abstract non-atomic sc-agent of automated knowledge extraction from technical documentation

⇒ *decomposition of abstract sc-agent:**:

{• *abstract sc-agent of document preparation*

• *abstract sc-agent of standard description*

• *abstract sc-agent of ontology description*

• *abstract sc-agent of LLM communication*

• *abstract sc-agent of LLM responce parsing*

}

Figure 2. Decomposition of the abstract non-atomic sc-agent of automated knowledge extraction

The sc-agent of document preparation is responsible for the preliminary preparation of the document and determining the type of standard to which it belongs. The document is converted into text form. This process may be accompanied by the use of image analysis tools (if images are present in the document)

The sc-agent of standard description is capable of returning a description of the standard to which the document in question was defined. It conveys information about the specific organization of text in documents of a particular standard.

The sc-agent of ontology description agent is capable of returning a description of the ontology of the subject area (for example, “probabilistic technological production processes”): a set of concepts, their description and a description of the connections between them.

The sc-agent of LLM communication is responsible for generating a prompt from information collected from previous agents and sending it as a request to the LLM.After sending, the agent expects to receive a response. The agent is not tied to a specific implementation or location of a large language model.

The sc-agent of LLM response parsing is the key element responsible for checking the correctness of the response received from the large language model (LLM). This agent checks the LLM response for compliance with the data format and consistency with the domain ontology. Its functionality is necessary to account for possible situations of "hallucination"or errors that may occur in the operation of the knowledge extractor-LLM. If the answer is successfully verified, the extracted information can be used to populate the knowledge base. Given the complex nature of the task that this agent needs to solve it may operate also using an LLM.

V. An example of extracting information from documentation using LLM

An example of the information extraction approach can be demonstrated using the project documentation for the PLCnext testbed [22] (Fig. 3, 4, 5).

The documentation contains extensive text on software settings for working with the stand (Fig. 6). Using the described prompt engineering, you can use LLM (ChatGPT 3.5 [17]) to structured knowledge extraction for each configuration step in JSON format, where each instruction step is correctly separated into a separate element (Fig. 7 and 8).

Figure 3. General view of the stand

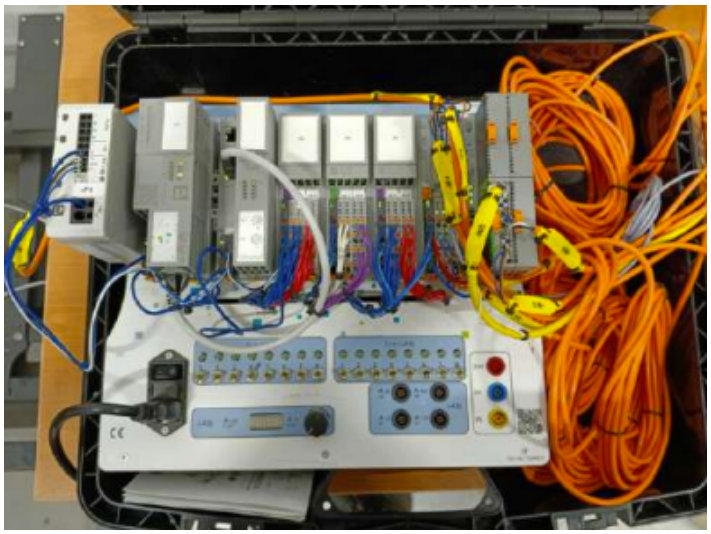

Figure 4. Controller and input/output node of the stand

VI. Conclusion

This paper proposes an approach to solving an important practical problem - automated knowledge extraction for

constructing enterprise knowledge bases, based on the application of an agent-based approach within the framework of OSTIS Technology and the use of large language models. The logic of operation of the corresponding component of the OSTIS Ecosystem is presented, which allows you to systematize and retrieve information from various data sources. The described approach can be applied not only in the field of enterprises, but also in other subject areas where work with standardized documentation is required as a source for populating knowledge bases.

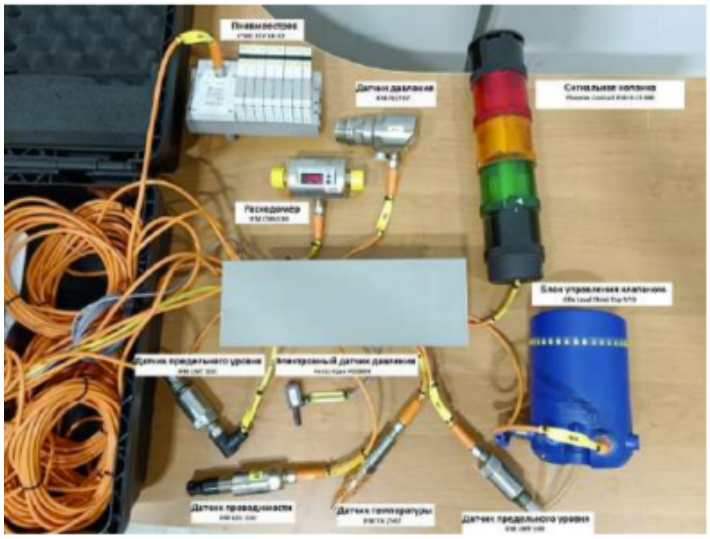

Figure 5. Sensors and actuators

```
# Тестовый стенд **"T1_PLCNext_Demo"**

| Тегов | Строк кода | Аппаратов | Агрегатов | Устройств | IO-Link I/O | Узлов | Модулей |
|---|---|---|---|---|---|---|---|
|![](statistics/tags_total.svg) |![](statistics/lines_total.svg)
|![](statistics/units_total.svg)
|![](statistics/agregates_total.svg)
|![](statistics/devices_total.svg)
|![](statistics/io_link_usage.svg)
|![](statistics/io_couplers_total.svg)
|![](statistics/io_modules_total.svg)

## Внешний вид демостенда

![общий вид](images/demo_stand.jpg)
<p align="center"> Рисунок 1 - Демостенд</p>

## Содержание

Запуск в режиме эмуляции (эмулятор "на столе").

1. [Клонирование проекта](#клонирование-проекта)
2. [Установка Visual Studio Code](#установка-visual-studio-code)
3. [Установка Git LFS](#установка-git-lfs)
4. [Настройка DbEditorXML](#настройка-dbeditorxml)
5. [Установка и настройка EasyServer](#установка-и-настройка-easyserver)
6. [Установка и настройка Monitor](#установка-и-настройка-monitor)
7. [Запуск и сборка проекта](#запуск-и-сборка-проекта)
8. [Взаимодействие с физическим устройством Bus Coupler](#взаимодействие-с-физическим-устройством-bus-coupler)

Запуск на контроллере.
```

Figure 6. Original view of the document text

```
{
    "Запуск проекта через консоль": [
        "Установить Toolchain нужной версии (подробная инструкция по установке здесь).",
        "Подключиться к контроллеру через клиентское программное обеспечение WinSCP.",
        "Перейти в каталог для своего контроллера ptusa_main\\bin\\install\\AXCF3152\\Release\\bin\\ и оттуда скопировать исполняемый файл ptusa_main в свой каталог на контроллере, например, /opt/main/.",
        "Далее в этот же каталог скопировать файлы с расширением lua и plua, а также директории spec и sys из проекта T1-PLCnext-Demo.",
        "После чего открыть Kitty, перейти в каталог со всеми скопированными файлами:",
        "cd /opt/main/",
        "И запустить исполняемый файл ptusa_main:",
        "./ptusa_main main.plua sys_path ./sys/",
        "В случае возникновения ошибки Permission denied, добавить файлу флаг -x, чтобы сделать его исполняемым. Это можно реализовать либо через интерфейс WinSCP, либо через консоль с помощью команды chmod.",
        "chmod +x ptusa_main"
    ]
}
```

Figure 7. Structured Information Retrieval

```
{
    "Подключение к контроллеру в программе PLCnext Engineer": [
        "Переходим на вкладку с настройками контроллера.",
        "Далее нажимаем на кнопку подключения.",
        "Далее вводим имя пользователя (admin) и пароль (напечатан на корпусе контроллера). После успешного подключения отображаются текущие параметры контроллера."
    ],
    "Обновление библиотеки": [
        "Скачать файл последней версии библиотеки (*.PCWLX).",
        "Скопировать данный файл с заменой в папку PLCnext необходимого проекта.",
        "Откройте клиентское программное обеспечение SFTP (например WinSCP).",
        "Авторизуйтесь как администратор (admin - обратитесь к администратору).",
        "Копировать файл с заменой файла библиотеки из /AXCF2152_22.0.4.144/Release/lib/libPtusaPLCnextEngineer.so в каталог /opt/plcnext/projects/PCWE/Libs/Ptusa, где AXCF2152_22.0.4.144 - соответствующая версия библиотеки.",
        "Откройте командную оболочку с помощью инструмента командной строки (например, KiTTY).",
        "Авторизуйтесь как администратор (admin - обратитесь к администратору).",
        "Перезапустите службу plcnext (команда: sudo /etc/init.d/plcnext restart)."
    ]
}
```

Figure 8. Structured Information Retrieval

References

[1] V. Golenkov, Ed., Tehnologija kompleksnoj podderzhki zhiznennogo cikla semanticheski sovmestimyh intellektual'nyh komp'juternyh sistem novogo pokolenija [Technology of complex life cycle support of semantically compatible intelligent computer systems of new generation]. Bestprint, 2023.

[2] C. D. Manning, H.Schütze Foundations of statistical natural language processing, 1999, MIT press.

[3] V. Taberko, D. Ivaniuk, V. Smorodin, V. Prokhorenko Adaptive Control System for Technological Process within OSTIS Ecosystem,

Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2023, pp. 291-298.

[4] V. Taberko, D. Ivaniuk, N. Zotov, M. Orlov, O. Pupena, N. Lutska Principles Of Building A System For Automating The Activities Of A Process Engineer Based On An Ontological Approach Within The Framework Of The Industry 4.0 Concept, Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2021,pp. 209-218.

[5] "ISA5.1 Standard," https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa5-1, (accessed 2024, Jan)

[6] "ISA-88 standard," Available at: https://www.isa.org/isa88, (accessed 2024, Jan).

[7] "ISA-95 standard," https://www.isa.org/standards-andpublications/isa-standards/isa-standards-committees/isa95, (accessed 2024, Jan)

[8] J. Lee, B. Choi A survey of standardization efforts for serviceoriented architecture, Expert Systems with Applications, 2009, 36(4), pp.7844-7855.

[9] M. Salama, S.K. Mostefaoui, A. Nour A survey on IoT interoperability and standardisation issues, International Journal of Ad Hoc and Ubiquitous Computing, 13(3), 145-155.

[10] V. Taberko, D. Ivaniuk, D. Shunkevich, O. Pupena Principles For Enhancing The Development And Use Of StandardsWithin Industry 4.0, Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2020, pp. 167-174.

[11] A. Radford, K.Narasimhan, Improving Language Understanding by Generative Pre-Training, 2018

[12] J. Devlin, M. Chang, K. Lee, K. Toutanova, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, 2019 North American Chapter of the Association for Computational Linguistics.

[13] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, Dario Amodei Language models are few-shot learners. arXiv preprint arXiv:2005.14165, 2020

[14] D. Xu, Wei C., W. Peng, C. Zhang, T. Xu, X. Zhao, X. Wu, Y. Zheng, E. Chen Large Language Models for Generative Information Extraction: A Survey, 2020, arXiv preprint arXiv:2005.14165.

[15] A. Zagorskiy Integration Of Third-Party Functional Services On A Unified Semantic Basis, Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2023, pp. 207-212.

[16] A. Cherkas, A. Kupo Ostis Technology Integration With ThirdParty NLP Services, Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems], 2023, pp. 207-212.

[17] "Introducing ChatGPT", Available at: https://openai.com/blog/chatgpt, (accessed 2023, January).

[18] P.Kohli, R.Gupta, S.Saxena, Robustness Challenges in Language Models: A Survey. arXiv preprint arXiv:2103.01021, 2021

[19] J. Dodge, S.Gururangan, D.Card, R.Schwartz, N. A.Smith, S. R.Bowman Fine-tuning language models from human preferences. arXiv preprint arXiv:1909.08593, 2020

[20] A. Holtzman, J. Buys, J. Du, M. Forbes, Y.Choi, The curious case of neural text degeneration. arXiv preprint arXiv:1904.09751, 2020

[21] Y. Liu, M.Ott, N. Goyal, J. Du, M.Joshi, D.Chen, V. Stoyanov, Roberta: A robustly optimized BERT pretraining approach, arXiv preprint arXiv:1907.11692, 2019

[22] "T1-PLCnext-Demo", Available at: https://github.com/savushkinr-d/T1-PLCnext-Demo, (accessed 2023, January).

NLP ПОДХОД К ПОСТРОЕНИЮ БАЗЫ ЗНАНИЙ ПРЕДПРИЯТИЯ С ИСПОЛЬЗОВАНИЕМ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ

Таберко В. В., Иванюк Д. С.
Смородин В. С., Прохоренко В. А.

Предложен подход к автоматизированному формированию базы знаний о предприятии на основании анализа существующей технической документации с применением больших языковых моделей. Данный подход позволяет обеспечить возможность интеграции предлагаемого решения с другими разработками, программными средствами предприятия для обеспечения построения интеллектуальных систем автоматизированного управления, рекомендательных систем и систем поддержки принятия решений, систем информационного обеспечения персонала предприятия.

Received 13.03.2024