

Ali Ghasempouri

Currently PhD student in CSE (Cycle 39), DIPARTIMENTO DI INFORMATICA - SCIENZA E INGEGNERIA, University of Bologna, Bologna, Italy

✉ seyedali.ghasempour2@unibo.it | github.com/ghasempouri1984

Through my ongoing doctoral research, I'm developing a multi-modal intelligent assistant that leverages AI to enhance human memory of daily experiences, which requires careful consideration of data privacy. My research combines expertise in data modeling, natural language processing, and multimedia data management. I'm driven to create innovative, user-centric AI solutions that bridge the gap between technology and human cognition. My interdisciplinary background equips me to consider the socio-cultural implications of AI.

Education

PhD (in progress)

Bologna, Italy

Alma Mater Studiorum - University of Bologna - DIPARTIMENTO DI INFORMATICA - SCIENZA E INGEGNERIA

Nov. 2023 - Current

- **Field of study:** Computer Science and Engineering
- **Tutor of PhD:** Prof.ssa Ilaria Bartolini
- **Research title:** "RememberMe: Enhancing User Memory of Daily Experiences through Multi-modal Memory Assistant". The goal of this research is to create a smart assistant that helps people remember and organize their daily experiences. A personal assistant that can remind us of what we like, what we have done, and how we felt about it. Whether it's a favorite recipe, a memorable vacation, or a shopping list, this project aims to create a system that understands and helps us recall those memories.

Master

Bologna, Italy

Alma Mater Studiorum - University of Bologna - Department of Classical Philology and Italian Studies

Sep. 2021 - Oct. 2023

- **Field of study:** Digital Humanities and Digital Knowledge
- **Thesis:** Title: "Integrating SPARQL Querying with Chart Creation and Statistical Summarization". It was centered on bridging a gap in the field of data visualization and semantic web by developing a comprehensive tool capable of querying SPARQL endpoints, generating charts, and providing statistical summary by utilizing Python and Flask.

Bachelor of Science

Ghaemshahr, Iran

Azad University - Ghaemshahr Branch

Oct. 2006 - Sep. 2009

- **Field of study:** Mathematics Education
- **Thesis:** The Comparison of Geometry Skills Learning in Dynamic Geometry Educational Environment Based on Computer Algebraic System and Traditional Educational Environment in Secondary School In Babol, Iran

Associate of Science

Tehran, Iran

University of Science and Technology

Sep. 2002- Jul. 2006

- **Field of study:** Applied Mathematics

Professional Experience

FULL STACK DEVELOPER - Master's Thesis Internship

University of Bologna, Italy

The Digital Humanities Advanced Research Centre (/DH.arc)

Feb. 2023 - Sep. 2023

- Developed the backend of a Flask application, focusing on handling and visualizing SPARQL queries.
- Utilized SPARQL queries for knowledge extraction and entity labeling from WikiData within the Flask application.
- Implemented robust error handling and user input validation
- Designed and developed the user interface for the Flask application

VR/AR DEVELOPER

Freelancer

Jan. 2018 - Current

- Developing Marker-based and Markerless AR apps with Vuforia and AR Foundation SDKs
- Developing Android and Head-mounted VR (HTC-VIVE) apps and VR 360 Video
- Creating serious game applications for the virtual spatial navigation tasks in VR
- Developing motion-based mini-game using Unity3D, integrating OpenCV to capture hand movements from WebCam for interactive gameplay.

FILM PRODUCTION AND DIRECTOR ASSISTANT

Tehran

Freelancer

May. 2018 - Current

- Hands-on experience in Screenwriting, visual storytelling, low-budget filmmaking
- In the role of project manager, managed and coordinated short films and live TV program projects, demonstrating strong organizational and leadership skills.
- Utilized digital tools and gears for video capturing and photography, lighting, and documentary video editing.

VR Developer and Tutoring

Tehran, Iran

INSTITUTE FOR COGNITIVE SCIENCE/SHAHID BEHESHTI UNIVERSITY

Jan. 2018 - OCT. 2019

- Collaborated in the development of a serious game with Unity3D and CSharp for the virtual spatial navigation task in VR (HTC-VIVE) under the supervision of Dr.Farzaneh Safavimanesh.
- Delivered lectures during two workshops on "Task Design and Implementation for Spatial Cognition Studies in Unity3D" at Shahid Beheshti University in Tehran, Iran, held on June 28th and July 18th, 2019.

Principle of Task Design in Behavioral Assessment of Cognitive Functions

Tehran, Iran

INSTITUTE FOR COGNITIVE SCIENCE

Jan. 2018

- Essentials of Psychophysics and three methods: Methods of adjustment, limits, and constant stimuli

Curriculum Developer

Tehran/Babol/Sari, Iran

Freelancer

Nov. 2010, Oct. 2017

- Developed and implemented a plan to balance multiple conflicting educational objectives, using TRIZ Inventive Principles to produce potential solutions in learning environments.
- NaturalMath LLC projects with Dr. Maria Droujkova (founder and director) from NC USA: I have participated in curriculum development efforts helped to build, test, and improve courses and books, and translated curricular materials for Persian families. Natural Math is a community for families, math circles, and other learning groups.
- Holding Inventive Mathematical Problem-Solving workshops for fourth grade in the non-government primary school (Andisheh Primary School at Babol).
- Designing and running science project-based learning program based on LEGO Education, Computer Algebra System (GeoGebra) in Institutes for the Intellectual Development of Children and Young Adults
- Developing CAS-based learning content with Geogebra software for grades 6-12 mathematics education for International GeoGebra Institute, Tehran.

E-Learning Content Developer

Tehran, Iran

IRAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Jan. 2005, Jan. 2010

- Developing the educational content and mathematical visualizations for the Calculus course at the e-Learning center of Iran University of Science and Technology using Mathematica and Powerpoint

Fellowship & awards

Bando, 2023, Per La Selezione Di 10 Studenti Iscritti Alla Lm 43 Per La Partecipazione Al Convegno Aiucd 2023 Dal 5 Al 7 Giugno 2023

Grant, 2019, From the Cognitive Sciences and Technologies Council of the Vice Presidency, Iran, for organizing and holding the Workshop series on Task Designs, Spatial Cognition Studies in the Virtual Environments.

Skills

Research and Documentation

- **Research Methods:** Rapid Prototyping, pilot studies, qualitative & quantitative User research, usability testing, user interviews and surveys research, physiological analysis, academic writing
- **Technical Skills:** TEI-compliant XML documents, Jupyter Notebook, LaTeX

Constrained Optimization

- **Multi-Objective Ranking:** Evaluating the integration of multiple constraints into traditional machine learning ranking methods to enhance their applicability in real-world scenarios, such as product search, Identifying computational challenges and proposing potential improvements to existing algorithms.

Data Ethics & Fairness

- **Algorithmic Fairness Analysis:** Identifying and mitigating biases in machine learning models, particularly in recommender systems.
- **Disparity Analysis:** Investigating and addressing unequal representation or outcomes in data-driven systems.
- **Fairness-Aware Machine Learning:** Implementing strategies to ensure equitable treatment and balanced exposure in algorithmic decision-making.

Hardware & software skills

- **Software Developer:** Proficiency in C# for XR Dev, and Proficiency in Python for Data Analysis, Good in C++ for Arduino, Good in Java for Android Dev.
- **ML Techniques:** ML models evaluation, Class imbalance handling, Feature scaling, Hyperparameter tuning, TF-IDF for feature selection
- **ML Algorithms:** Gradient Boosting, Random Forest, Support Vector Machines (SVM)
- **Deep learning frameworks:** Proficiency in TensorFlow, Keras, Pytorch
- **Web Development:** Proficiency in web information technologies (HTML, CSS, Bootstrap, JavaScript (Leaflet, D3.js), Proficiency in creating web application with Streamlit and Flask
- **Natural Language Processing (NLP):** Tokenization, part-of-speech tagging, lemmatization, sentiment analysis, word-sense disambiguation algorithms, Text-Classification using BERT
- **Design & prototyping:** Good in Figma, Adobe Photoshop and Adobe AfterEffects
- **Statistical analysis:** Good in R, Proficiency in Python and Mathematica.
- **AR/VR Development:** Proficiency in Unity(including Multiplayer Dev), Vuforia, MRTK, ARFoundation, Steam VR, Hololens, HTC Vive, Android

Design

- **User Experience (UX):** Design Thinking, User Interface Design, Interaction Design
- **Prototyping:** Sketches and mockup videos, wireframe, storyboard

Soft Skills

- **Project Lead:** Conducting research and leading a multidisciplinary research team.
- **Problem-Solving:** Recognizing and addressing challenges in data modeling and software development
- **Collaboration:** Working in diverse teams across various projects
- **Communication:** Effective in conveying technical information through documentation and presentations.

Multimedia Skills

- **Creative Skills:** Screenwriting, visual storytelling, low-budget film-making
- **Project Management:** Managing and coordinating short films and live TV programs projects, demonstrating organizational and leadership skills
- **Technical Proficiency:** Adobe Premiere Pro, using professional cameras and lighting equipment

Comms

- **English:** B2 (Working Proficiency)
- **Farsi:** C2 (Native Speaker)
- **Italian:** A2 (Elementary Proficiency)

Projects

Practical Image Analysis Software Development Project

Tehran, Iran

Description: As part of an intensive 8-week training course, the project focused on the practical application of modern C++ software development principles and technologies.

July. 2018 - Sep. 2018

- **Purpose:** The project encompassed a wide range of topics including Modern C++, QT5-QML, Software Architecture, and more, culminating in the development of a comprehensive software solution.
- **Roles and Contributions:**
 - **Modern C++ Application:** Utilized advanced features of C++11 and C++14 such as auto declarations, range-based for loops, and lambda expressions to optimize the software performance.
 - **QT5-QML for GUI Development:** Designed and implemented a user-friendly GUI using QT5-QML, demonstrating cross-platform development skills.
 - **Software Architecture:** Applied the principles of 'Clean Architecture' for robust software design, ensuring scalability and maintainability.
 - **Design Patterns and OOD:** Incorporated various design patterns following SOLID principles and Object-Oriented Design methodologies.
 - **BOOST Library and STL:** Leveraged the BOOST library and Standard Template Library (STL) for efficient memory management and algorithm implementation.
 - **Test-Driven Development:** Adopted TDD practices using BOOST.UnitTest and Mock Turtle libraries, ensuring high code quality through unit and integration tests.
 - **Database Integration:** Integrated SQLite database, applying ORM and Repository patterns for efficient data management.
 - **Image Processing with OpenCV:** Implemented image processing algorithms using OpenCV, focusing on features like segmentation and circle detection.
 - **Version Control and Issue Tracking:** Managed the project using Git for version control and GitLab for issue tracking, demonstrating effective team collaboration skills.
- **Outcome:** Developed a functional software application "Automated Image Analysis System". GUI development with QT5-QML, and advanced image processing capabilities with OpenCV.

A VR application that challenges the user to accomplish a navigational task inside a building characterized by a complex curved layout and misaligned maps

Tehran, Iran

Description:

Oct. 2018 - Jan. 2019

- Unity 3D, SteamVR SDK, and SketchUp software were used for designing and making the environment. Twenty-six students evaluated the system using HTC Vive and their navigational behavior was studied. The data analysis was performed using R.

An Interactive Streamlit Tool for Analyzing Modern Texts

Bologna, Italy

Description: the Course of "Text Retrieval, Analysis and Mining" by Professor Fabio Tamburini, part of the Master Degree offer in Digital Humanities and Digital Knowledge(DHDK) at FICLIT Department, University of Bologna

March. 2022 - April. 2022

- **Purpose:** Develop an interactive tool using Streamlit that allows users to analyze and explore modern texts, such as lyrics or plays. The tool provides detailed insights into text attributes, named entities, sentiment, temporal-spatial aspects, and more.
- **Roles and Contributions:**
 - **Text Token Analysis:** Analyzes tokens in a given text, extracting attributes like part-of-speech tags, lemmas, and more.
 - **Named Entity Recognition:** Extracts and visualizes named entities within a text.
 - **TEI-Compliant XML Generation:** Converts a spaCy document into a TEI-compliant XML document.
 - **Sentiment Analysis:** Performs sentiment analysis on texts, returning polarity and sentiment labels.
 - **Temporal-Spatial Analysis:** Maps texts to specific locations and dates.
 - **Part-of-Speech Distribution Plotting:** Visualizes the distribution of part-of-speech tags.
 - **Interactive Map Creation:** Generates an interactive map with circle markers based on location data.
- **Outcome:** The project has combined various text retrieval, analysis, and mining techniques into a single, web-based platform using Streamlit with Python. It provides diverse functionalities that could be valuable for students in related fields.

Exploring and Evaluating Word-Sense Disambiguation Algorithms

Bologna, Italy

Description: the Course of "Natural Language Processing" by Professor Fabio Tamburini, part of the Master Degrees offer in DHDK and Computer Science, University of Bologna

In progress

- **Purpose:** To implement and analyze the efficiency of different word-sense disambiguation algorithms. The implementation will utilize the SemCor corpus from nltk.corpus and will aim to improve the understanding and practical application of word-sense disambiguation in natural language processing.
- **Roles and Contributions:**
 - **Data Preparation and Preprocessing:** Preprocessing the data, including tokenization, lemmatization, and other necessary preparations. Then, splitting the data into training and testing sets for supervised learning methods.
 - **Naive-Bayes Supervised Method Implementation:** Training the model on the training data and evaluating it using the test set. Next, analyzing the performance, including precision, recall, and F1-score.
 - **Lesk-based Lexical Algorithm Implementation:** Applying the algorithm to disambiguate words in the corpus. Then, comparing the results with other methods and analyzing the efficiency.
 - **Non-Supervised Method based on the EM Algorithm Implementation:** Applying the method to the SemCor corpus and analyzing its efficiency, comparing its performance with the other methods.
- **Outcome:** Implementation of three different word-sense disambiguation algorithms, demonstrating a clear understanding of the underlying theories and practical application, and highlighting strengths and weaknesses. Moreover, exploring possible integration with other NLP tasks and applications such as document classification, sentiment analysis, or semantic textual similarity.

TellMeWhere: What's My Next Destination?

Bologna, Italy

Description: the Course of "Data Modelling and Multimedia Databases" by Professor Ilaria Bartolini, part of the Master Degree offer in Digital Humanities and Digital Knowledge, University of Bologna

Feb. 2023 - Jun. 2023

- **Purpose:** The primary aim of the project is to facilitate and improve access to personal multimedia collection for general users. For this reason, we study the 'Imagination' case study, to incorporate Content-based and Semantic-based image retrieval (using query by example and query by keyword approaches). Then, the goal is to design a system to engage users in conversations about their daily experiences, utilizing emotions, facial expressions, and sentiment analysis.
- **Roles and Contributions:**
 - **Data Structuring:** Identified and described different types of multimedia applications, such as structured, semi-structured, and unstructured data.
 - **Data Modeling:** Modeled the data according to relational and XML models, providing an accurate definition of the low-level features for unstructured data.
 - **Textual Documents Analysis:** Created term-document matrices, identified relevant queries, and plotted graphs for Average Precision vs. Nret documents.
 - **Visual Examples and Interfaces:** Enriched the description with visual examples, comparing low-level features like global vs. local features.
- **Outcome:** The 'Imagination' case study enhanced my skills in Data Modelling and Multimodal Data Management through the application of graph-based approaches and K-NN links for analyzing image and region similarities.

RememberMe: Towards Multimodal Conversational Diary

Bologna, Italy

Description: the Course of "Multimedia Data Management" by Professor Ilaria Bartolini, part of the

May. 2023 - Jun. 2023

Master Degree offer in Computer Science and Engineering at DISI Department, University of Bologna

- **Purpose:** The inclusion of the "Multimedia Data Management" course in my study path was a strategic decision. This course addressed the growing complexities of multimedia systems and filled a knowledge gap in managing and browsing multimedia data effectively. Incorporating efficient algorithms and real-world applications in designing user-centric solutions.
- **Roles and Contributions:**
 - Based on 'TellMe Where' project in previous course of 'Data Modelling and Multimedia Databases', this project explored how multimodal conversational AI can be used to correct user misunderstandings, provide context-aware responses, and what are challenges dealing with handling ambiguity.
- **Outcome:** This course enlarged my skills in MM data management by two significant subjects: efficient algorithms for retrieving multimodal (MM) data (sequential retrieval of MM data and index-based retrieval of MM data), as well as browsing MM data collections, exemplified by the Personalizable Image Browsing Engine and Scenique. The state-of-the-art techniques and hands-on experience during the course, significantly contributed to shaping my current PhD proposal.

Knowledge Engineering

Bologna, Italy

Description: the Course of "Knowledge Engineering" by Professor Valentina Presutti, part of the Master

In progress

Degree offer in Artificial Intelligence, University of Bologna

- **Purpose:** Extending Smashub to model lyrics
- **Roles and Contributions:**
 - Designing a music ontology related to the annotation type
 - Deploying the ontology within the Polifonia Ontology Network
 - Integrating a selection of music datasets providing those annotations
 - Creating a Musical Knowledge Graph from your integrated data and ontology
 - Creating example sparql queries that showcase stories from the dataset
- **Outcome:** Different aspects of music through a common ontology is linked. By utilizing the ontology and integrated dataset, a Musical Knowledge Graph is constructed

Predicting Year Label of Email Using Imbalanced Models Based-on Enron Email Dataset

Bologna, Italy

Description: the Course of "Applied Machine Learning" by Professor Daniele Bonacorsi, part of the

May. 2023 - Aug. 2023

Master Degree offer in Bioinformatics at FABIT Department, University of Bologna

- **Purpose:** Trained an imbalanced model to predict the year an email was sent using the Enron Email Dataset. Implemented three approaches: BERT Model, Logistic Regression model, and Feature Scaling and Hyperparameter Tuning for a Logistic Regression model.
- **Roles and Contributions:**
 - **Class Imbalance:** Faced challenges with unbalanced classes, leading to results not much better than always guessing the most populated class. Explored ad-hoc procedures but recognized the need for further investigation into methods to address this issue.
 - **Data Exploration:** Engaged in data preparation and feature selection, including the use of TF-IDF, but identified the need for more in-depth exploration and explanation of these choices.
 - **Model Comparison:** Recognized the value of including a simple base model for comparison and the potential for comparing models on filtered and unfiltered data to deepen understanding.
- **Outcome:**
 - **BERT Model:** Configured and trained a BERT model.
 - **Logistic Regression Model:** Utilized TF-IDF, handled class imbalance with SMOTE, and trained a Logistic Regression model.
 - **Feature Scaling and Hyperparameter Tuning:** Scaled features, conducted hyperparameter tuning, and retrained the Logistic Regression model.
 - **Some Python Libraries Used:** pandas, transformers, Scikit-learn, TF-IDF, SMOTE, MaxAbsScaler, GridSearchCV.

Discovering Italian Paintings at Metropolitan Museum of Art

Bologna, Italy

Description: the Course of "Information Visualization" by Professor Marilena Daquino, part of the Master

Dec. 2022 - Jan. 2023

Degree offer in DHDK at FICLIT Department, University of Bologna

- **Purpose:** A comprehensive analysis of over 500 Italian paintings at the Metropolitan Museum of Art (MET), including works by renowned artists such as Bellini, Raffaello, Tiziano, and Tiepolo. The project aimed to reconstruct the trajectories of the paintings' journeys over time and identify key cultural actors involved in this process.
- **Roles and Contributions:**
 - **Data Preparation and Analysis:** Utilized two datasets (MET Dataset in CSV and Zeri and LODE Dataset in RDF) and employed SQL and SPARQL for querying. Linked datasets through Object Number to explore past trajectories of artworks.
 - **Network Analysis:** Created a directed graph to visualize the collaborative relationships between art galleries and collections, highlighting the centrality of MET among other collections.
 - **Temporal Visualization:** Analyzed and visualized the distribution of artworks' end dates and accession years, identifying significant periods and gaps.
 - **Transfer Types Analysis:** Investigated different modalities of transferred artworks, including the roles of auction houses and art dealers, and visualized the findings using bar charts, maps, and word clouds.
- **Outcome:** The project successfully illuminated the complex journey of Italian paintings at MET, revealing insights into art history, and the role of various actors in the art market. The findings were presented through engaging visualizations and storytelling, catering to a non-specialist audience, and offering potential for further research and content re-use.

Representation and Distribution of SSH Journals in OpenCitations Meta Database

Bologna, Italy

Description: the Course of "Open Science" by Professor Silvio Peroni, part of the Master Degree offer in DHDK at FICLIT Department, University of Bologna

May. 2023 - Aug. 2023

- **Purpose:** Investigate the representation and distribution of SSH journals, emphasizing Open Access status, disciplines, and countries.
- **Roles and Contributions:**
 - **Data Curation:** Prepared Data Management Plan using ARGOS service.
 - **Methodology:** Developed research methodology using protocols.io platform based-on OpenCitations Meta database, ERIH PLUS, and DOAJ index.
 - **Software Development:** Created custom Python software for data processing, integrating the following functionalities: (1) Extracting essential metadata, (2) Handling the countries involved in the publications, (3) Extracting specific disciplines represented within the SSH journals, (4) Comparing the contributions between the EU and US, both in terms of disciplines and countries.
 - **Visualization:** Utilized Streamlit, Matplotlib, Plotly, and Folium in Python.
 - Writing: Contributed to the original draft and review.
- **Outcome:** 78.1% of SSH journals in ERIH-PLUS included in the database; 38% are Open Access; Psychology leads in publications; U.S. and U.K. are top contributors.

Database Processing and Querying Software

Bologna, Italy

Description: the Course of "Data Science" by Professor Silvio Peroni, part of the Master Degree offer in DHDK at FICLIT Department, University of Bologna

March. 2022 - - July. 2022

- **Purpose:** Developing a software solution to process and upload data stored in different formats for simultaneous querying in distinct databases. The project included creating a detailed Jupyter notebook to explain data characteristics, potential issues, and software organization.
- **Roles and Contributions:**
 - Working with publication data stored in CSV files and additional details in JSON files, transforming and preparing them for integration into a relational database.
 - Took charge of data analysis and processing, relational database table creation and population, and the implementation of specific query modules for targeted information retrieval. Leveraged a variety of tools including Python's pandas library and SQL for database management.

Publications

Map-based Missalignment effect in Virtual Spatial Navigation

Shirin Hajahmadi, Farzaneh Safavimanesh, Ali Ghasempouri

8th basic and clinical neuroscience congress, Tehran, Iran, 2019

Exploring Symbolic Narratives in Virtual Spaces: Leveraging Curiosity-Driven Design and the Attention-Value Model for Educational VR Museum Experiences

Shirin Hajahmadi, Seyedali Ghasempouri, Gustavo Marfia

Mimesis Journal 13.2 (2024) pp. 499–511. 2024

Context-aware Agents and Recommender Systems

Seyedali Ghasempouri, Iliaria Bartolini

ACM Computing Surveys, Under Review (2024). 2024

Incremental Learning in Graph-Based Recommender Systems from a Complex Network Perspective

Seyedali Ghasempouri, Iliaria Bartolini

SIGIR2025, Under Review, 2025

References available upon request.