ARTIQ - AI Centres of Excellence

Application for a Host Institution

Institution	National Centre for Research and Development, National Science Centre
Project Joint National Project:	ARTIQ – AI Centres of Excellence
Deadline for the submission of applications	8th of April-11th of May 2021

I. HOST INSTITUTION DATA

Identification data of the Host Institution

Name (full)	***Łukasiewicz Research Network– Institute of Innovative Technologies EMAG***
Name (short)	***Łukasiewicz-EMAG***
Name of the main organisational unit (where applicable)	***386888354***
Address of the registered office	
Street	Leopolda
Building No.	*31*
Office No.	-
Postal code	*40-189*
City/district	*Katowice*
Post office	*Katowice*
Municipality	*Katowice*
County	*Katowice*
Province	*Śląskie*

EPUAP [Electronic Platform for Public Administration Services] mailbox	**/ibEMAG/SkrytkaESP**
Legal form	***Research Institute***
The person appointed for contact with NCBR and with the potential Leader/Project Manager	
First name	***Marek***
Last name	***Sikora***
Position	***Area leader***
Phone number	***32/2007 805***
E-mail address	***marek.sikora@emag.lukasiewicz.gov.pl***
The person authorised to represent the applicant	
First name	***Artur***
Last name	***Kozłowski***
Function/Position	***Director***

II. CAPACITY OF THE HOST INSTITUTION TO PERFORM THE PROJECT

1. Description of major research achievements in the scope of implementation of R&D projects, as well as the commercialisation of deliverables of such projects regarding artificial intelligence for the last 5 years prior to or in the year of the application along with a list of the most important publications and patents of the applicant (max. 1 A4 page).

The last three years were a transformation period for the Łukasiewicz-EMAG Institute, which had been an institute providing solutions for the mining industry so far but changed its profile completely during that time. Now its main research streams include cyber security and artificial intelligence (more information available in section 5). The Institute has performed a number of AI-related projects and has launched co-operation with many scientists working in this domain, both in the institutes grouped in the Łukasiewicz Research Network (ILIM, PIARP), universities (Silesian University of Technology, Jagiellonian University in Kraków, University of Warsaw, Bournemouth University (UK)) and from commercial companies (3 Soft, QED Software).

Łukasiewicz-EMAG's AI-related projects (listed in section 2) concern a wide range of AI issues. Technical and industrial projects have focused on the development of AI/ML platforms for on-line monitoring of processes and natural hazards and for predictive maintenance (PdM) tasks. Currently, Łukasiewicz-EMAG has a complex AI/ML platform for hazards prediction and PdM. Medical projects are mainly focused on the treatment of childhood acute lymphoblastic leukemia. In these projects, Łukasiewicz-EMAG was responsible for the development of a data collecting application (including genetic data collection) and analytical applications for data fusion and reasoning (including survival analysis of patients). In the range of cyber security, Łukasiewicz-EMAG has gained experience in the development of hybrid solutions for detecting attacks, based on domain (expert) knowledge and data stream analysis. In addition, recently the Institute has been running projects in the areas of image processing and natural language processing.

Several Łukasiewicz-EMAG's employees have been working on doctoral theses on AI-related topics. Please find below selected publications by the members of Łukasiewicz-EMAG's AI/ML team:

1. Kozielski M., Henzel J., Wróbel Ł. Łaskarzewski Z., Sikora M.: A sensor data-driven decision support system for LPG suppliers. Applied Sciences 1198), 3474 (70 p. IF 2.474)
2. Gudyś A., Sikora M., Wróbel Ł.: RuleKit: A Comprehensive Suite for Rule-Based Learning. Konwledge Based Systems 194, 1-2, 2020 (200 p. IF 5.101).
3. Sikora M., Gudyś A., Wróbel Ł.: GuideR: A guided separate-and-conquer rule learning in classification, regression, and survival settings. Knowledge Based Systems 173, 1-14, 2019 (200 p. IF 5.358).
4. Ślęzak D., Grzegorowski M., Janusz A., Kozielski M., Nguyen S.H., Sikora M., Stawicki S., Wróbel Ł.: A framework for learning and embedding multi-sensor forecasting models into a decision support system: A case study of methane concentration in coal mines. Information Sciences 451-452, 112-133, 2018 (200 p. IF 5.524).
5. Wróbel Ł., Gudyś A., Sikora M.: Learning rule sets from survival data. BMC Bioinformatics 18(1): 285:1-285:13 2017 (100 p. IF 2.448).
6. Gruca A., Sikora M.: Data and expert-driven rule induction and filtering framework for functional interpretation and description of gene sets. Journal of Biomedical Semantics 8(1), 1-14, 2017 (70 p. IF 1.582).
7. Janusz A., Grzegorowski M., Michalak M., Wróbel Ł., Sikora M., Ślęzak D.: Predicting seismic events in coal mines based on underground sensor measurements. Engineering Applications of Artificial Intelligence 64:83-94, 2017 (140 p. IF 3.526).

2. A list of 5 research and development projects within national and international competitions in the area of artificial intelligence and implemented within the last 5 years prior to or in the year of the application (title, manager, source of financing, amount of financing) (max. 1 A4 page).

Projects:

- Avatar – Polish Sign Language automatic translator with the use of an avatar – the project concerns entirely AI. It deals with automatic translation of texts (particularly formal texts) into the Polish Sign Language. Project value: 7,989,437.50 PLN, financing source: National Centre for Research and Development (NCBR), project manager: Adam Piasecki. The project is under way.
- DISESOR – Integrated framework for supporting decisions for process-, device- and hazard monitoring systems – the project concerned entirely AI. It dealt with the development of a platform which used domain knowledge and AI/ML methods to monitor production processes and natural hazards in the mining industry. Project value: 3,463,087 PLN, financing source: National Centre for Research and Development (NCBR), project manager: Marek Sikora. The project has been partially commercialized.
- PersonALL – Personalization of childhood acute lymphoblastic leukemia treatment in Poland – Łukasiewicz-EMAG's role in the project was to develop a data collection system and an analytical platform which enabled to integrate clinical and genetic data, to perform data exploration processes in the range of survival analysis, and to identify strong, unknown and interesting dependencies in the data. Project value: 13,738,483 PLN, financing source: National Centre for Research and Development (NCBR), project manager: Marek Sikora. The project has been continued in the form of clinical, non-commercial research CAL-POL financed by Medical Research Agency. Łukasiewicz-EMAG has been developing and maintaining the applications developed within the PersonALL project. The CAL-POL research is planned to be completed in 2026.
- RegSoc – Regional Cyber Security Centre – Łukasiewicz-EMAG's role in the project is to develop – in co-operation with Wrocław University of Science and Technology – a library to identify outliers in data streams describing web traffic. The library is to be used in the Regional Security Operation Centre for small and middle-size companies. The centre is planned to be launched this year. The project makes use of known AI/ML methods dedicated to such tasks. Project value: 12,779,290 PLN, financing source: National Centre for Research and Development (NCBR), project manager: Andrzej Białas.

- POMPY – Developing a prototype of the system for current measurement of the degree of pulverized coal dust in pulverized coal boilers – AI/ML methods were used here to develop algorithms whose aim was to determine the degree of coal dust grinding based on a set of measurement data. The developed algorithm became the basis for a patent application. The process of commercializing the project results is under way. Project value: 1,186,136 PLN, financing source: National Centre for Research and Development (NCBR), project manager: Jarosław Smyła.

It is also worth mentioning that between 2013 and 2016 there were three projects (Coal&Steel Fund) related to the prediction of fires in mines, prediction of methane hazards and coal dust explosion hazards. All these projects made use of thresholding methods based on AI/ML, including stream data analysis.

3. Available research equipment, apparatus/infrastructure and intangible assets held in the context of implementation of a project regarding artificial intelligence (max. 1 A4 page).

As far as research infrastructure is concerned, we base on our own solutions and cloud solutions (in the case when one-time sources are needed to fulfill analytical tasks requiring high computing power).
Our own infrastructure includes the following:

- Dell PowerEdge M820 server, 2 Intel Xeon E5-4617 processors (2.90GHz, 6 cores, 128GB RAM, 1600MHz 146GB SAS
- PowerEdge M640, Intel Xeon Gold 6130, 12x32 RDIMM 2666MHz, 2x200GB
- Dell PowerEdge M610 Blade Server, Intel Xeon E5620, 32GB for 2 processors (8x4GB in two-bank RDIMM modules) 1333MHz, 2x146GB SAS
- DELL PowerEdge M710, 2 x XEON X5660 2.8Ghz 6 cores, 64GB DDR3 RDIMM (8 x 8GB), 2 x 146 GB HotPlug SAS
- DELL PowerEdge M710, 2 x XEON X5660 2.5Ghz 6 cores, 64GB DDR3 RDIMM (8 x 8GB), 2 x 146 GB HotPlug SAS
- DELL PowerEdge R610 server, 2 x 6-core INTEL XEON E5649 2,53GHZ, 64GB DDR3, RDIMM (8 x 8GB) 2 x 146 GB HotPlug SAS
- DELL PowerEdge R610 server, 2 x 6-core INTEL XEON E5649 2.53GHZ, 64GB DDR3, RDIMM (8 x 8GB) 2 x 146 GB HotPlug SAS
- DELL PowerEdge T110, 4-core Intel Xeon X3430 2.40GHZ, 8GB (2X4GB), 4 x 500GB SATA
- DELL PowerEdge M520 server, Intel Xeon E52440 (6 x 2.4ghz, 15MB Cache), 8 x 16GB(1333Mhz) 8GB SD card
- PowerEdge M630 DELL, Intel Xeon E5-2640 v3 2.6GHz, 8x16GB RDIMM, 2133MT, with two 2GB SD cards

Cluster: (used chiefly for the development of distributed processing architectures)

- 5 x server with 8 x Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz (16 Logical processors) 160GB RAM
- 4 x server with 8 x Intel(R) Xeon(R) Silver 4110 CPU @ 2.10GHz (16 Logical processors) 128GB RAM

Additionally, Łukasiewicz-EMAG has at its disposal a full range of platforms facilitating the implementation of software projects, such as Confluence, BitBucket, Jira, Enteprise Architect, full-licence MS Teams, Sharepoint, DataStorage – with a total capacity of 80TB.
Łukasiewicz-EMAG has implemented an information security policy, back-up policy and business continuity policy.

All employees have computer equipment at their own disposal, mostly high-class notebooks. In addition, each employee has at least one extra computer screen. As far as AI/ML software is concerned, we base on open-source solutions, which is now a standard procedure in this domain. In the range of system software and tools, we have an MSDN agreement with Microsoft.

4. Facilities or incentives to establish an AI Centre of Excellence in the entity (max. 1 A4 page).

In order to establish an AI Centre of Excellence on its premises, Łukasiewicz-EMAG can provide office space – renovated and equipped with modern computer appliances – as well as conference space with equipment necessary to run trainings and conferences (stationary, on-line and hybrid). Some office space will be arranged for the growing AI/ML team, some for people co-operating with Łukasiewicz-EMAG within concrete AI projects. Two conference rooms (one big, one medium-size) will host working meetings, conferences and trainings.

5. Other information concerning internationalisation of the entity, foreign scientists employed in this institution, availability of seminars in English, etc. (max. 1 A4 page).

Łukasiewicz-EMAG developed wide-ranging co-operation with international research institutions engaged in the mining industry. Between 2000-2015 there were several projects completed, mainly within the Coal&Steel programmes. The co-operation concerned partners from Germany, UK and Spain. In the AI domain the Institute relies on direct contacts (Bournemouth University) or finds partners indirectly, e.g. through the Polish Artificial Intelligence Society and commercial companies. In addition, Łukasiewicz-EMAG has applied to join new EU initiatives – we will attempt to participate in the EdgeAI programme. All these activities aim to widen the range of undertaken projects but also, more importantly, to increase the number of direct contacts with international researchers.

6. Other significant information confirming the experience and resources of the institution (max. 1 A4 page).

Since 2020 Łukasiewicz-EMAG has been improving its AI/ML competences intensively. The Institute has entered into co-operation with the Polish Artificial Intelligence Society (as a supporting partner). Łukasiewicz-EMAG's employees have been participating in AI working groups operating under the auspices of the Ministry of Economic Development, Labour and Technology. Łukasiewicz-EMAG is a partner in a consortium which has submitted a grant application within the Digital Innovation Hub (in the AI area). Since

2020 three doctoral programmes have been opened in the AI domain (one of them within the Doctoral School of Silesian University of Technology, the other two within the implementation doctorates programme of this school).

Over 30 Łukasiewicz-EMAG's employees are competent in AI. This team includes scientists, data scientists, software developers, and testers. Łukasiewicz-EMAG's staff had a chance to participate in trainings on such issues as times series analysis, containerization and project management.
New AI-related projects performed in Łukasiewicz-EMAG include Avatar (see section 2), SPINET and BUMECH. In the latter one, AI/ML methods will be applied in on-line diagnosis of a developed vehicle. In addition, several projects are waiting for evaluation – they concern fake news detection (Infostrateg) and medical images analysis.

Łukasiewicz-EMAG operates within EU's third largest network of research institutes – Łukasiewicz Research Network. The Łukasiewicz-EMAG Institute is a member of the Digital Transformation Research Group within this network. The Group consists of 12 Łukasiewicz institutes. The offer presented in this application also includes access to resources (equipment and human resources) of institutes working in the group, particularly the Industrial Research Institute for Automation and Measurements (Łukasiewicz-PIAP) and the Institute for Logistics and Storage (Łukasiewicz-ILIM). Łukasiewicz-PIAP is a leading institute in the domain of special-purpose autonomous vehicles. Łukasiewicz-ILIM is a leader in logistics, storage and optimization (including supply chains optimization). Both institutes conduct research on the application of AI solutions in these areas. The Digital Transformation Research Group includes 12 institutes out of 32 united in the whole Łukasiewicz network.