Human Understanding Capabilities of Symbiotic AI systems in FAIR (Future Artificial Intelligence Research)

Pierluigi Amodio[1], Linda Antonella Antonucci[2], Pierpaolo Basile[3], Marco de Gemmis[3], Felice Iavernaro[1], Pasquale Lops[3], Francesca Mazzia[3], Cataldo Musto[3], Marco Polignano[3,*], Giovanni Semeraro[3], Lucia Siciliani[3], Cristiano Tamborrino[3] and Paolo Taurisano[2]

[1]University of Bari Aldo Moro, Dept. of Mathematics, Bari, 70125, Italy
[2]University of Bari Aldo Moro, Dept. of Translational Biomedicine and Neuroscience, Bari, 70125, Italy
[3]University of Bari Aldo Moro, Dept. of Computer Science, Bari, 70125, Italy

Abstract

Nowadays, we need AI systems capable of engaging and working with people, perceiving and acting within changing contexts, being aware of their limits and adapting to new scenarios, behaving correctly in complex social settings, being aware of their security and trust perimeters, and of being aware of the environmental and societal effect that their implementation and execution may imply. In summary, we require a formerly undiscovered AI. Symbiotic-AI systems should disclose human cognitive capabilities to improve the effectiveness of information access and decision-making. This is achieved by combining methods for determining who is interacting with the system and how. The former includes the definition of strategies for acquiring and exploiting users' personal information gathered by combining different strategies and heterogeneous sources. In contrast, the latter includes detecting and interpreting human signals acquired from various sources, such as advanced machine learning (particularly deep learning) and natural language processing.

Keywords

Artificial Intelligence, Future Research, FAIR, Symbiotic-AI

1. Introduction

Future Artificial Intelligence Research (FAIR) is the Italian AI research community's answer to the National Strategic Program. FAIR accepts the challenge of defining the research agenda for tomorrow's AI approaches and techniques. Despite AI's progress, its acceptance has been primarily in low-risk applications, with medium/high-risk applications, such as healthcare, public administration, and safety-critical industries, still lagging behind forecasts. This is why fundamental, interdisciplinary research is required to shape future AI. FAIR would like to work on future types of artificial intelligence: 1. human-centered: co-evolve with humans "in-the-loop" both individually and collectively; 2. integrative: a link between various AI approaches; 3. resilient: the ability to operate in real-world environments; 4. adaptive: the ability to respond in rapidly changing circumstances; 5. high-quality: safety-critical applications; 6. symbiotic: encourages efficient human-machine contact and collaboration; 7. edge/exascale: operates on the edge and on the cloud; 8. pervasive: function ubiquitously; 9. environmentally conscious: consider the environmental dimension; 10. bio-cognitive and sustainable: imitate the concepts of biological systems at multiple scales. This work introduces the groundwork topics of Symbiotic-AI (SAI), as well as its three primary goals: 1) enabling SAI system design using HCI principles and methodologies; 2) enabling SAI development using both data-driven and model-driven approaches to endow AI systems with human understanding capabilities and improve AI systems' performance with user input; 3) enabling SAI applicability by improving AI system understandability.

2. Symbiotic-AI

The research is founded on the novel concept that AI augments (and values) human cognitive talents rather than replacing them and hence supports and facilitates human activities [1]. In terms of human comprehension, we investigate strategies for learning who is engaging with the system and comprehending the semantics of user signals obtained from various sources using deep/machine learning and Natural Language Processing (NLP). In terms of AI system understandability, we will look into

Ital-IA 2023: 3rd National Conference on Artificial Intelligence, organized by CINI, May 29–31, 2023, Pisa, Italy
*Corresponding author.
pierluigi.amodio@uniba.it (P. Amodio); linda.antonucci@uniba.it (L. A. Antonucci); pierpaolo.basile@uniba.it (P. Basile); marco.degemmis@uniba.it (M. d. Gemmis); felice.iavernaro@uniba.it (F. Iavernaro); pasquale.lops@uniba.it (P. Lops); francesca.mazzia@uniba.it (F. Mazzia); cataldo.musto@uniba.it (C. Musto); marco.polignano@uniba.it (M. Polignano); giovanni.semeraro@uniba.it (G. Semeraro); lucia.siciliani@uniba.it (L. Siciliani); cristiano.tamborrino@uniba.it (C. Tamborrino); paolo.taurisano@uniba.it (P. Taurisano)

human mental models and their key ability to deal with ambiguity and imprecision in decision-making processes. The activities will take two paths: i) Explainable AI for Symbiotic-AI [2]; (ii) methodologies for assessing and conveying uncertainty and imprecision. Research on the acceptability of Symbiotic-AI will use an interdisciplinary approach, with academics from AI, Law, and Ethics participating. The long-term viability of Symbiotic-AI is explored from two viewpoints: model recycling to conserve computing resources for model learning and tuning and data collection, labeling, and representation effort reduction.

2.1. Data curation and ingestion

This research task involves creating, organizing, and maintaining large volumes of data or corpora that are used to train and test large language models. The process of data curation and ingestion is a complex and time-consuming task, but it is necessary to ensure the quality of the model's output and to support explainability, transparency, and interoperability [3, 4, 5]. Additionally, data security and privacy need to be taken into account, such as gaining participants' consent, anonymizing or de-identifying data, and ensuring that it is securely kept and managed. Data curation and intake are critical components of developing Large Language Models (LLMs) that can effectively understand and respond to human language [6]. This is especially crucial to allow for diversity in terms of the languages and topics covered [7]. Recognizing possible biases in the data curation and ingestion processes and taking proactive steps to remove them is critical. Data security, ethics, and privacy must be considered, especially when working with sensitive data. It is critical that we protect people's privacy rights and prevent data exploitation [8].

2.2. Combination of exogenous and endogenous semantics

This research task involves a combination of exogenous and endogenous semantics, utilizing multimodal embeddings and generative tasks through attentive, self-supervised training and large-scale-based transfer learning [9]. Exogenous semantics refers to external information that is used to train the model, while endogenous semantics refers to internal information that is subsequently learned by the model from the training data. Multidimensional embeddings involve representing words and sentences using multiple modalities, allowing the model to capture different facets of language, leading to improved performance [10]. Linguistic knowledge can be added into embeddings to increase the representations' quality [11]. This can lead to better results in tasks like sentiment analysis [12, 13, 14, 15, 16, 17], machine translation, and question answering. Multimodal embeddings may be constructed using a variety of methods, including fusion-based and alignment-based approaches. The input from several modalities is integrated to generate a single embedding vector in fusion-based approaches. Deep learning architectures like CNNs and RNNs are common approaches for constructing multimodal embeddings. Multimodal embeddings have been employed in a variety of applications, including picture captioning, visual question answering, and multimodal sentiment analysis [18]. They allow the models to reason across multiple modalities and give a more thorough grasp of the underlying data.

2.3. Training, fine-tuning, and prompting of Large Language Understanding Models

The study of strategies for training, fine-tuning, and motivating Large Language Models is part of this research task. It is critical to train adaptive LLMs, which can learn and alter their knowledge depending on new data and feedback [19, 6]. This must be done while keeping multimodality in mind. The task goal is to solve this issue and examine options that may be available for less frequently spoken languages. Fine-tuning and prompting are two approaches often employed in Big Language Understanding, a subset of Natural Language Processing. These strategies are used to boost the performance of pre-trained language models like GPT-3 for specific tasks or domains. Fine-tuning entails picking a smaller dataset relevant to the job and then training the model on it. Except for the last layer, the weights of the pre-trained model usually are frozen. Fine-tuning can help with various NLP tasks, including sentiment analysis, question answering, and text categorization. Prompting, on the other hand, is sending a specific prompt or input to a pre-trained language model in order for it to respond. The prompt might be a single brief word or sentence that offers context for the model to respond to. The purpose of prompting is to increase the model's relevance and accuracy. Task-specific prompts can be created for writing a news story summary or answering a specific question.

2.4. Exploiting LLMs for intelligent information access

LLMs may be used to improve a variety of information-access applications, including information extraction, question answering, information search and seeking discovery, decision-making, and recommendation. One way of extracting information from large language models is named entity recognition (NER) [20], while another is relationship extraction (RE). Large language models have shown tremendous potential in the field of information extraction and are expected to play an increasingly important role in automating structured information extraction. Large language models, such as GPT-3 and BERT, have revolutionized quality assurance by achieving cutting-edge performance on a wide range of benchmarks and real-world applications. They can also be used with other tools and technologies to create more robust information search and retrieval systems. Large language models can aid in discovering new information and insights from data. They can aid in identifying patterns and insights that traditional data analysis methods may overlook, providing tailored suggestions and insights to help users make decisions, and assisting decision-makers in making intelligent and data-driven decisions [21]. They may also be used in finance to analyze market data and estimate stock prices or investment opportunities, in healthcare to review patient data, and in natural language processing applications to evaluate consumer comments, social media postings, and other forms of communication [22].

2.5. Functional connectivity patterns

This project aims to identify different functional connectivity patterns associated with different aspects of higher-order cognitive functions, with particular attention to the different stages of memory processing. During functional magnetic resonance imaging (fMRI) scanning, participants will perform a memory task between two resting state sessions. To investigate whether different participants recruit different thalamic hubs as a function of memory performance, the percentage of overlap between each individually-segmented thalamic subdivision will be extracted from individual spatial maps of a frontoparietal network. Multiple support vector regression models will be used to validate and extend the method to reach the goal to whole-brain functional connectivity beyond the thalamus and to other cognitive functions beyond memory.

2.6. User Modeling

In this task, we will create techniques for automatically eliciting information about users' preferences and requirements from disparate data sources. Initially, we will look into how such data points may be extracted from conversational data. This data, often available in natural language format, is essential for modeling what a user wants and requires when interacting with an intelligent system. Unfortunately, obtaining this information is not easy. Hence approaches for exact elicitation are becoming increasingly important. Similarly, we will look into how Knowledge Graphs (KGs) and LLMs might be used for this purpose [10]. Regarding KGs, we will look into ways to automatically map qualities that represent preferences and requirements to KG components. In this approach, reasoning and learning mechanisms may be used to infer more general and accurate information. Lastly, LLMs may be utilized to describe preferences and requirements accurately. Conversational data, KGs, and LLMs can represent information sources. In terms of conversational data, the massive volume of information already accessible may be utilized to forecast more accurate information about the user [23, 24]. For instance, the psychological attributes of the user (i.e., personality) can be learned and inferred from such a data point [13]. There are now various ways to predict such qualities automatically, but the concept of analyzing conversational data for these purposes remains unexplored. Similarly, sources such as KGs and LLMs might be valuable for learning about the characteristics of users. For example, by using a user's preferences and requirements stored as an embedding, it is feasible to fine-tune models that allow reliable prediction of such attributes [25]. All information about what users desire, as well as their characteristics, will be utilized to give customized information access to them. In this situation, conversational data will also be crucial in learning new strategies and models [26, 27, 28].

2.7. Data preprocessing with outlier/anomaly detection

Outliers are frequently caused by measurement mistakes or data corruption, but in some instances, the source of the outliers is unknown. During the statistical analysis of the dataset, a robust automated approach for detecting outliers turns out to be critical. Novel techniques based on statistical regression processes and using the idea of entropy to detect outliers contextually when fitting a model to data will be addressed. The approach, already developed for polynomials [29], will be extended to the B-spline basis and nonlinear models and will be used to remove noise from signals generated either synthetically or by specific sources of interest.

Noise reduction, imputing, and smoothing of long time series will also be accomplished by using numerical approaches based on B-spline quasi-interpolation [30]. In general, quasi-interpolation refers to a method for constructing efficient local approximants to a given data collection or function. There are two primary reasons for using quasi-interpolation. In practice, data acquired from natural settings are frequently contaminated by mistakes. Hence, interpolation methods may be overly rigorous, in addition to suffering from the overfitting issue. Furthermore, since quasi-interpolation is based on local construction, the computational cost is significantly lower when compared to global alternatives such as interpolation. The same methodologies will be examined, together with statistical methods, for a novel anomaly detection strategy in massive datasets. This data preprocessing phase is critical because it enables the use of higher-quality data as input to the algorithms that will be built.

2.8. Matrix and tensor decomposition

The frontier of machine learning and deep learning in Symbiotic-AI is to evaluate enormous amounts of data from many sources, such as text, photos, and signals from the IoT environment, to construct models capable of assisting human understanding, providing answers, knowledge, and synthesis. Nevertheless, sustaining such models is too expensive in terms of computational and data storage costs. Therefore, it remains fundamental to use techniques capable of extracting information and patterns from different types of data by reducing their dimensionality while preserving the most important information. This is in the sense of learning closer to the human, which requires little information to learn, build new knowledge, or provide answers. Within this task, new Matrix and Tensor decomposition techniques based on SVD will be investigated [31, 32, 33]. This is essential to obtain data in a low dimensional space that contains the salient information discarding those redundant or those that don't contain crucial information. These techniques turn out to be an important step when used for supervised and unsupervised learning processes. Indeed, numerous works in this field have shown that matrix and applied tensor decomposition in the task of classification, clustering, change detection, or salience detection algorithms are able to increase performance, not only from a computational point of view but also from the expected results [34, 35]. Lately, the Tensor and Matrix decomposition have been also used as a generative model to obtain synthetic data with the same characteristics as the real data. This is another aspect that can be taken into account, especially when there is a need to form a model, and there is not enough real data available.

2.9. Explore NN with a new family of deep neural network models

In the last decades, a wide range of neural networks has been proposed to address challenges in deep learning processes, achieving great success in a wide spectrum of applications. Despite the prominent success achieved by the neural network, this approach still lacks theoretical direction for designing an effective neural network model, and evaluating the performance of a model requires unnecessary resources. Recent research has shown that many existing models may be thought of as various numerical discretizations of differential equations [36, 37]. Deep learning is also making rapid progress in natural language processing (NLP), speech recognition, and computer vision. Deep neural networks' excellent function approximation capabilities are at the root of these achievements. Modeling sequential data is a difficult challenge in NLP applications [38, 39]. Much work has been done to overcome the challenge of sequential data modeling. Because of its recurrent nature, recurrent neural networks (RNNs) outperform the others. Notwithstanding the capabilities of the RNNs, training them is a challenging task. When the input sequence is large, the multiplication term generated by the chain derivation method is numerically unstable, causing RNN to suffer from gradient explosion and gradient vanishing. To train the model, backpropagation through time (BPTT) is required. In this challenge, we will investigate a novel family of deep neural networks that entail the usage of Ordinary Differential Equations (ODEs) inspired by the relationship between RNNs and ODEs (ODE). Specific models in the ODE class are worth considering, such as delay ordinary differential equations [40] and integrodifferential equations. Such delayed dynamics are significant in characterizing emergent phenomena in many real-world systems, including physical, chemical, ecological, and physiological systems. This study area will be examined in conjunction with data representation utilizing tensors and multiway arrays that account for the data structure. Moreover, the usage of decompositions to address data dimensionality will be discussed.

3. Conclusion

The FAIR approach is comprehensive and multidisciplinary, with the goal of profoundly rethinking the foundations of AI while also studying the societal consequences of emerging kinds of AI. Depending on the path that the AI revolution takes, AI will either empower or reduce human agency; expand or replace the human experience; create new forms of human activity or reduce jobs; help distribute well-being for many or increase the concentration of power and wealth in the hands of a few; expand or endanger democracy in our societies; help fight climate change or increase emissions. FAIR researchers want AI to be part of the solution, not the issue, to the world's social, financial, ethical, and environmental concerns. The merging of human intelligence with artificial intelligence (AI) to improve both capabilities is thought of as Symbiotic-AI. This form of collaboration has the potential to forge strong alliances by allowing people to profit from AI's speed, efficiency, and processing capacity while giving machines human-like judgment, reasoning, and creativity.

4. Acknowledgements

We acknowledge the support of the PNRR project FAIR - Future AI Research (PE00000013), Spoke 6 - Symbiotic AI (CUP H97G22000210007) under the NRRP MUR program funded by the NextGenerationEU.

References

[1] P. Brusilovsky, M. De Gemmis, A. Felfernig, P. Lops, M. Polignano, G. Semeraro, M. C. Willemsen, Joint workshop on interfaces and human decision making for recommender systems (intrs'22), in: Proceedings of the 16th ACM Conference on Recommender Systems, 2022, pp. 667–670.

[2] C. Musto, N. Tintarev, O. Inel, M. Polignano, G. Semeraro, J. Ziegler, Workshop on explainable user models and personalized systems (exum 2021), in: Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization, 2021, pp. 211–212.

[3] C. Musto, A. Delic, O. Inel, M. Polignano, A. Rapp, G. Semeraro, J. Ziegler, Workshop on explainable user models and personalised systems (exum), in: Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization, 2022, pp. 160–162.

[4] M. Polignano, V. Basile, P. Basile, G. Gabrieli, M. Vassallo, C. Bosco, A hybrid lexicon-based and neural approach for explainable polarity detection, Information Processing & Management 59 (2022) 103058.

[5] M. Polignano, G. Colavito, C. Musto, M. de Gemmis, G. Semeraro, Lexicon enriched hybrid hate speech detection with human-centered explanations, in: Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization, 2022, pp. 184–191.

[6] K. Florio, V. Basile, M. Polignano, P. Basile, V. Patti, Time of your hate: The challenge of time in hate speech detection on social media, Applied Sciences 10 (2020) 4180.

[7] M. Polignano, P. Basile, M. De Gemmis, G. Semeraro, V. Basile, et al., Alberto: Italian bert language understanding model for nlp challenging tasks based on tweets, in: CEUR Workshop Proceedings, volume 2481, CEUR, 2019, pp. 1–6.

[8] F. Lorè, P. Basile, A. Appice, M. de Gemmis, D. Malerba, G. Semeraro, An ai framework to support decisions on gdpr compliance, Journal of Intelligent Information Systems (2023) 1–28.

[9] P. Lops, C. Musto, M. Polignano, Semantics-aware content representations for reproducible recommender systems (score), in: Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization, 2022, pp. 354–356.

[10] M. Polignano, C. Musto, M. de Gemmis, P. Lops, G. Semeraro, Together is better: Hybrid recommendations combining graph embeddings and contextualized word representations, in: Proceedings of the 15th ACM Conference on Recommender Systems, 2021, pp. 187–198.

[11] D. Nozza, L. Passaro, M. Polignano, et al., Preface to the sixth workshop on natural language for artificial intelligence (nl4ai), in: CEUR Workshop Proceedings, 2022.

[12] M. Polignano, P. Basile, M. de Gemmis, G. Semeraro, A comparison of word-embeddings in emotion detection from text using bilstm, cnn and self-attention, in: Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization, 2019, pp. 63–68.

[13] M. Polignano, F. Narducci, M. de Gemmis, G. Semeraro, Towards emotion-aware recommender systems: an affective coherence model based on emotion-driven behaviors, Expert Systems with Applications 170 (2021) 114382.

[14] M. Polignano, P. Basile, M. De Gemmis, G. Semeraro, Hate speech detection through alberto italian language understanding model., in: NL4AI@ AI* IA, 2019, pp. 1–13.

[15] M. Polignano, M. de Gemmis, F. Narducci, G. Semeraro, Do you feel blue? detection of negative feeling from social media, in: AI* IA 2017 Advances in Artificial Intelligence: XVIth International Conference of the Italian Association for Artificial Intelligence, Bari, Italy, November 14-17, 2017, Proceedings 16, Springer, 2017, pp. 321–333.

[16] M. Polignano, P. Basile, G. Rossiello, M. de Gemmis, G. Semeraro, Learning inclination to empathy from social media footprints, in: Proceedings of the 25th conference on user modeling, adaptation and personalization, 2017, pp. 383–384.

[17] M. Polignano, P. Basile, Hansel: Italian hate speech detection through ensemble learning and deep neural networks, EVALITA Evaluation of NLP and Speech Tools for Italian 12 (2018) 224.

[18] M. G. de Pinto, M. Polignano, P. Lops, G. Semeraro, Emotions understanding model from spoken language using deep neural networks and mel-frequency cepstral coefficients, in: 2020 IEEE conference on evolving and adaptive intelligent systems (EAIS), IEEE, 2020, pp. 1–5.

[19] M. Polignano, M. de Gemmis, G. Semeraro, Contextualized bert sentence embeddings for author profiling: The cost of performances, in: Computational Science and Its Applications–ICCSA 2020: 20th International Conference, Cagliari, Italy, July 1–4, 2020, Proceedings, Part IV 20, Springer, 2020,

pp. 135–149.

[20] M. Polignano, M. de Gemmis, G. Semeraro, Comparing transformer-based ner approaches for analysing textual medical diagnoses., in: CLEF (Working Notes), 2021, pp. 818–833.

[21] M. Polignano, V. Suriano, P. Lops, M. de Gemmis, G. Semeraro, A study of machine learning models for clinical coding of medical reports at codiesp 2020., CLEF (Working Notes) 640 (2020).

[22] P. Basile, D. Croce, V. Basile, M. Polignano, et al., Overview of the evalita 2018 aspect-based sentiment analysis task (absita), in: EVALITA Evaluation of NLP and Speech Tools for Italian, CEUR, 2018, pp. 1–10.

[23] C. Musto, M. Polignano, G. Semeraro, M. de Gemmis, P. Lops, Myrror: a platform for holistic user modeling: Merging data from social networks, smartphones and wearable devices, User Modeling and User-Adapted Interaction 30 (2020) 477–511.

[24] C. Musto, G. Semeraro, M. Polignano, A comparison of lexicon-based approaches for sentiment analysis of microblog posts., in: DART@ AI* IA, Citeseer, 2014, pp. 59–68.

[25] C. Musto, F. Narducci, M. Polignano, M. de Gemmis, P. Lops, G. Semeraro, Towards queryable user profiles: Introducing conversational agents in a platform for holistic user modeling, in: Adjunct publication of the 28th ACM conference on user modeling, adaptation and personalization, 2020, pp. 213–218.

[26] C. Musto, F. Narducci, M. Polignano, M. De Gemmis, P. Lops, G. Semeraro, Myrrorbot: A digital assistant based on holistic user models for personalized access to online services, ACM Transactions on Information Systems (TOIS) 39 (2021) 1–34.

[27] M. Polignano, P. Lops, M. de Gemmis, G. Semeraro, Helena: An intelligent digital assistant based on a lifelong health user model, Information Processing & Management 60 (2023) 103124.

[28] M. Polignano, F. Narducci, A. Iovine, C. Musto, M. De Gemmis, G. Semeraro, Healthassistantbot: a personal health assistant for the italian language, IEEE Access 8 (2020) 107479–107497.

[29] D. Giordano, F. Iavernaro, Maximal-entropy driven determination of weights in least-square approximation, Mathematical Methods in the Applied Sciences 44 (2021) 6448 – 6461. doi:10.1002/mma.7197.

[30] A. Falini, F. Mazzia, C. Tamborrino, Spline based hermite quasi-interpolation for univariate time series, Discrete and Continuous Dynamical Systems - S 15 (2022) 3667–3688. URL: /article/id/621c8f822d80b7479e4357ab. doi:10.3934/dcdss.2022039.

[31] M. Vandecappelle, L. De Lathauwer, From multilinear svd to multilinear utv decomposition, Signal Processing 198 (2022) 108575. URL: https://www.sciencedirect.com/science/article/pii/S0165168422001128. doi:https://doi.org/10.1016/j.sigpro.2022.108575.

[32] L. Lathauwer, B. De Moor, A multi-linear singular value decomposition, Society for Industrial and Applied Mathematics 21 (2000) 1253–1278. doi:10.1137/S0895479896305696.

[33] T. G. Kolda, B. W. Bader, Tensor decompositions and applications, SIAM Review 51 (2009) 455–500. URL: https://doi.org/10.1137/07070111X. doi:10.1137/07070111X.

[34] A. Falini, G. Castellano, C. Tamborrino, F. Mazzia, R. M. Mininni, A. Appice, D. Malerba, Saliency detection for hyperspectral images via sparse-non negative-matrix-factorization and novel distance measures, in: 2020 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS), 2020, pp. 1–8. doi:10.1109/EAIS48028.2020.9122749.

[35] C. Tamborrino, F. Mazzia, Classification of hyperspectral images with copulas, Journal of Computational Mathematics and Data Science 6 (2023) 100070. URL: https://www.sciencedirect.com/science/article/pii/S277241582200030X. doi:https://doi.org/10.1016/j.jcmds.2022.100070.

[36] P. Kidger, On neural differential equations, 2022. arXiv:2202.02435.

[37] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, D. K. Duvenaud, Neural ordinary differential equations, in: S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, R. Garnett (Eds.), Advances in Neural Information Processing Systems, volume 31, Curran Associates, Inc., 2018. URL: https://proceedings.neurips.cc/paper_files/paper/2018/file/69386f6bb1dfed68692a24c8686939b9-Paper.pdf.

[38] M. Liu, W. Luo, Z. Cai, X. Du, J. Zhang, S. Li, Numerical-discrete-scheme-incorporated recurrent neural network for tasks in natural language processing, CAAI Transactions on Intelligence Technology (2023) n/a–n/a. doi:10.1049/cit2.12172.

[39] B. Li, Q. Du, T. Zhou, Y. Jing, S. Zhou, X. Zeng, T. Xiao, J. Zhu, X. Liu, M. Zhang, Ode transformer: An ordinary differential equation-inspired model for sequence generation, 2022, pp. 8335–8351. doi:10.18653/v1/2022.acl-long.571.

[40] F. Stelzer, S. Yanchuk, Emulating complex networks with a single delay differential equation, The European Physical Journal Special Topics 230 (2021). doi:10.1140/epjs/s11734-021-00162-5.