NEXT GENERATION INTERNET
INTERNET OF HUMANS

# NGI SEARCH – D1.1 WORK PROGRAMME AND ADVISORY BOARD

| | |
|---|---|
| Start date of project: | 02/09/2022 |
| WP Leader | UMU |
| WP n° and title | WP1 Ecosystem |
| Document Type | Report |
| Dissemination level | Pubic |
| Due date | 30/11/2022 |
| Delivery date | 28/11/2022 |
| Responsible Author(s) | Aurora González Vidal (UM), Mirko Presser (AU) |
| Contributors (s) | |
| Version | V1.0 |

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101069364

# TABLE OF CONTENTS

| TABLE OF LIST OF ABBREVIATIONS AND DEFINITIONS | |
|---|---|
| OC | Open Call |
| NGI | Next Generation Internet |
| NLU | Natural Language Understanding |
| AR/VR | Augmented Reality / Virtual Reality |
| NLP | Natural Language Processing |
| IoT | Internet of Things |

# 1 Work Programme Open Call 1

NGI Search looks for proposals to address deep-tech development and research-based solutions that address NGI at the core of their developments whilst supporting the outcomes of the above-mentioned vision and mission, as well as following the principles. The following list of topics is a set of problems that the consortium has identified upfront under the topic of Search and Discovery for the first OC, which welcomes start-up entrepreneurs, tech geeks, developers, socially engaged people and researchers with a focus on technical developments.

We expect that all developments will be open source and provide real demonstrations either as research demonstrators showcasing novel concepts; minimum viable products showcasing a potential business; and/or community projects that are in a deployable state. We will not only accept Open Source, but also Open Core is a possibility if the developments within the project work on the core. All developments are expected to be well documented and uptake of the open source code is widely encouraged.

## 1.1 The next generation of intelligent voice-based assistants

### Description

Voice assistants have been successfully adopted for simple, routine tasks, such as asking for the weather or making a phone call. They can have a social function since they can decrease depression and simulate interest in physical activity among other positive social influences[1]. However, as familiarity increases, we will start demanding more complex tasks, such as exploratory search. Advice on starting a new hobby, suggestions on books on a certain topics or activities to do abroad are examples of exploratory search queries that cannot be answered with a single-shot answer, and this will require the addition of intelligence to the assistants[2]. Traditional written query formulations are usually simpler than voice-based searches and analysing the word choices and interactions provides more context about the intent of the user. Some challenges relate to situationally induced impairments and security, since the search can be done while performing other activities and in public spaces (the assistant demands private information), to queries that mix languages (multilingual) and to mixed modal interactions, where questions and answers not only have voice content but also images, text etc.

### NGI search motivation

The Search and Discovery that is done through voice assistants is a resource that many people use nowadays, making their life easier by accessing information and services. We seek ambitious projects to overcome the various technical challenges that are associated with NGI values such as privacy and trust in search retrieval, global language support by including and creating open source training data for minority languages, and advancing the field of NLU.

### Keywords

voice assistants, voice search, security in voice search, contextual search, induced impairments, mixed modal interactions, image search, image query, search query, textual search, NLU in search, cross lingual search, immersive technology (AR/VR) in search, inclusive search.

---

1 Kachouie, R., Sedighadeli, S., Khosla, R., & Chu, M. T. (2014). Socially assistive robots in elderly care: a mixed-method systematic literature review. *International Journal of Human-Computer Interaction*, *30*(5), 369-393.

2 Ma, X., & Liu, A. (2020, July). Challenges in Supporting Exploratory Search through Voice Assistants. In Proceedings of the 2nd Conference on Conversational User Interfaces (pp. 1-3).

## 1.2 Natural language processing

### Description

Natural language processing (NLP) methods are widely in use in production in the area of machine translation. The majority of search engines use the huge amounts of previously accumulated user requests for predicting the search output without taking into account the user's intention[3]. Model complexity of the current state-of-the-art models is increasing and implies the use of large amounts of energy for computation. At the same time, the state-of-the-art models assume that each device would have full access to powerful processors, generous memory, and, generally, cloud connectivity. This may not be possible for many edge devices. We want to initiate the paradigm of tinyNLP by searching for ways to adapt NLP methods to edge and fog computing, studying ways to apply transfer learning in NLP and any other methodologies that can improve the energy efficiency of the current NLP approaches towards simpler and more sustainable NLP research and practices.

### NGI Search motivation

NLP models enable a semantic search yielding optimized search results. NLP has the power to deal with searches that use the meaning of whole sentences. However, methods need to be adapted to the device where they are used, its resources and context in order to be seamlessly integrated into real systems, and for that reason, we look for simpler and more sustainable decentralized NLP practices that are not gender biased.

### Keywords

NLP, energy efficient NLP, deep learning NLP, federated NLP, edge computing NLP

## 1.3 Semantic analysis

### Description

Semantic data integration generates a common representation of concepts and their relations using domain knowledge formalisms in the form of ontologies and reasoning capabilities and therefore can aid in the integration of heterogeneous data[4]. Information about a subject or topic might be spread across different data sources so there exists the need for the integration of knowledge derived from large-scale human and machine-built repositories. Question answering and data analytics can make use of such knowledge which in turn can be applied for decision-making, which should be many times based on near real-time data. The next steps in the semantic analysis field could be to search dynamic relations between concepts using "fresh" data and to minimize query execution while maximizing answer completeness based on federated principles. Federated query processing techniques integrate data from autonomous, distributed, and heterogeneous data sources in a uniform way[5].

### NGI Search motivation

Engines for knowledge discovery that can perform searches that return contextualized, organized and dynamic (on-the-go) enriched results by means of semantic analysis are motivated by this topic. We look for semantic privacy protection implementations to improve trust that are resilient,

3 Chernyshov, A., Balandina, A., & Klimov, V. (2018, August). Intelligent processing of natural language search queries using semantic mapping for user intention extracting. In Biologically Inspired Cognitive Architectures Meeting (pp. 56-61). Springer, Cham.

4Al-Lahham, A. (2020). *Ontology-based context-aware model for event processing in an IoT environment* (Dissertation, University of Salford).

5Diego Collarana Vargas (2018) Strategies and Techniques for Federated Semantic Knowledge Retrieval and Integration (Doctoral dissertation, Universitat Bonn)

i.e. have no limits with regard to semantics parameters such as frequency of words or complexity of structured data.

**Keywords**

Federated query processing, semantic analysis, semantic search, knowledge graphs, dynamic semantic search

## 1.4 Social Computing

**Description**

The development of technologies that require interaction with humans imposes an interesting challenge since they have to succeed in improving motivation, encouraging participation and enhancing the learning process for their success. The interaction between social behaviour and technologies needs to be addressed in order to reach substantial changes in the behaviour of the adopters[6]. Human-related data presents big data characteristics and therefore, edge social computing should be considered in these scenarios in order to process and filter data of the network to reduce bandwidth costs, storage and energy consumption[7]. The implementation of edge social computing by means context-aware learning, collaborative learning and other proposals in this direction are encouraged.

**NGI Search motivation**

This topic is motivated by the improvement of search and discovery of relations using multiple types and sources of social information and social interaction with machines. Several technical challenges can be addressed with regards to data sharing for collaborative learning, context awareness, etc. so as to create a collective intelligence without geographical or temporal limitations encouraging collective decision-making, learning, online collaboration, etc. building a strong, vibrant and inclusive internet community.

**Keywords**

Social computing, context-aware learning, collaborative learning, edge social computing, human computing, collective intelligence

## 1.5 Data visualization

**Description**

Data visualization has attracted much attention recently, calling for joint actions in different research fields such as information visualization, human-computer interaction, machine learning, data management and mining, and computer graphics[8]. We seek interactive tools and mechanisms that allow visualizations for machine learning results that can provide user recommendations and support user-driven actions. This includes new applications of visually driven analysis of spatiotemporal, textual and other kinds of data, (such as, but not limited, data analysis within an immersive environment) progressive visualizations (in batches) and other kinds of scalable and efficient solutions[9,10].

6García, Ó., Alonso, R. S., Prieto, J., & Corchado, J. M. (2017). Energy efficiency in public buildings through context-aware social computing. *Sensors*, *17*(4), 826.

7Sittón-Candanedo, I., Alonso, R. S., García, Ó., Muñoz, L., & Rodríguez-González, S. (2019). Edge computing, iot and social computing in smart energy scenarios. *Sensors*, *19*(15), 3353.

8Andrienko, G., Andrienko, N., Drucker, S., Fekete, J. D., Fisher, D., Idreos, S., ... & Sharaf, M. (2020, March). Big data visualization and analytics: Future research challenges and emerging applications. In *BigVis 2020-3rd International Workshop on Big Data Visual Exploration and Analytics*.

### NGI Search motivation

The visual navigation of data and of the results of queries in order to find detect patterns, gain insight, and answer highly specific questions using scalable and efficient data visualization techniques is the motivation of this topic. NGI Search will support visualization solutions that help understand results provide a more inclusive setup for technological engagement and trust.

### Keywords

Information visualization, computer graphics, visualization, graphs, animated, infographics, spatiotemporal visualization, textual visualizations, immersive (VR/AR) data visualization

## 1.6 Enabling new ways of discovering and accessing information

### Description

Due to the rapid development of the Internet of Things (IoT) and consequently, the variability of more and more data sources, mechanisms for searching and integrating data become essential to leverage all relevant knowledge for improving processes and services[11]. New ways of discovering and accessing information need to be created in the form of platforms and products that deal algorithmically with data. The integration of data-driven machine learning with human knowledge (common priors or implicit intuitions) can effectively lead to explainable AI[12] that would provide us ways to discover and access information where only raw data is present. The Challenge is to develop new algorithms and methodologies to discover and access information by combining Big Data technologies.

### NGI Search motivation

This topic gathers any kind of techniques that cannot be classified in the prior topics but yet it is interesting to create new ways of discovering and accessing information with regard to nowadays challenges associated with data and computing towards open technologies that respond to people's fundamental needs on search and information retrieval.

### Keywords

Information visualization, computer graphics, visualization, graphs, animated, infographics

# 2 Advisory Board

The Advisory Board is a group of experts in the field and issues addressed by the NGI Search project. They are also potential users of projects funded by NGI Search. Experts come from a variety of backgrounds including academia, industries, and SDOs. We are pleased to have such a

9 Silva, N., Blascheck, T., Jianu, R., Rodrigues, N., Weiskopf, D., Raubal, M., & Schreck, T. (2019, June). Eye tracking support for visual analytics systems: foundations, current applications, and research challenges. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications* (pp. 1-10).

10 Caillou, P., Renault, J., Fekete, J. D., Letournel, A. C., & Sebag, M. (2020). Cartolabe: A Web-Based Scalable Visualization of Large Document Collections. *IEEE Computer Graphics and Applications*, *41*(2), 76-88.

11 Iggena, T., Bin Ilyas, E., Fischer, M., Tönjes, R., Elsaleh, T., Rezvani, R., ... & Holmgård Christophersen, S. (2021). IoTCrawler: Challenges and Solutions for Searching the Internet of Things. *Sensors*, *21*(5), 1559.

12 Zhuang, Y. T., Wu, F., Chen, C., & Pan, Y. H. (2017). Challenges and opportunities: from big data to knowledge in AI 2.0. Frontiers of Information Technology & Electronic Engineering, 18(1), 3-14.

diversity of profiles, as their collective feedback will guide NGI Search contributors and steer the community in the right direction.

**Payam Barnaghi** - Professor of Machine Intelligence Applied to Medicine

at Imperial College London; Keywords: W3C, machine data (social media, IoT, images); machine learning; deep tech research.

**Anja Bechmann** – Professor at Aarhus University; Keywords: Social Media, Social science and humanities research.

**Simone Fischer-Hübner** - Professor at Kalstad University on Privacy and Security; Keywords: IFIP; Security, GDPR, privacy, research, open source.

**Bob Goudriaan** – Director of NLNet (NGI0 Search and Discovery); Keywords: Community, Open Source.

**Michael Granitzer** – Chair of Data Science at University of Passau and coodinator of openwebsearch.eu; Keywords: Data science, Open Web Index, deep tech research.

**Marie-Jose Montpetit** - Digital Technologies - Ferme d'hiver / Winter Farm; Keywords IRTF and IETF, open source, semantic web, distributed systems, core-edge continuum, open platforms, research.

**Josiane Parreira** - Senior Research Scientist at Siemens AG Austria; Keywords: Vertical Use Cases, energy, semantic data, linked open data, data streams, machine learning, applied research, industry.

**Stefan Voigt** – Open Search Foundation, participant to openwebsearch.eu; Keywords: Community, Open Source.

# Advisory Board

**Payam BARNAGHI**
*Professor of Machine Intelligence applied to Medicine*
Imperial College, London
Academia
UK

**Anja BECHMANN**
*Professor of Social Sciences*
Aarhus University
Academia
Denmark

**Simone FISCHER-HÜBNER**
*Professor of Computer Science*
Karlstad University
Academia
Sweden

**Bob GOUDRIAAN**
*General Director*
NLNet Foundation
Entrepreneurship, Consulting
Netherlands

**Michael GRANITZER**
*Chair of Data Sciences*
Passau University
Academia
Germany

**Marie-Jose MONTPETIT**
*Affiliate Professor*
Concordia University
Academia, SDOs
Canada

**Josiane Xavier PARREIRA**
*Senior Research Scientist*
Siemens AG
Industry
Austria

**Stefan VOIGT**
*Chief Executive Officer*
Open Search Foundation
Industry
Germany

Figure 1 Advisory Board