



COLOSS B-RAP Expert Evaluation of Beekeeping Advice from ChatGPT, Part 2

Lotta Fabricius Kristiansen, Linde Morawetz, Robert Chlebo, Alison Gray, Ewa Mazur, Bjørn Dahle, Ana Diéguez-Antón, Georg Schaunitzer, Kristina Gratzner, Hannes Oberreiter, Dirk Louis P. Schorkopf, Johann Fischer & Robert Brodschneider

To cite this article: Lotta Fabricius Kristiansen, Linde Morawetz, Robert Chlebo, Alison Gray, Ewa Mazur, Bjørn Dahle, Ana Diéguez-Antón, Georg Schaunitzer, Kristina Gratzner, Hannes Oberreiter, Dirk Louis P. Schorkopf, Johann Fischer & Robert Brodschneider (28 Oct 2024): COLOSS B-RAP Expert Evaluation of Beekeeping Advice from ChatGPT, Part 2, Bee World, DOI: [10.1080/0005772X.2024.2418224](https://doi.org/10.1080/0005772X.2024.2418224)

To link to this article: <https://doi.org/10.1080/0005772X.2024.2418224>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 28 Oct 2024.



[Submit your article to this journal](#)



Article views: 88



[View related articles](#)



[View Crossmark data](#)

COLOSS B-RAP Expert Evaluation of Beekeeping Advice from ChatGPT, Part 2

Lotta Fabricius Kristiansen , Linde Morawetz , Robert Chlebo , Alison Gray , Ewa Mazur , Bjørn Dahle , Ana Diéguez-Antón , Georg Schaunitzer , Kristina Gratzner , Hannes Oberreiter , Dirk Louis P. Schorkopf , Johann Fischer and Robert Brodschneider 

The artificial intelligence (AI) chatbot or advanced language model (LLM) ChatGPT is capable of understanding and generating human-like text. This article is a sequel to our first article on this topic, in which we imagined an AI tool that would serve as a beekeeping advisory tool in the future after having been trained with information stemming from high quality public domain resources on beekeeping and bee diseases (Morawetz et al., 2024). In that article, we presented the first part of the COLOSS B-RAP AI challenge, in which we asked ChatGPT-3.5 for advice regarding the serious honey bee disease *American Foulbrood*, which is a notifiable disease in many countries. We concluded that the tested AI chatbot version is a good teacher for basic or general beekeeping knowledge. It summarizes common knowledge about bee diseases at a level of detail that can be found in general beekeeping textbooks or in a beginners' course in beekeeping. As a summarization tool, it was, for example, used to write the recently published book on bee diseases by Walker (2024) to provide a comprehensive and structured overview of the topic. Nonetheless, it was concluded by the expert panel that advice from ChatGPT should be avoided for critical situations needing immediate action by the beekeeper, such as notifiable or severe diseases (Morawetz et al., 2024). Worryingly, part of the advice given was identified as misleading, wrong, or of limited usefulness in the given context and may easily lead to the spread of notifiable diseases and to severe consequences for the beekeeper and the beekeeping neighbors.

In this follow-up paper, we now present further results from the session held at the

COLOSS core project B-RAP held in Olomouc, Czechia, in February 2024 (Fabricius Kristiansen et al., 2022) concerning questions submitted to AI chatbots regarding specific beekeeping practices. Hence, we discuss those ChatGPT-3.5-generated answers that were evaluated during the event by a panel of 13 present experts (researchers, beekeeping advisers, veterinarians), many of them beekeepers themselves. Here we challenged ChatGPT with a very practical question on queen excluders that is commonly directed towards beekeeping advisers in Germany (see Table 1). Additionally, as the capability of novel AI chatbot versions is evolving extremely fast, we included in this second article the answers of three additional systems: ChatGPT-4o and two ChatGPT-based variations created explicitly for beekeeping questions: Wizbee (RAG (retrieval-augmented generation) integration with public resources about beekeeping; www.btree.at) and Beetrice (ChatGPT-3.5 and Meow App; beezum.online). Thus, Wizbee and Beetrice utilized high-quality beekeeping databases to retrieve and generate their answers.

The ratings of the answers' quality from the different AI tools are shown in Table 1 and were divided into answer topics/categories (for full answers see Supplementary Material). All four AI tools tested gave a mixture of correct advice and incorrect or incomplete advice. Furthermore, we found no difference in advice quality between the two general AI chatbots and the two chatbots specifically designed to answer beekeeping questions. Wizbee and ChatGPT-4o both gave more complete answers compared to the other two tools.

Comparing the older version ChatGPT-3.5 with the more recent version ChatGPT-4o (full access fee required), we conclude that the latest version provides a more complete and correct answer compared with the older version. Beetrice did not provide answers that were as detailed as the other tools gave and suggested some very basic and unhelpful information that did not contribute to solving the problem raised in the question. However, it was the only tool to suggest, as a final piece of advice, to contact a local beekeeping expert for advice if the suggested solutions did not work. The two AI tools utilizing beekeeping databases did not provide a more complete or more action-oriented answer than the two general AI tools.

In an agricultural study by De Clercq et al. (2024) it was found that while AI chatbots in the style of ChatGPT hold promise for improving agricultural efficiency, fostering innovation, and guiding informed policy decisions, several challenges must be addressed, most importantly the spread of misinformation (Chowdhury et al., 2023). The fast-paced development of LLM technology emphasizes the importance of agricultural policy makers crafting thoughtful frameworks and guidelines to ensure the responsible integration of LLMs in farming practices. Addressing these issues pre-emptively is crucial before these technologies become deeply entrenched, making policy intervention more difficult (De Clercq et al., 2024).

Our findings for the tested AI Chatbots regarding the variability in quality of advice for more specific topics and handling in practical beekeeping are broadly consistent with studies conducted

Table 1. Summary of the outputs by the four tested AI chatbots and experts' rating of these outputs for the question: "In May in Germany, a honey bee colony does not pass the queen excluder to the honey supers. What can I do?"; Rating of the given information: C = correct and valuable information, sometimes minor details are missing, U = correct, but useless information for the question, M = correct, but important details are missing, I = incorrect information, — = topic not discussed (for detailed outputs and opinions see Supplementary Material).

Advise (shortened version)	ChatGPT-3.5	ChatGPT-4o	Wizbee	Beetrice AI
Check the placement of the queen excluder	C	—	—	U
Ensure that the queen excluder is clean, so that bees can pass without problems	—	C	C	—
Remove the queen excluder temporarily	—	C (I)	C (I)	—
Ensure strong nectar flow	M	C	C (I)	U
Check if brood nest is too small	I	—	I	C
Check if brood nest is congested because of surplus of honey	—	C	I	—
Add frames of drawn combs or honey/nectar to the honey supers	C	C	C	—
Add frames with brood to the honey supers	—	M (I)	—	—
Monitor colony health	C	C	—	—
Check for queen presence	—	—	C	—
Ensure proper hive ventilation	—	I	—	—
Monitor the colony closely after the changes	—	—	C	—
Be patient and monitor the progress	M	M	—	M
If no solution works, consult a local beekeeping expert	—	—	—	C

(I): this method works but may have problematic consequences in your further beekeeping depending on the bees' behavior and your beekeeping methods.

for highly specific topics and methods in a wide range of other disciplines and professional settings where ChatGPT or other AI chatbots have been tested (e.g., Cappellani et al., 2024; Choi et al., 2024; Sandmann et al., 2024).

It is noteworthy that only one of the four chatbots (Beetrice) named and explained the most common reason (as seen by our expert panel) for the given problem correctly: a small brood nest combined with sufficient space in the brood box, which gives the bees no reason to move upwards. It is to be expected that a competent beekeeping advisor would ask a first question about the size of the brood nest before even thinking about less commonly encountered reasons. This example shows a central limitation of the tested AI chatbots—not asking follow-up questions to clarify exactly what the beekeeper wishes to know or to retrieve details of the situation the questions relate to to assess the beekeeper's needs more precisely (Figure 1).

All described contexts may be given to the AI chatbot by a person experienced in prompt-engineering—that is providing the AI chatbot with the necessary input for answering the question correctly. However, a layperson would be confronted with ChatGPT giving answers inadequate for the situation, area, or beekeeping skill. The



Figure 1. AI generated pictures from Bing Image. It is based on the following prompt: 'A beekeeper finds problem in the hive and ask an AI beekeeping advisor'. Citation: Fabricius Kristiansen L.

beekeeper can ask ChatGPT more detailed or specific questions—however, that additionally requires sound beekeeping knowledge to detect flaws and gaps in the provided answer.

While AI tools such as ChatGPT are potentially more useful and more likely to be used by less knowledgeable and/or less experienced beekeepers, in practice not all beekeepers, particularly this less knowledgeable group, may be able to cope with the complexity of interacting with AI chatbots to get accurate and useful answers. Tools are needed that contain some general prompts (e.g., which sources to use, which audience to expect) and/or that are able to ask the user key questions to form their own prompts optimally. Interfaces should provide some basic information; e.g., where do you keep your bees? How long have you been a beekeeper? Do you produce organic honey?

In summary, the AI tools tested provide guidance on the posed beekeeping question and show some promise as a quick way of getting an answer on what to do. However, they require further development to be useful and reliable for beekeepers in practice.

Disclosure statement

HO is owner of the b.tree app which includes wizbee. No other potential conflict of interest was reported by the author(s).

Supplementary material

Supplementary content is available via the “Supplementary” tab on the article’s online page (<https://doi.org/10.1080/0005772X.2024.2418224>).



References

- Cappellani, F., Card, K. R., Shields, C. L., Pulido, J. S., & Haller, J. A. (2024). Reliability and accuracy of artificial intelligence ChatGPT in providing information on ophthalmic diseases and management to patients. *Eye*, 38(7), 1368–1373. <https://doi.org/10.1038/s41433-023-02906-0>
- Choi, J., Kim, J. W., Lee, Y. S., Tae, J. H., Choi, S. Y., Chang, I. H., & Kim, J. H. (2024). Availability of ChatGPT to provide medical information for patients with kidney cancer. *Scientific Reports*, 14(1), 1542. <https://doi.org/10.1038/s41598-024-51531-8>
- Chowdhury, A., Kabir, K. H., Abdulai, A.-R., & Alam, M. F. (2023). Systematic review of misinformation in social and online media for the development of an analytical framework for agri-food sector. *Sustainability*, 15(6), 4753. <https://doi.org/10.3390/su15064753>
- De Clercq, D., Nehring, E., Mayne, H., & Mahdi, A. (2024). Large language models can help boost food production, but be mindful of their risks. *arXiv:2403.15475*. <https://doi.org/10.48550/arXiv.2403.15475>
- Fabricius Kristiansen, L., Kristiansen, P., Vejsnæs, F., & Morawetz, L. (2022). Is COLOSS an ivory tower of beekeeping science? Efforts to bridge research and practice (B-RAP). *Bee World*, 99(1), 5–7. <https://doi.org/10.1080/0005772X.2021.1993612>
- Morawetz, L., Fabricius Kristiansen, L., Brus, J., Dahle, B., Danihlík, J., Diéguez-Antón, A., Fischer, J., Gray, A., Mazur, E., Schaunitzer, G., Brodschneider, R. (2024). COLOSS B-RAP expert evaluation of beekeeping advice from ChatGPT, Part I. *Bee World*, 101(1), 2–4. <https://doi.org/10.1080/0005772X.2024.2348305>
- Sandmann, S., Riepenhausen, S., Plagwitz, L., & Varghese, J. (2024). Systematic analysis of ChatGPT, Google search and Llama 2 for clinical decision support tasks. *Nature Communications*, 15(1), 2050. <https://doi.org/10.1038/s41467-024-46411-8>
- Walker, P. G. (2024). *AI in the apiary* © Honeybee diseases (p. 198). Northern Bee Books.


Lotta Fabricius Kristiansen*

Department of People and Society,
Swedish University of Agricultural
Sciences, Uppsala, Sweden
*Email: lotta.fabricius@slu.se
Lotta Fabricius Kristiansen  <http://orcid.org/0000-0002-6551-2133>


Linde Morawetz and Dirk Louis P. Schorkopf

Department for Apiculture and Bee
Protection, Austrian Agency for Health
and Food Security,
Vienna, Austria
Dirk Louis P. Schorkopf  <http://orcid.org/0000-0003-0664-3050>
Linde Morawetz  <http://orcid.org/0000-0002-6771-4660>


Robert Chlebo

Institute of Animal Husbandry, Slovak
University of Agriculture in Nitra,
Nitra, Slovakia
Robert Chlebo  <http://orcid.org/0000-0001-8715-0578>

Alison Gray

Department of Mathematics and
Statistics, University of Strathclyde,
Glasgow, UK
Alison Gray  <http://orcid.org/0000-0002-6273-0637>


Ewa Mazur

Laboratory of Bee Diseases, Institute of
Veterinary Medicine, Warsaw University
of Life Sciences, Warsaw, Poland
Ewa Mazur  <http://orcid.org/0000-0002-3606-6082>



Bjørn Dahle

Norwegian Beekeepers Association,
Kløfta, Norway
Bjørn Dahle  <http://orcid.org/0000-0002-5131-2405>


Ana Diéguez-Antón

Department of Vegetal Biology and Soil
Sciences, Faculty of Sciences, University
of Vigo, Ourense, Spain
Ana Diéguez-Antón  <http://orcid.org/0000-0001-6142-2512>


Georg Schaunitzer and Robert Brodschneider

Department of Biology,
University of Graz,
Graz, Austria
Robert Brodschneider  <http://orcid.org/0000-0002-2535-0280>
Georg Schaunitzer  <http://orcid.org/0009-0003-1338-3469>

Kristina Gratzner

Department of Biology, University of
Graz, Graz, Austria
Kristina Gratzner  <http://orcid.org/0000-0002-2338-3708>

Hannes Oberreiter

Independent Researcher, Graz, Austria
Hannes Oberreiter  <http://orcid.org/0000-0003-1019-1245>

Johann Fischer

Institute for Bee Research and
Beekeeping, Veitshöchheim, Germany