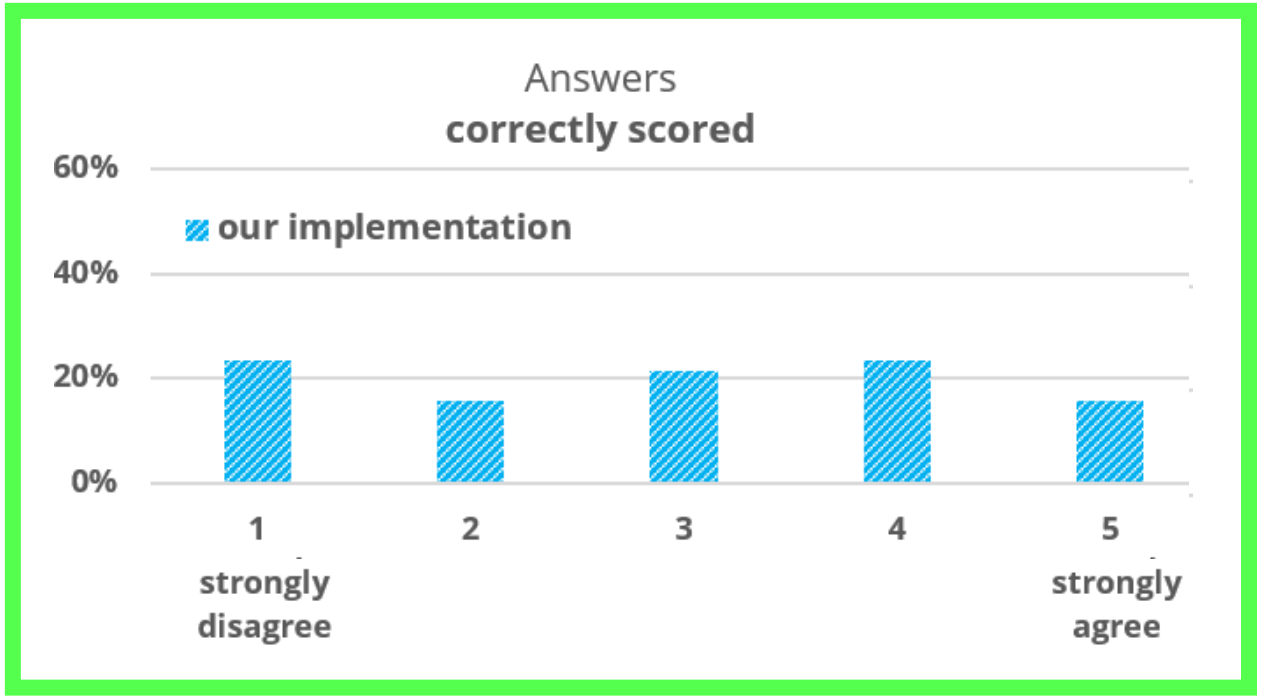
SURVEY – NLP MODELS (2)

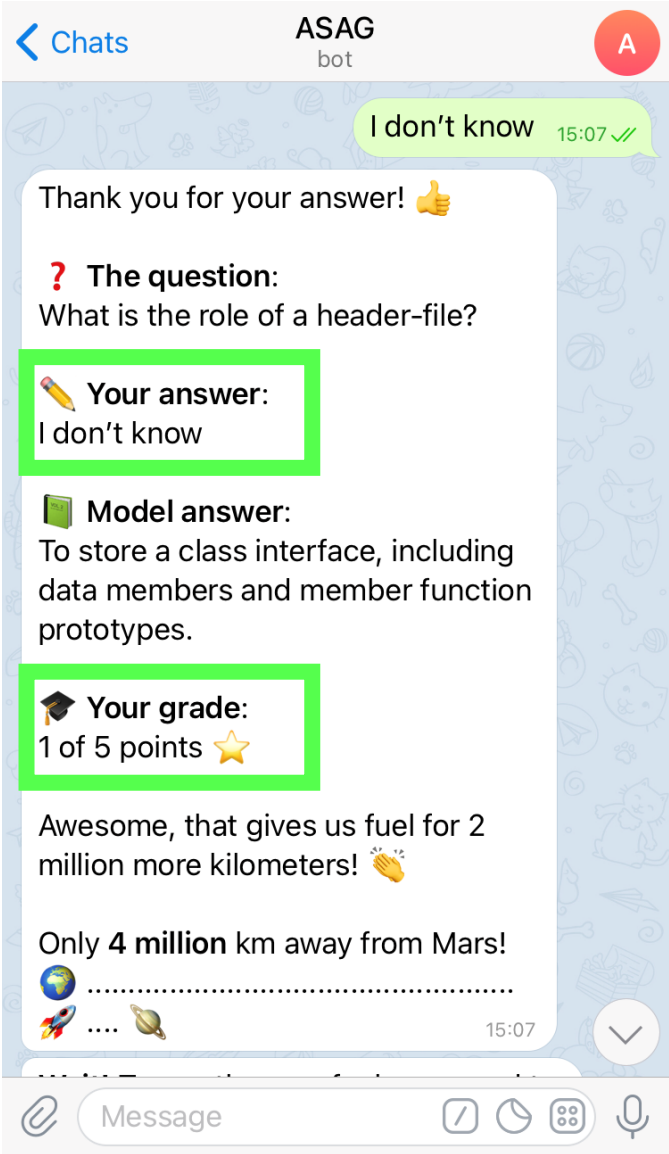


➔ Automatic scoring of the student answers was only rated average

Graphic Source: Schlippe & Sawatzki (2021).

SURVEY – NLP MODELS (2)

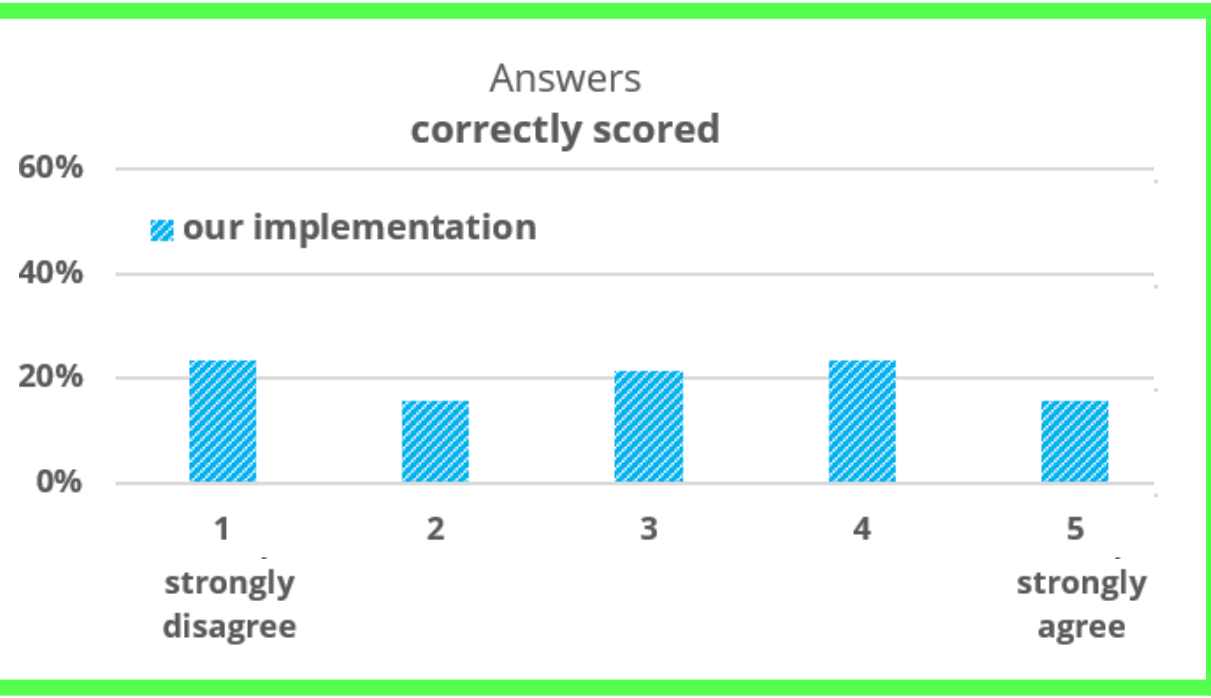


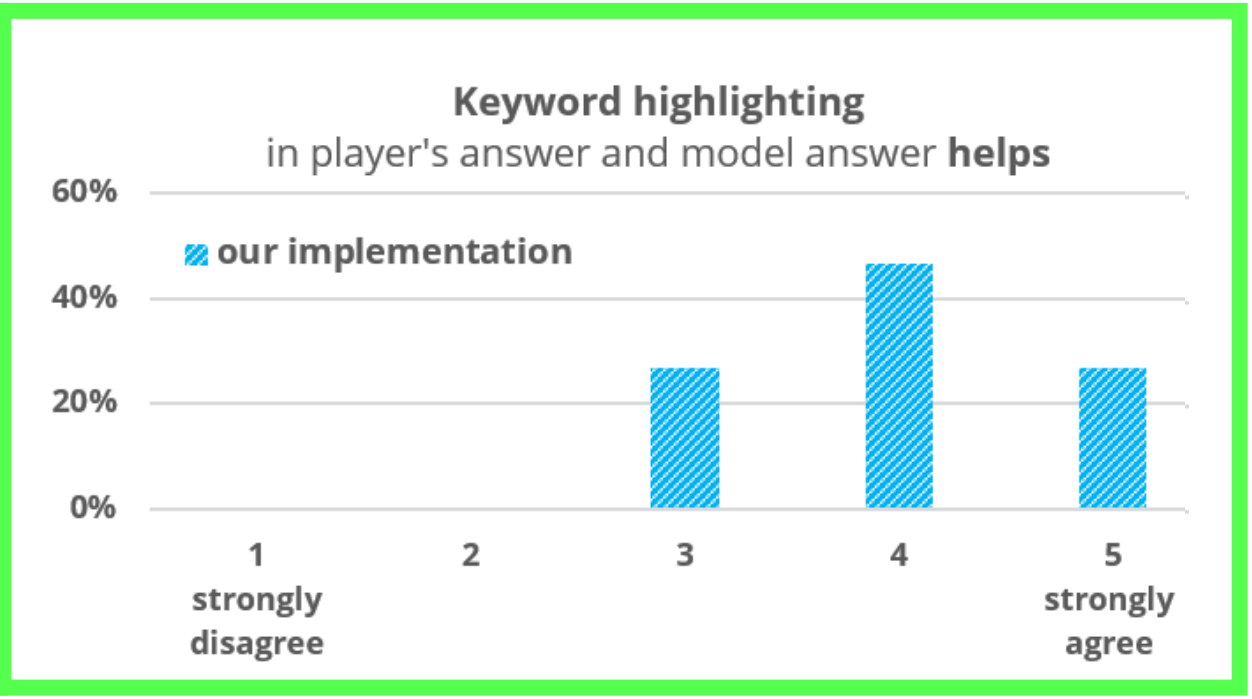


➔ Automatic scoring of the student answers was only rated average

Graphic Sources: Schlippe & Sawatzki (2021); Custom Depiction.

SURVEY – NLP MODELS (2)

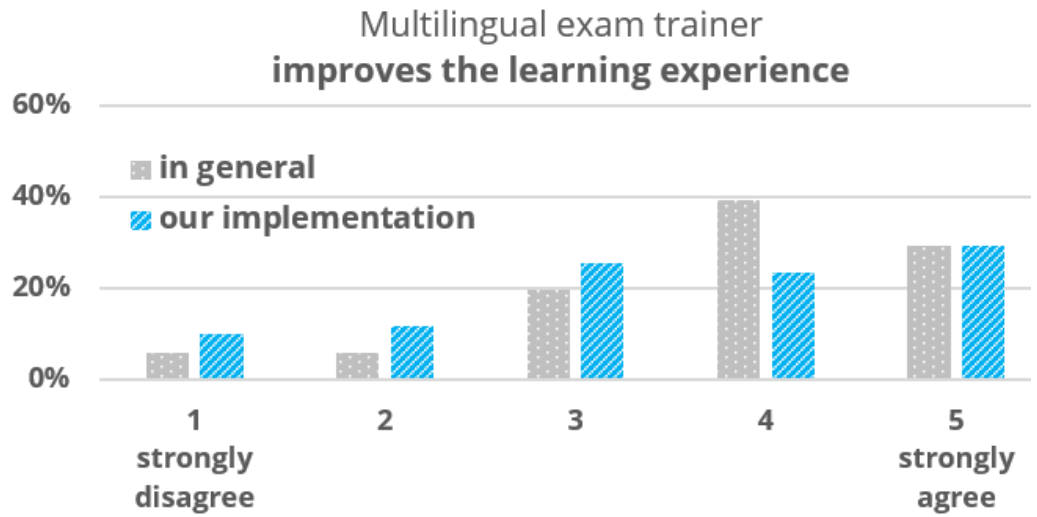


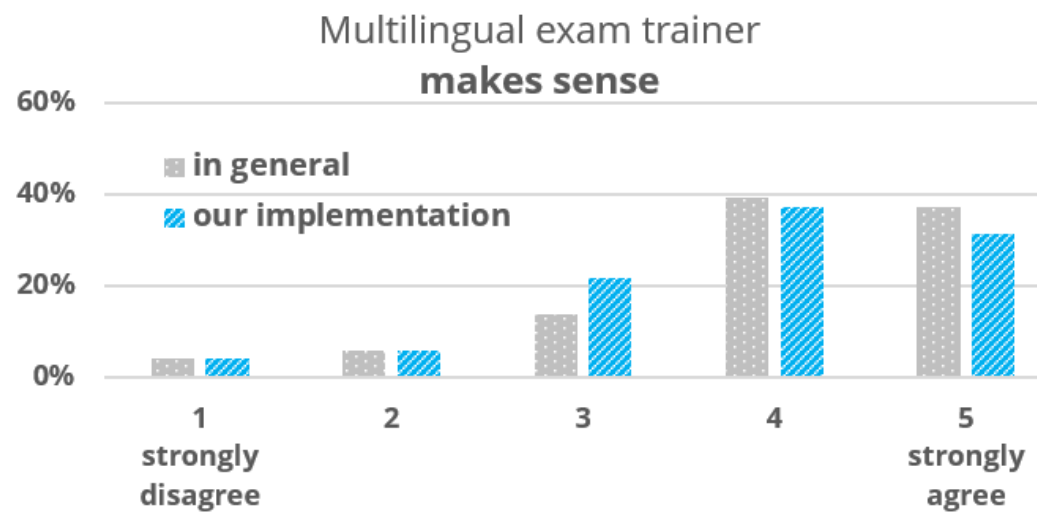


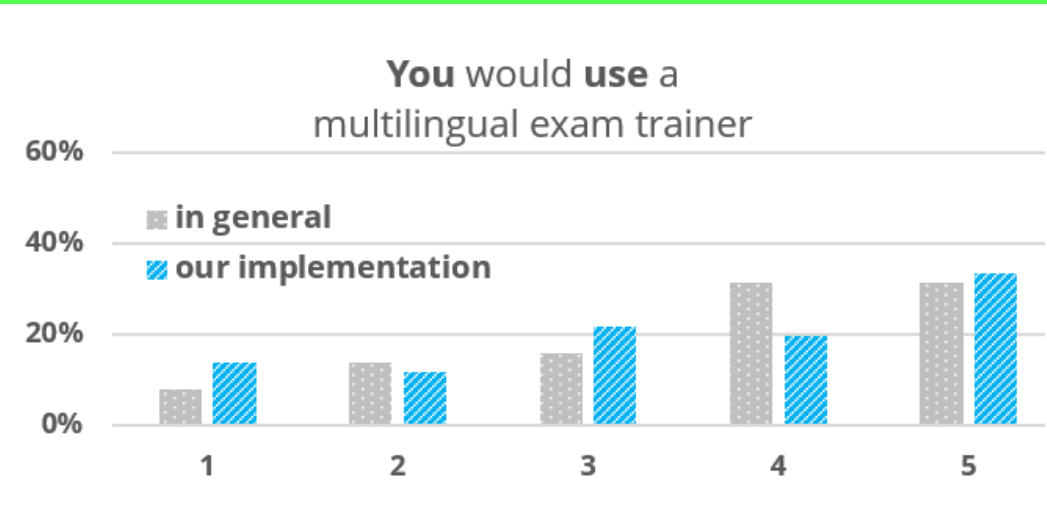
- ➔ Automatic scoring of the student answers was only rated average (Problem: Random answers)
- ➔ Explainability approach with keyword highlighting was well rated

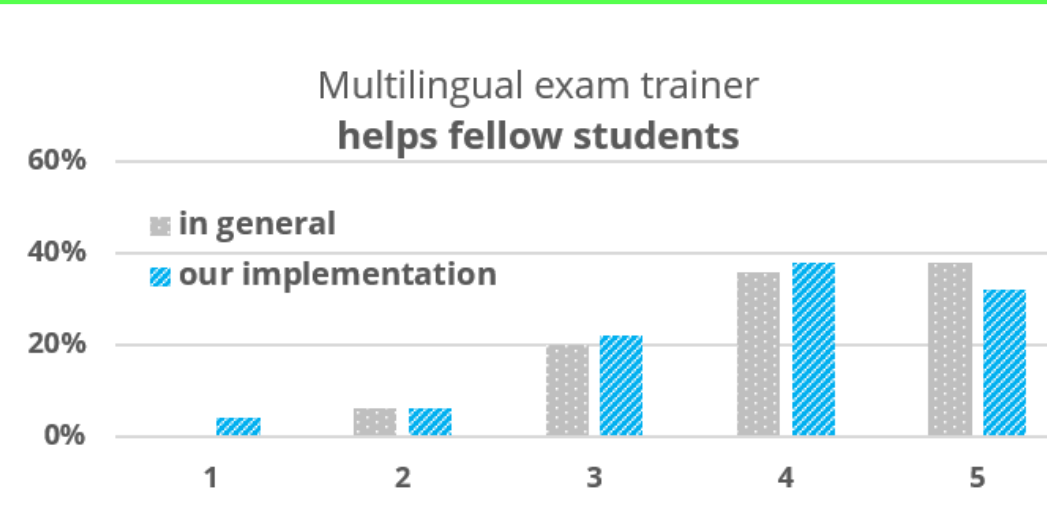
Graphic Sources: Schlippe & Sawatzki (2021).

SURVEY – LEARNING EXPERIENCE (1)



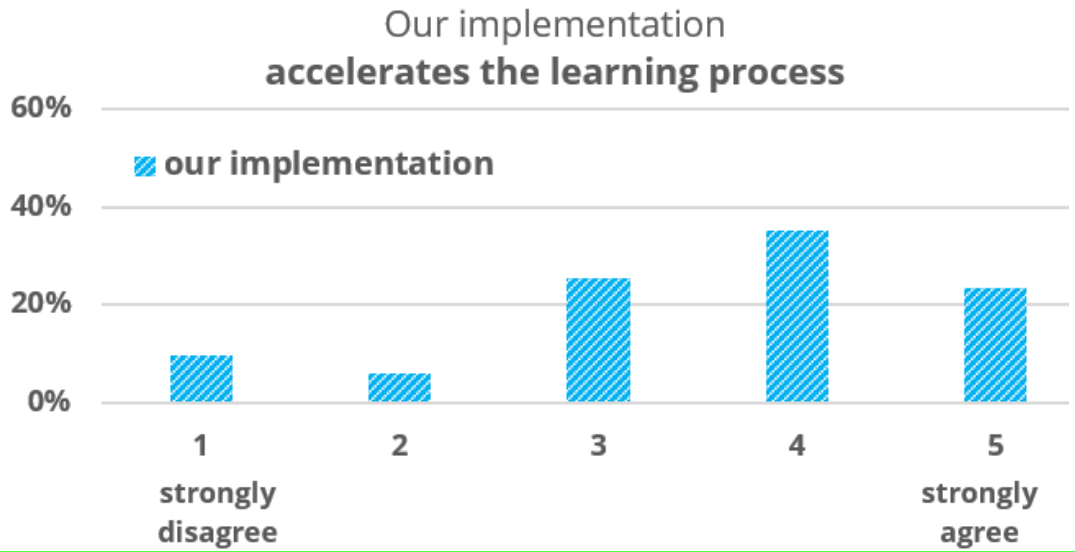


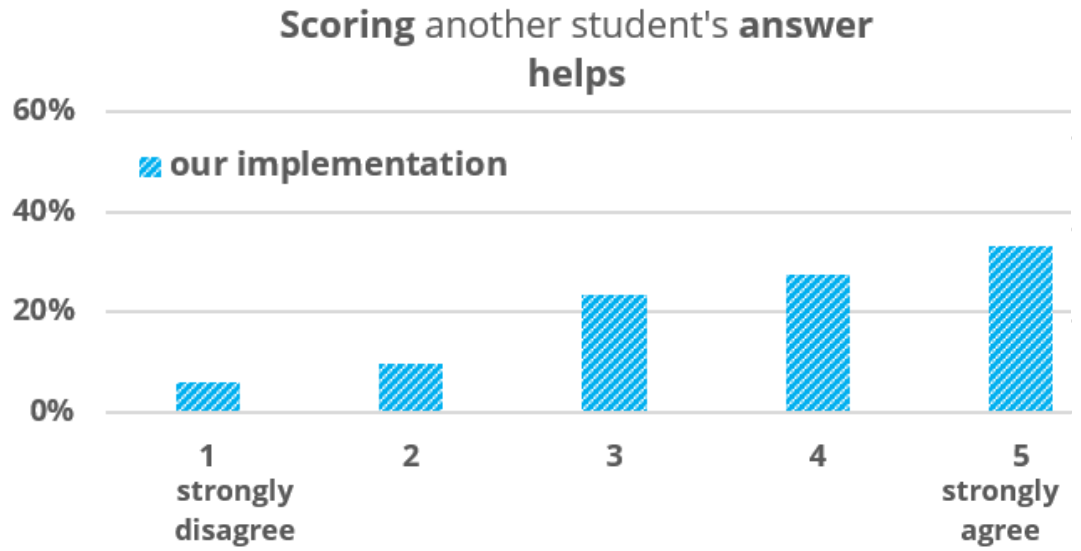


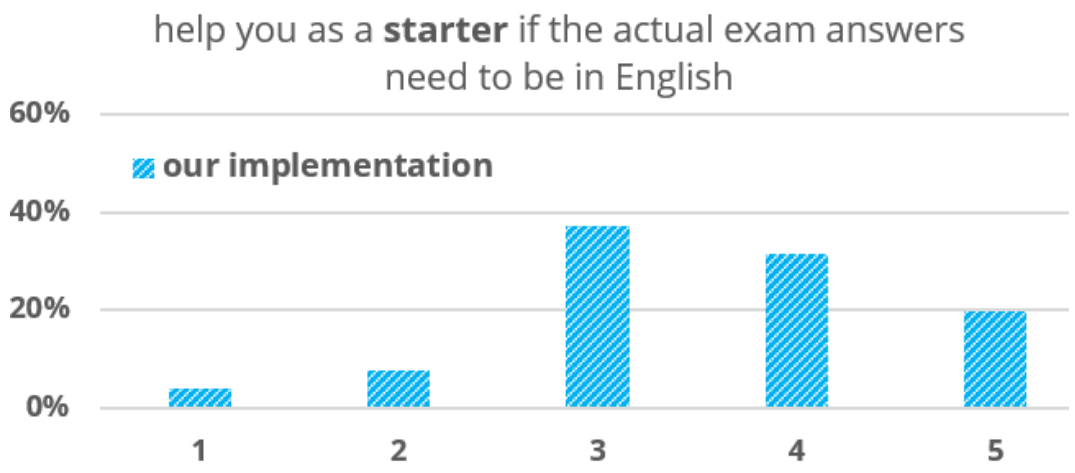


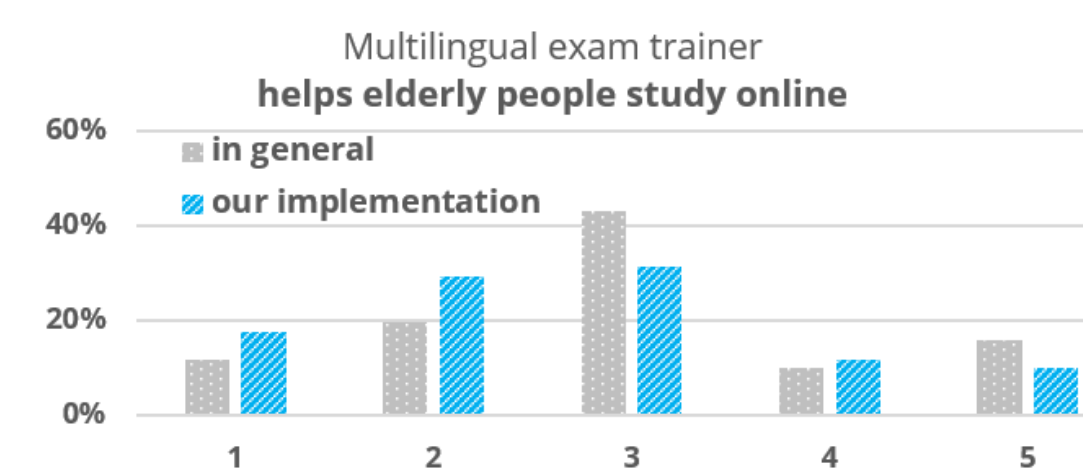
Graphic Sources: Schlippe & Sawatzki (2021).

SURVEY – LEARNING EXPERIENCE (2)









Graphic Sources: Schlippe & Sawatzki (2021).

05

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

CONCLUSION AND FUTURE WORK

CONCLUSION AND FUTURE WORK

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

Conclusion

- Multilingual interactive conversational AI tutoring system for exam preparation
- combines multilingual NLP components, ASAG, conversational AI, keyword extraction, learning analytics, crowdsourcing, and gamification
- positive feedback in a survey regarding learning experience, user experience, motivation, quality of NLP models, and gamification
- proof-of-concept where users have tested 6 languages so far

CONCLUSION AND FUTURE WORK

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

Conclusion

- Multilingual interactive conversational AI tutoring system for exam preparation
- combines multilingual NLP components, ASAG, conversational AI, keyword extraction, learning analytics, crowdsourcing, and gamification
- positive feedback in a survey regarding learning experience, user experience, motivation, quality of NLP models, and gamification
- proof-of-concept where users have tested 6 languages so far

Furture Work

- extension to other languages
- address the issue of explainability to provide even better support to the users
- more emotional dialog, e.g., (Wölfel et al., 2015; Schlippe et al., 2020)

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

THANK YOU

Tim Schlippe
tim.schlippe@iu.org

REFERENCES

Literature

- **United Nations: Sustainable Development Goals: 17 Goals to Transform our World (2021)**: https://www.un.org/sustainabledevelopment/sustainable-development-goals
- **Statista: The Most Spoken Languages Worldwide in 2019 (2020)**: https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide
- **Schlippe, T. & Sawatzki, J. (2021)**: *AI-based Multilingual Interactive Exam Preparation*. In: Proceedings of The Learning Ideas Conference 2021 (14th annual conference), ALICE - Special Conference Track on Adaptive Learning via Interactive, Collaborative and Emotional Approaches, New York, New York, 14-18 June 2021. In: Advances in Intelligent Systems and Computing, Springer, 2021.
- **Devlin, J., Chang, M.W., Lee, K., & Toutanova, K. (2019)**: *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. In: NAACL-HLT, Minneapolis, Minnesota.
- **Pires, T., Schlinger, E., & Garrette, D. (2019)**: *How Multilingual is Multilingual BERT?* In: ACL. Florence, Italy.
- **Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019)**: *RoBERTa: A Robustly Optimized BERT Pretraining Approach*. arXiv:1907.11692.

Images

- **Images provided by OpenClipart-Vectors/154119/Pixabay**. (https://pixabay.com/vectors/international-project-world-154119 [last access: 16.16.2021])
- **Coursebook (NLP and Computer Vision)**: IU International University of Applied Sciences. Version No.: 001-2020-1211.

REFERENCES

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

Literature

- **Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2018)**: *Unsupervised Cross-lingual Representation Learning at Scale*. arXiv:1911.02116.
- **Burrows, S., Gurevych, I., & Stein, B. (2014)**: *The Eras and Trends of Automatic Short Answer Grading*. In: IJAIED 25, 60–117.
- **Camus, L. & Filighera, A. (2020)**: *Investigating Transformers for Automatic Short Answer Grading*. AIED. Cyberspace.
- **Sawatzki, J., Schlippe, T., & Benner-Wickner, M. (2021)**: *Deep Learning Techniques for Automatic Short Answer Grading: Predicting Scores for English and German Answers*. In: AIET, Wuhan, China.
- **Schlippe, T. & Sawatzki, J. (2021b)**: *Cross-Lingual Automatic Short Answer Grading*. In: AIET, Wuhan, China.
- **Wölfel, M. (2021)**: *Towards the Automatic Generation of Pedagogical Conversational Agents from Lecture Slides*. In: EAI ICMTEL. Cyberspace.
- **Reyes, R., Garza, D., Garrido, L., De la Cueva, V., & Ramirez, J. (2019)**: *Methodology for the Implementation of Virtual Assistants for Education Using Google Dialogflow*. In: MICAI. Xalapa, Mexico.
- **Setiaji, H. & Paputungan, I.V. (2017)**: *Design of Telegram Bots for Campus Information Sharing*. In: ICITDA. Yogyakarta, Indonesia.
- **Hasan, K.S. (2014)**: *Automatic Keyphrase Extraction: A Survey of the State of the Art*. In: ACL. Baltimore, Maryland, USA.

REFERENCES

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

Literature

- **Merrouni, Z.A., Frikh, B., & Ouhbi, B. (2020)**: *Automatic Keyphrase Extraction: A Survey and Trends*. In: JIIS. vol. 54, pp. 391–424.
- **Chandrasekaran, D. & Mago, V. (2021)**: *Evolution of Semantic Similarity - A Survey*. arXiv: 2004.13820.
- **Wu, Y., Schuster, M., Chen, Z., Le, Q.V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, L., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., Stevens, K., Kurian, G., Patil, N., Wang, W., Young, C., Smith, J., Riesa, J., Rudnick, A., Vinyals, O., Corrado, G., Hughes, M., & Dean, J. (2016)**: *Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation*. CoRR abs/1609.08144.
- **Aiken, M. (2019)**: *An Updated Evaluation of Google Translate Accuracy*. Studies in Linguistics and Literature 3, 253.
- **Luo, H., Robinson, A. & Park, J.-Y. (2014)**: *Peer Grading in a MOOC: Reliability, Validity, and Perceived Effects*. Online Learning: Official Journal of the Online Learning Consortium. 18. 1-14. 10.24059/olj.v18i2.429.
- **Wölfel, M., Schlippe, T., & Stitz, A. (2015)**: *Voice Driven Type Design*. In: SpeD. Bucharest, Romania.
- **Schlippe, T., Alessai, S., El-Taweel, G., Wölfel, & M., Zaghouani, W. (2020)**: *Visualizing Voice Characteristics with Type Design in Closed Captions for Arabic*. In: Cyberworlds. Caen, France.