Future Artificial Intelligence Research (FAIR)

Proposta per un Partenariato Esteso (PE) nell'area tematica 1

Il CNR si è candidato come soggetto proponente per presentare una proposta per il PE1 "Intelligenza Artificiale: Aspetti Fondazionali", da realizzare in collaborazione con il Laboratorio Nazionale AIIS (Artificial Intelligence and Intelligent Systems) del CINI (Consorzio Inter universitario Nazionale per l'Informatica). La proposta, denominata "Future Artificial Intelligence Research (FAIR)", in linea con il Programma Strategico nazionale per l'Intelligenza Artificiale (AI) 2022-2024, si pone l'obiettivo di contribuire ad affrontare le domande di ricerca, le metodologie, i modelli, le tecnologie e anche le regole etiche e legali per costruire sistemi di AI capaci di interagire e collaborare con gli umani, di percepire e ad agire all'interno di contesti in continua evoluzione, di essere coscienti dei propri limiti e capaci di adattarsi a nuove situazioni, di essere consapevoli dei perimetri di sicurezza e fiducia, e di essere attenti all'impatto ambientale e sociale che la loro realizzazione ed esecuzione può comportare.

Come previsto dal bando, la proposta si basa sul modello Hub & Spoke.

L'Hub sarà una Fondazione di partecipazione.

Le attività di ricerca saranno svolte nell'ambito di 10 spoke che coinvolgeranno 350 ricercatori. Ciascuno spoke è caratterizzato da una specifica area tematica e da un proprio insieme di sfide di ricerca con l'obiettivo di affrontare le sfide di FAIR da diversi punti di osservazione. Gli spoke sono: *Human-centred* AI, *Integrative* AI, *Resilient* AI, *Adaptive* AI, *High-quality* AI, *Symbiotic* AI, *Edge-exascale* AI, *Pervasive* AI, *Green-aware* AI, *Sustainable Bio-socio-cognitive* AI. Si veda l'Appendice 1 per la lista degli Spoke, i referenti scientifici e i principali temi di ricerca affrontati da ciascuno spoke.

Gli spoke agiranno da catalizzatori sia dello sviluppo di tecnologie AI innovative che di nuovi servizi di AI in settori stategici per il sistema Paese, coinvolgendo il settore industriale, sia a livello delle grandi aziende che delle piccole e medie imprese innovative, utilizzando soprattutto il meccanismo dei bandi a cascata riservati agli enti/imprese che NON siano parte dell'Hub o degli Spoke.

Gli spoke avranno anche un ruolo di amplificatore delle opportunità di formazione in AI a tutti i livelli. In particolare, per l'alta formazione il partenariato FAIR collaborerà strettamente con il Dottorato Nazionale in Intelligenza Artificiale (www.phd-ai.it), coordinato da CNR, Campus Biomedico di Roma, Politecnico di Torino, Sapienza Università di Roma, Università di Napoli "Federico II", e Università di Pisa.

Coerentemente con la Strategia Nazionale Italiana in Intelligenza Artificiale, la proposta FAIR coniuga un approccio inclusivo e interdisciplinare con la valorizzazione delle eccellenze diffuse nella comunità scientifica nazionale in Intelligenza Artificiale.

ALLEGATO 1: AREE TEMATICHE DEGLI SPOKE

SPOKE 1: HUMAN-CENTERED AI – UNIVERSITÀ DI PISA [Rif: DINO PEDRESCHI]

Human-centered AI promises to expand our capacity to make high-stakes decisions and everyday choices, empowering humans both as individuals and as collectives, but requires a deep rethinking of the foundations of AI. Spoke 1 is committed to build the foundations of Human-centered AI along three main challenges:

- **"human-in-the-loop" machine learning and reasoning**: allowing humans to understand and steer learning and reasoning of AI systems and interact synergistically.
- **social-aware AI**: understanding and governing the societal outcomes of large-scale, networked socio-technical systems of humans and AIs, e.g., social media and online marketplaces.
- **responsible design of trustworthy AI systems**: methods for the responsible (co-)design, development, validation and use of trustworthy AI systems, including certification, to incorporate "by-design" European laws and ethical values (interdisciplinary links with ethics and moral philosophy, law, software engineering)

SPOKE 2: INTEGRATIVE AI – FBK [Rif: PAOLO TRAVERSO]

Current research in AI addresses with success specific tasks of a rather narrow scope (or complex problems that can be decomposed in simpler independent sub problems). In order to overcome the barrier of complexity, we need a substantial change in the approach to research in AI, i.e., we need research that goes beyond the current approach based on vertical and separated AI areas. We need an integrated approach, that we call **Integrative-AI**. We propose an overarching interdisciplinary approach oriented to theories and techniques that integrate different scientific methods, technologies, disciplines and competences. We will investigate the integration of heterogeneous computational representation and reasoning/learning (inference) techniques, such as the integration of data driven and model based approaches, symbolic and sub-symbolic representations.

SPOKE 3: RESILIENT AI – UNIVERSITÀ DI NAPOLI "FEDERICO II" [RIF. CARLO SANSONE]

The challenges posed to AI systems share a crucial aspect: the use of machine learning algorithms trained with real-world data that are unstructured, noisy, often incomplete/limited in number, and partially inconsistent. The achievement of adequate performance of the systems obtained in these fields cannot be separated from the study of specific AI foundational methodologies that are aimed at processing data in-the-wild, making the performance of AI resilient and robust in these challenging contexts.

The research activities to be carried out will include

- the definition of appropriate data augmentation techniques, when the data are incomplete or not adequately representative, while analyzing, monitoring, and improving the fairness of the machine learning algorithms
- to make algorithms both resilient and robust with respect to possible external attacks (also deriving from training with "malicious" data), thus also having to define how to measure these characteristics in an appropriate manner
- the investigation of the implications related to the design, validation & verification, evolution and operation of the software that implements machine or deep learning algorithms, when they have to work in-the-wild

SPOKE 4: ADAPTIVE AI – POLITECNICO DI MILANO [RIF: NICOLA GATTI]

Adaptivity has been one of the primary goals of AI since its beginning and is now pervasively required not only in the virtual world but also in the physical world. In particular, adaptivity is related to the capability of an entity to interact with the environment, perceive the context and the related information that changes in time, and promptly take actions. In particular, **machine learning** has emerged so far as the main key enabling technology **to design adaptive artificial systems**. Despite its ubiquitous adoption, there are serious gaps in our theoretical **understanding of how learning systems can provide guarantees and how they can effectively be combined with other AI paradigms** when designing adaptive artificial systems. Challenges to be addressed include: the foundations of learning theory for adaptivity, the understanding of nonconvex optimization for deep learning models (taking into account the interaction between the statistical and the algorithmic aspects of the training process), the study of notable environment structures that can be used to provide better adaptivity guarantees.

SPOKE 5: High Quality AI – SAPIENZA UNIVERSITÀ DI ROMA [RIF: GIUSEPPE DE GIACOMO]

AI techniques are currently deployed in the real world, however such applications are often limited to systems that are considered low-risk such as search engines and recommender systems. Recent advances in AI tools show their potential to be used in a variety of other applications, including safety-critical systems such as autonomous vehicles and high-cost industrial processes such as power plants, scientific experiments, handling of public administration data. Deploying AI systems in such high-stake applications demands algorithms and tools to meet high quality requirements. This line of research aims at characterizing the various **qualities of interest in AI-based systems**, their **measurability**, **assessment** and **certifiability**, across all the whole AI, including perception, actuation, learning, modeling, reasoning and planning.

SPOKE 6: Symbiotic AI – UNIVERSITÀ DI BARI [RIF: DONATO MALERBA]

Symbiotic AI investigates scientific, social, economic, legal and ethical challenges related to the growing symbiosis between humans and AI. As AI systems are becoming a part of our daily lives, the question of how to improve the current shortcomings and limitations in human-machine collaboration has taken on a new urgency. The primary challenge of an AI agent that is functioning alone is how effectively and flawlessly it achieves its goal. However, in a team of humans and AI assisting each other for a team goal, the challenges are not limited to the goal itself, since the AI system should also have the ability to not only reason about the human's actions but their mental models. Symbiotic AI promises to boost human-machine collaboration, with mutually beneficial relationships, by augmenting (and valuing) human cognitive abilities rather than replacing them.

SPOKE 7: Edge-Exascale AI – POLITECNICO DI TORINO [RIF: BARBARA CAPUTO]

The vast majority of modern AI is computationally demanding at all stages, from algorithm design and development, to training and deployment: designing new algorithms often resort to heuristics and long trial and errors; training them requires data processing on a very large scale, and guaranteeing a strong final performance is tightly linked to an extensive search and optimization for the best hyperparameters; This is true for all kinds of hardware supporting AI, from cloud and HPC with large scale centralized data storage, to edge devices with limited computational and memory resources. These two extreme computing frameworks – the infinitely small on the edge, and the infinitely big on HPC – are the emerging two computing paradigms of the next decades. Designing and studying new AI algorithms,

able to leverage by construction over the intrinsic properties of edge and exascale hardware is an open and crucial challenge. There is a need for a new generation of sparse, robust, adaptive and accurate algorithms for optimization and learning, able to support intelligent systems and autonomous agents, with varying degrees of supervision, where applications go from industrial robotics to banking, mobility and defense, energy management and health.

SPOKE 8: PERVASIVE AI – UNIVERSITÀ DI BOLOGNA [RIF. MICHELA MILANO]

The notion of Pervasive AI does not pertain only smart objects, but refers to the possibility of introducing intelligent models and algorithms into the (socio-technical) contexts in which we live: from work, to cities, from political decisions to public services, from business processes to the production of artifacts in industry and the arts. The notion of pervasiveness implies that we need to model, predict, decide on large-scale and complex socio-technical systems, where every component is itself intelligent, possibly self-interested and we need to take into account artificial infrastructures, natural phenomena, economic aspects, environmental impacts, social dynamics and human behaviour. To make AI positively pervasive, it is necessary to holistically address challenges related to the robustness and accuracy of models, the efficiency of implementations, the cultural acceptability and sustainability of AI: thus, algorithmic and modeling, scientific and technological, educational, legal, socio-economic, ethical and cultural challenges arise in this context.

SPOKE 9: GREEN-AWARE AI – UNIVERSITÀ DELLA CALABRIA [RIF. NICOLA LEONE]

The challenges posed by the sustainable development goals, such as those of the 2030 Agenda, require substantial methodological advances in current AI techniques. On the one hand, the crucial role that Artificial Intelligence can play to deploy a wide range of applications to promote the goals of the European Green Deal has been largely recognized in the last few years. On the other hand, we are becoming more and more conscious of the adverse impacts of AI on the planet, ranging from the direct negative environmental effects of the increased consumption of resources and energy in current computational infrastructures, to systemic effects originating as consequences of intended or unintended changes in the behavior of consumers, users or producers because of the adoption of AI systems. The Spoke will conduct foundational research to define a novel generation of AI systems and algorithms, in which – no matter of the specific application domain being considered – the green dimension will be considered as a first-class citizen in order to mitigate the two kinds of risks defined above.

SPOKE 10: BIO-SOCIO-COGNITIVE AI – IIT [RIF. Massimiliano Pontil]

Human intelligence is capable of seamlessly, continuously, and efficiently learning from limited supervision by adapting to the environment while interacting with other agents. For AI to tackle this ambitious objective, it is necessary to design and develop sense-perceive-act systems that go beyond the mere "imitation by inspiration" paradigm. These systems must integrate learning, planning and discovery by strictly mimicking their biological counterpart at multiple scales from single neurons to more complex brain networks to cognitive and social mechanisms. These are the key features that **sustainable bio-inspired AI systems must have to learn as humans do, producing efficient (data/energy) and robust outcomes.** The efficiency and robustness dimensions are in stark contrast to the recent breed of AI systems whose success comes at the expense of an unsustainable amount of resources while often being brittle.