Artificial Intelligence for Hadron Spectroscopy and Interactions (AI4HSI)

Letter of Intent for HORIZON-INFRA-2025

1) Research objectives

Nuclear Physics (NP) and High Energy Physics (HEP) experiments at leading accelerator facilities generate vast volumes of data aimed at uncovering the properties and interactions of matter's fundamental constituents—leptons and quarks. These data are collected during dedicated runs by large collaborations that design, build, and operate sophisticated detectors. The recorded detector signals are stored and transformed into kinematic variables (e.g., momentum, energy), filtering out noise and background to reconstruct the four-momenta of final-state particles. These are used to compute physics observables such as yields, distributions, multiplicities, cross sections, asymmetries, and correlations. To interpret the results, physics-informed models are developed, combining established theory with unknown components of microscopic interactions. By comparing models to data, researchers validate theoretical predictions and extract key parameters, advancing our understanding of hadron physics. At each stage, from data collection to analysis, extraction, and interpretation, Artificial Intelligence (AI) has the potential to transform workflows, offering deeper insights and more efficient, innovative analysis of these large and complex datasets.

With AI4HSI (Artificial Intelligence for Hadron Spectroscopy and Interactions), we propose to build and coordinate a team of researchers with expertise across all phases of the analysis chain to apply AI techniques to the study of big data in hadron spectroscopy. Providing physical and virtual access to nuclear and particle physics accelerator facilities, AI4HSI will support and enhance ongoing efforts, leveraging existing investments and infrastructure while introducing novel capabilities through AI integration.

In the following, we describe in more detail the data analysis framework and then outline the specific contributions of each research group.

a) Data Collection (Bonn U., Bochum U., Giessen U., Glasgow U., GSI, INFN, Mainz U., York U.)

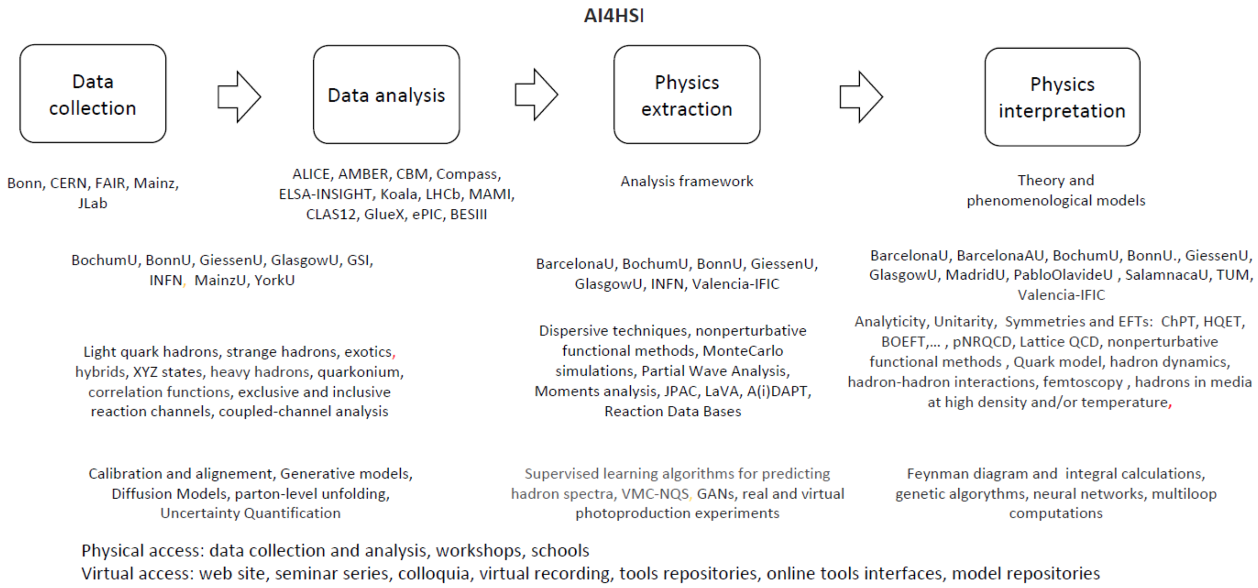
Experimental collaborations contributing to data taking at Bonn (ELSA), CERN (ALICE, AMBER, COMPASS, LHCb), FAIR (CBM, KOALA), Mainz (MAMI), Jefferson Lab (CLAS12, GlueX), BNL (ePIC pseudodata) and other world facilities (e.g., BESIII) will provide access to data associated with publications. Data will be collected using various probes (protons, electrons, ions) across a broad range of kinematic conditions (from GeV to TeV energies) and spanning a wide rapidity range, with diverse final states. AI4HSI will work in close collaboration with these experiments to develop AI-assisted tools for efficient data skimming and transformation into a common data format. The objective is to create the conditions to store selected datasets in a centralized repository, accessible for further analyses.

b) Data analysis (Bonn U., Bochum U., Giessen U., Glasgow U., GSI, INFN, Mainz U., York U.)

Selected reactions will be used to validate AI-supported analysis tools designed to unfold detector effects, such as smearing, acceptance, and efficiency, and to isolate the signal of interest from both physical and instrumental background. The goal is to provide clean samples of vertex-level four-momenta across various reaction channels. Where possible, reactions yielding the same final state in overlapping or complementary kinematic regions will be selected from different data sets. This approach will enable combined analyses with enhanced statistical power and broader coverage of the accessible phase space.

c) Physics extraction (Barcelona U., Bochum U., Bonn U., Giessen U., Glasgow U., INFN, Valencia-IFIC)

Smart tools will be developed to process reconstructed vertex-level data and extract key physics observables such as angular and momentum distributions, cross sections, decay observables, spin-, angular-, and CP-asymmetries, and correlation functions, essential for physics interpretation. In coordination with the experimental collaborations, these AI-supported tools will facilitate the computation and validation of such observables. The resulting data products will be stored in a shared repository and made virtually accessible to all partners through an online web interface, enabling streamlined access and further interpretation.



d) Physics interpretation (Barcelona U., Barcelona A.U., Bochum U., Bonn U., Giessen U., Glasgow U., Madrid U., Pablo de Olavide U., Salamanca U., TUM, Valencia-IFIC)

AI4HSI theoretical groups will develop the theoretical framework and (AI-)tools to interpret physics observables extracted in the previous steps, relying only on fundamental theoretical principles. Thus, advanced and innovative theoretical techniques (EFTs, analyticity constraints, unitarity re-summations, dispersion relations, lattice-QCD simulations, etc.) will be developed and applied for a robust interpretation of the experimental results. We will also address some problems of interest for the low-energy nuclear community, including experimental groups working in facilities like GANIL/SPIRAL2, GSI/FAIR, HIE-ISOLDE, and LNL-SPES. In this context, the use of artificial neural networks to represent quantum many-body wave functions could be used to develop variational Monte Carlo methods with a polynomial scaling in the number of nucleons. Machine learning methods will also be developed for lattice QCD computations and for solving the evolution equations of hard probes in hot and dense media.

2) Deliverables

- Web site with tools for data analysis, physics extraction, and theoretical interpretation.
- Virtual repository for documentation (papers, seminars, and colloquia recordings, …).
- Collaborative venues: two topical and one general workshops.
- Education and dissemination: AI4HSI Summer School.

3) Connection to Transactional Access Infrastructure and Virtual Access projects

AI4HSI relies on data collected from the proposed pre-selected infrastructures. The project will provide and develop online services, facilitate partner access to infrastructures, and make these resources available to the entire hadron physics community upon project completion. ECT* is an ideal venue to interact with researchers participating in workshops organized in the center.

AI4HSI is synergistic with the Nucleon Structure Knowledge and Analysis Toolkit (NuSKAT) project. Although focused on different aspects of hadron physics, both projects will maximize the impact of AI-supported tools in hadron physics by sharing procedures for data analysis and physics extraction.

4) Resources and budget: 750k€

- 4 (exp) + 4 (theory) years postdocs (assuming EU 60k€/per year)	480k€
- Travels	220k€
- Schools, workshops	50k€

5) Participating and partner institutions

Barcelona University, Autonomous University of Barcelona, Bochum University, Bonn University, FAIR, Giessen University, Glasgow University, INFN, Madrid University, Pablo de Olavide University, Salamanca University, Technical University Munich, Valencia-IFIC (Valencia University & CSIC), York University, involving several tens of senior researchers. Partners: JLab, Indiana University, Old Dominion University, University of Virginia