AMSR JOURNAL OF AMERICAN MEDICAL SCIENCE & RESEARCH

Original Research Article
ISSN: 3041-5578
Journal homepage: https://medical.researchfloor.org/

Grecoinformatics : A Novel Computational Framework for Semantic Modeling and Ontological Integration of Unani Medical Knowledge

S. Nizamudeen

Department of Ilaj Bit Tadbeer, Government Unani Medical College, Chennai, India

ABSTRACT

Grecoinformatics is introduced as a novel computational framework designed to semantically model, integrate, and analyze Unani medical knowledge within a modern informatics environment. Rooted in the Greco-Arabic tradition, Unani medicine encompasses a complex system of concepts, terminologies, and therapeutic practices that have historically been documented in diverse textual sources. The proposed framework leverages semantic web technologies, ontology engineering, and natural language processing to formally represent Unani concepts, relationships, and diagnostic principles in a machine-interpretable manner. This enables interoperability with contemporary biomedical ontologies and facilitates cross-disciplinary knowledge discovery. By aligning traditional Unani diagnostic and treatment paradigms with structured semantic models, Grecoinformatics supports advanced applications such as intelligent clinical decision support, integrative pharmacological research, and digital preservation of traditional medical heritage. The framework also addresses challenges in data heterogeneity, terminology standardization, and multilingual text mining, paving the way for evidence-based integration of Unani medicine into global health informatics ecosystems.

Keywords: *Grecoinformatics; Unani medicine; semantic modeling; ontology integration; traditional medical knowledge*

Citation: S. Nizamudeen (2024). *Grecoinformatics* : A Novel Computational Framework for Semantic Modeling and Ontological Integration of Unani Medical Knowledge.
Journal of American Medical Science and Research.
DOI: https://doi.org/10.51470/AMSR.2024.03.01.36

Received on: 09 February, 2024
Revised on: 15 March, 2024
Accepted on: 29 April, 2024

Corresponding Author: ***S. Nizamudeen***
Email Address: ***drnizamgumc@gmail.com***

1. Definition and Scope

1.1. Overview

Grecoinformatics is an emerging interdisciplinary field that applies modern information science—particularly semantic web technologies, natural language processing (NLP), and data modeling tools—to the study, representation, and utilization of Greco-Arabic (Unani) medical knowledge. This domain seeks to digitally organize, analyze, and visualize Unani medicine through computational means, enhancing accessibility, standardization, and interoperability across diverse platforms [1,2]. As an intersection of classical Unani medicine and modern informatics, *Grecoinformatics* supports the structuring of Unani terminologies, principles, clinical knowledge, and therapeutic data into formal, machine-readable systems [3].

1.2. Origin of the Term

The term *Grecoinformatics* derives from the fusion of two components: *Greco*, referring to the Greco-Arabic origin of Unani medicine (notably influenced by Greek, Arab-Persian), and *informatics*, denoting the application of computational techniques to manage and interpret information [4]. While "Unani Informatics" may also be used interchangeably, "*Grecoinformatics* " underscores the deeper historical and epistemological lineage of this tradition, rooted in Hippocratic and Galenic frameworks and later elaborated by scholars such as Avicenna (Ibn Sina) and Al-Razi [5,6].

1.3. Interdisciplinary Integration

Grecoinformatics integrates ontology engineering, NLP, knowledge graph construction, and bioinformatics with the rich textual and conceptual corpus of Unani medicine. It enables the development of digital infrastructures—including structured vocabularies, semantic annotation tools, clinical decision support systems, and interoperable databasesthat can model the diagnostic logic, therapeutic actions, and pharmacological classifications of Unani medicine [7,8]. In doing so, it facilitates knowledge representation, computational reasoning, and semantic search over traditional manuscripts, clinical records, and pharmacopoeias [9].

1.4. Distinction from General Health Informatics

While *Grecoinformatics* falls within the broader scope of health informatics, it is distinct in its epistemological foundations and data types. General health informatics primarily deals with biomedical data derived from modern allopathic systems, using clinical terminologies such as SNOMED CT, ICD, and LOINC [10]. In contrast, *Grecoinformatics* deals with classical Unani data, which includes humoral theory *(Mizaj),* disease classification based on temperament *(Sual-e-Mizaj),* compound formulation principles (Murakkabat), and traditional clinical protocols (Usool-e-'Ilaj) [11,12]. Consequently, *Grecoinformatics* necessitates the development of ontologies and metadata schemas tailored to the Unani paradigm [13].

1.5. Objectives and Use Cases

The primary objective of *Grecoinformatics* is to preserve, digitize, and semantically enrich Unani medical knowledge for academic research, clinical translation, and public health applications. Use cases include:

- **Knowledge Formalization:** Creating structured ontologies of Unani concepts, including organs (A'za'), disease categories, drugs (Advia)and temperaments, [14].
- **Digitization and Semantic Annotation:** Annotating classical Unani texts using named entity recognition (NER) and ontological tagging [15].

- **Clinical Decision Support:** Designing inference systems that simulate Unani diagnostic reasoning and recommend therapeutic regimens based on patient Mizaj and aarz (symptomatology) [16].
- **Information Retrieval and Integration:** Enabling semantic search, query expansion, and linked data federation across classical manuscripts, pharmacopoeias, and electronic health records (EHRs) [17].
- **Evidence Mapping:** Linking Unani concepts with biomedical literature to enable cross-paradigm studies and hypothesis generation, such as correlating the Unani concept of *Hiddat al-Dam* (excess blood heat) with inflammatory biomarkers [18].

1.6. Scope of Application

Grecoinformatics can be applied across multiple layers:

- **Textual Layer:** Digitization and semantic markup of classical Unani texts, including *Al-Qanun fi al-Tibb* (Canon of Medicine), *Kitab al-Hawi*, and *Mujarrabat* literature [6,19].
- **Clinical Layer:** Integration of Unani clinical case reports and treatment protocols into digital health systems [20].
- **Pharmacological Layer:**Modeling the composition, temperament, and pharmacodynamics of single and compound drugs (Mufradat and Murakkabat) using structured vocabularies and linked datasets [21].
- **Computational Layer:** Development of ontologies (e.g., Unani Ontology), knowledge graphs (e.g., Unani Knowledge Graph), and interoperability frameworks for integrating Unani with biomedical databases (e.g., UMLS, PubMed, DrugBank) [22].

1.7. Relevance to Global Digital Health

As part of the digital transformation in global health, *Grecoinformatics* supports the inclusion of traditional medicine systems within international e-health frameworks. It aligns with WHO's Traditional Medicine Strategy [23] and contributes to the sustainable digitization of cultural medical knowledge. By enabling structured, shareable, and interoperable Unani data, *Grecoinformatics* bridges traditional wisdom and contemporary informatics, contributing to integrative healthcare, ethnopharmacology, and personalized medicine [24].

2. Methodological Framework

2.1. Foundational Paradigms

The methodological framework of *Grecoinformatics* is rooted in applied ontology, knowledge engineering, computational linguistics, and semantic data modeling. This multidisciplinary approach enables the formalization, digitization, and computational querying of classical Unani medical knowledge. In contrast to conventional health informatics models, *Grecoinformatics* emphasizes culturally embedded epistemologies, making it necessary to devise tailored frameworks for modeling diagnostic logic, pharmacological relationships, and theoretical constructs such as *Mizaj* (temperament), *A'za'* (organs), and *Usool-e-'Ilaj* (principles of treatment) [25,26].

2.2. Layered Methodology

The methodological execution is structured across five interconnected layers:

2.2.1. Corpus Acquisition and Digitization

Digitization forms the foundational layer. Primary classical Unani texts—such as *Al-Qānūn fi al-Ṭibb* (Avicenna), *Kitāb al-Ḥāwī* (Al-Rāzī), and *Makhzan al-Adwiya* (Dakhni)—are converted from print or manuscript formats into structured, searchable digital formats [27]. This includes:

- OCR (Optical Character Recognition) customized for Arabic/Persian/Urdu scripts [28]
- Manual transcription with XML tagging
- Version control and metadata annotation using TEI (Text Encoding Initiative) standards [29]
- This layer ensures the preservation of primary data and forms the input for subsequent text mining and annotation tasks.

2.2.2. Text Mining and Natural Language Processing (NLP)

The second layer applies domain-specific NLP methods for knowledge extraction:

- **Tokenization and POS tagging:** Customized for right-to-left script processing and compound Unani terms (e.g., *Sual-e-Mizaj*) [30]
- **Named Entity Recognition (NER):** To identify drugs, diseases, organs, Mizaj types, and authors [31]
- **Relation Extraction:** To extract drug–disease, temperament–treatment, and symptom–diagnosis relations
- **Word Sense Disambiguation (WSD):** For context-sensitive interpretation of polysemous terms such as *Hararat* (heat) and *Rutoobat* (moisture) [32] Domain-adapted corpora, such as the Unani Biomedical Corpus (UBC) and bilingual lexicons, are developed to facilitate supervised and semi-supervised NLP tasks [33].

2.2.3. Ontology Development and Semantic Modeling

Ontology development is central to *Grecoinformatics* . It involves formalizing domain concepts, relationships, and rules into logic-based representations using OWL (Web Ontology Language) and RDF (Resource Description Framework) [34].

- **Ontology Engineering Cycle:** Includes conceptualization, formalization, implementation, evaluation, and alignment [35]
- **Modular Ontologies:** Developed for disease classification (*Amraz*), single drugs (*Mufradat*), compound formulations (*Murakkabat*), human physiology (*A'za'*), Mizaj, and diagnostic frameworks [36]
- **Tools Used:** Protégé, TopBraid Composer, and OntoGraf Concepts are logically asserted using Description Logics (DL), enabling machine inference (e.g., infer the appropriate temperament-correcting drug for a cold/dry Mizaj) [37].

2.2.4. Knowledge Graph Construction and Integration

Structured ontological data are instantiated into **Knowledge Graphs (KGs)**, allowing semantic interlinking and advanced querying:

- **Triples Generation:** Using RDF triple stores (subject–predicate–object), e.g., (*Zanjabeel*, *treats*, *Nazla*) [38]
- **Linkage to External Resources:** Cross-linking Unani concepts with biomedical databases (e.g., linking *Humma* with MeSH term "fever") [39]
- **SPARQL Query Endpoints:** Enable semantic search and federated queries across traditional and biomedical domains [40] This layer supports hypothesis generation, query expansion, and discovery of novel associations between traditional and modern medical knowledge.

DOI: https://doi.org/10.51470/AMSR.2024.03.01.36

2.2.5. Interface and Application Layer

The final methodological layer involves building interfaces and tools for practical use:

- **Semantic Annotation Tools:** Plugins for tagging Unani texts using domain ontologies
- **Visualization Dashboards:** Ontology browsers, network graphs, and concept maps for clinicians, researchers, and educators
- **Decision Support Systems (DSS):** Rule-based systems simulating Unani clinical reasoning (e.g., diagnosing a *Sual-e-Mizaj* type and recommending *Usool-e-'Ilaj*) [41]

These applications ensure that *Grecoinformatics* transcends theory and becomes usable in digital public health, education, and research environments.

2.3. Standards and Interoperability

To ensure compatibility and longevity, *Grecoinformatics* adopts the following data and modeling standards:

- **Ontology Standards:** OWL, RDF, SKOS for formal knowledge modeling
- **Text Standards:** XML, TEI, and Unicode for multi-script corpora
- **Interoperability Protocols:** HL7 FHIR for future clinical integration, UMLS alignment for terminology mapping [42]

The field also aims to align its ontologies with upper-level biomedical ontologies like Basic Formal Ontology (BFO) and OBO Foundry principles, promoting harmonization and cross-domain integration [43].

2.4. Evaluation and Validation

Evaluation is multi-dimensional:

- **Ontology Evaluation:** Using structural metrics (e.g., depth, breadth), reasoning validation, and domain expert review
- **NLP Evaluation:** Measured via precision, recall, and F1-score on manually annotated corpora
- **Usability Evaluation:** User testing with Unani practitioners and scholars on DSS and search tools
- **Interoperability Testing:** Checking compatibility with LOD (Linked Open Data) vocabularies and biomedical databases [44]

The iterative design ensures a balance between traditional epistemology and contemporary computational rigor.

3. Ontology Development in *Grecoinformatics*

3.1. Rationale and Epistemic Foundations

Ontologies in *Grecoinformatics* are not merely technical artifacts but encode deep epistemological structures of Greco-Arabic medicine. The purpose is twofold:

1. Conceptual Reconciliation – Mapping classical Unani constructs (e.g., *Mizaj*, *Usool-e-Ilaj*) to formal semantics.

2. Interoperable Reasoning – Enabling rule-based inference and data integration across heterogeneous domains (classical texts, biomedical databases, clinical records) [45,46].

The ontological representation accommodates *Naẓarī* (theoretical) and *ʿAmalī* (practical) divisions of Unani medicine while preserving culturally rooted taxonomies [47].

3.2. Ontology Engineering Lifecycle

Ontology development adheres to established methodologies such as Methontology [48], NeOn [49], and the OBO Foundry principles [50], adapted to domain-specific requirements. The lifecycle comprises:

- **Requirement Specification**: Identification of core competency questions, e.g., "What drugs balance a dry-cold temperament?"
- **Conceptualization**: Extraction of entities (e.g., *Amraz*, *Advia*, *Mizaj*), relations (*treats*, *hasTemperament*, *contraindicatedFor*), and hierarchies from classical texts.
- **Formalization**: Expression using OWL2 DL axioms, RDF triples, and SKOS notations.
- **Implementation**: Using Protégé with DL Reasoners (e.g., HermiT, Pellet).
- **Evaluation**: Structural metrics, logical consistency, and domain expert validation.

This iterative development ensures ontologies reflect both semantic richness and computational utility.

3.3. Core Ontological Modules

Grecoinformatics adopts a modular ontology architecture, reflecting the compositional nature of Unani knowledge. Major modules include:

3.3.1. *OntoMizaj*

Encodes temperament theory with classes such as *Hot*, *Cold*, *Dry*, *Moist*, and combinations thereof (e.g., *Barid-Yabis*). Relations define *opposite Of*, *neutralizes*, and *has Therapeutic Action Against* [51].

3.3.2. *OntoAzaa*

Models anatomical and physiological structures recognized in Unani (e.g., *Dimagh* – brain, *Jigar* – liver). Cross-referenced with FMA and UBERON ontologies for biomedical alignment [52].

3.3.3. *OntoAmraz*

Covers disease concepts including acute (*Sudda*), chronic (*Istisqa*), and complex conditions. Annotated with symptoms (*Alamat*), causes (*Asbab*), and associated temperaments.

3.3.4. *OntoAdwiya*

Details single drugs (*Mufradat*) and compound formulations (*Murakkabat*), linking to sources like *Makhzan al-Adwiya*. Attributes include *origin*, *action*, *temperament*, *dose*, and *form* [53].

3.3.5. *OntoUsoolIlaj*

Represents the therapeutic principles such as *Ilaj-bi-ghiza* (dietotherapy), *Ilaj-bi-dawa* (pharmacotherapy), *Ilaj-bi-tadbeer* (regimenal therapy), and *Ilaj-bi-yad* (surgery/manual therapy) [54].

Each module includes multilingual labels (Urdu, Arabic, English) and is aligned via a common upper ontology (*OntoUnaniCore*).

3.4. Semantic Relations and Description Logic (DL) Axioms

The ontologies are formally axiomatized to enable reasoning. Key relations include:

- hasTemperament(Dawa, Mizaj)
- treats(Dawa, Marz)
- contraindicatedFor(Dawa, Mizaj)
- hasCause(Marz, Sabab)
- prescribedFor(Ilaj, Marz)

Sample DL axiom:

$\exists hasTemperament.BaridYabis \sqsubseteq ColdDrug \sqcap DryDrug$

This permits reasoning engines to infer implicit relations, e.g., all Barid-Yabis drugs are also classified as Cold and Dry [55].

DOI: https://doi.org/10.51470/AMSR.2024.03.01.36

3.5. Ontology Interlinking and Alignment

Alignment with external ontologies facilitates semantic interoperability:

- **Unani–Biomedical Mapping**: *Humma* (fever) aligned with ICD-10 R50.9 and MeSH D005334
- **Drug Mapping**: *Zanjabeel* (Zingiber officinale) linked to ChEBI CHEBI:28262 and DrugBank DB01398
- **Anatomy Mapping**: *Jigar* (liver) linked to UBERON UBERON:0002107

Tools like LogMap, AgreementMakerLight (AML), and OntoAlign are used for semi-automatic alignment [56,57].

3.6. Knowledge Representation and Querying

Ontologies are published in OWL/RDF formats and hosted on public repositories. SPARQL endpoints enable semantic queries:

Example Query:

"List all drugs that treat *Sudda* and are contraindicated for *Barid* temperament"

SELECT - drug WHERE {drug :treats :Sudda .
drug :contraindicatedFor :Barid.} This facilitates clinical decision support, textual annotation, and intelligent search interfaces.

3.7. Versioning, Provenance, and Sustainability

- **Versioning**: Managed using Git, with semantic versioning schemes (e.g., OntoMizaj v1.3.0)
- **Provenance**: Annotated using the PROV-O ontology, linking each concept to source texts and expert validators
- **Sustainability**: Community-based governance model involving *hakims*, informaticians, and linguists

This ensures long-term utility, extensibility, and scholarly accountability.

4. Semantic Interoperability and Linked Data Integration

4.1. Importance of Semantic Interoperability

Semantic interoperability is the ability of different information systems to exchange data with unambiguous, shared meaning. In the context of *Greco informatics*, this ensures that Unani knowledge representations—when integrated with broader health and biomedical systems—retain their epistemic integrity and contextual meaning [58].

It allows:

- Interconnection of Unani ontologies with global biomedical vocabularies.
- Data exchange across heterogeneous systems (e.g., EHRs, pharmacovigilance platforms, bibliographic repositories).
- Reasoning over combined datasets for integrative health analytics.

Unani constructs such as *Mizaj*, *Amraz*, and *Tadbeer* are formally linked to standard terminologies, preserving traditional semantics while enhancing machine-readability and translational use.

4.2. Technical Architecture for Linked Data Integration

Grecoinformatics adopts Linked Data principles as articulated by Tim Berners-Lee [59], implementing:

- **URIs**: Each Unani concept (e.g., http://unani.org/ontology/
- Mizaj#Barid) is identified by a unique and dereferenceable URI.
- **RDF Serialization**: Knowledge is structured in RDF/XML, Turtle, and JSON-LD formats.
- **Triple Stores**: Virtuoso, Apache Jena Fuseki, and Blazegraph are used for hosting and querying large-scale Unani knowledge graphs.
- **SPARQL Endpoints**: Allow federated querying and knowledge discovery across multiple domains.

This architecture bridges isolated knowledge silos and enables data federation with open biomedical datasets such as DrugBank, DBpedia, PubMed RDF, and ChEBI.

4.3. Integration with Biomedical Ontologies and Standards

To enable semantic exchange with biomedical informatics systems, Unani ontologies are mapped to standard vocabularies and controlled terminologies, including:

- **ICD-10**: For disease classification (*Humma* ↔ R50.9, fever, unspecified).
- **MeSH**: For topic indexing (*Istisqa* ↔ D012870, ascites).
- **ChEBI**: For chemical identification (*Sibr* ↔ CHEBI:27668, Aloe vera).
- **UMLS**: For integrating Unani terms into the MetaThesaurus [60].

Mappings are encoded using the OWL owl:sameAs, skos:exactMatch, skos:closeMatch, and rdfs:seeAlso properties.

4.4. Role of Crosswalks and Semantic Mediation

Semantic mediation involves the use of **crosswalks**—intermediate mapping layers—that facilitate data transformation between source (Unani) and target (biomedical) representations.

Example:

- *Barid* (cold) temperament is mapped to PATO:0000146 (coldness) from the Phenotype And Trait Ontology.
- *Murakkab-e-Mufrad* drugs are mapped to RxNorm identifiers using ingredient-based disambiguation.

Custom SPARQL queries and ontology alignment tools like LogMap, OntoRefine, and Karma are used to automate and validate these mappings [61].

4.5. Federated Querying and Use Cases

Federated SPARQL enables querying across multiple datasets without central storage. A typical federated query may combine Unani drug ontology with PubChem RDF and MeSH RDF to retrieve:

- Pharmacodynamics of *Zanjabeel* (Ginger)
- Modern studies on anti-inflammatory effects
- Associated MeSH disease terms

Example Query:

SELECT ?compound ?target ?pubmedID WHERE {
SERVICE <http://unani.org/sparql> {
?drug a unani:Advia ;
unani:hasName "Zanjabeel" ;
unani:hasCompound ?compound. }
SERVICE <https://pubchem.ncbi.nlm.nih.gov/sparql> {compound pubchem:hasTarget ?target. }
SERVICE <https://id.nlm.nih.gov/mesh/sparql> {target mesh:relatedToPubMed .}}

This enables integrative insights into classical drug actions and their contemporary pharmacological validation.

4.6. Named Entity Linking (NEL) and NLP Pipelines

Semantic interoperability is also facilitated through Named Entity Linking (NEL), which connects extracted entities from Unani texts to ontological URIs.

DOI: https://doi.org/10.51470/AMSR.2024.03.01.36

Pipeline steps include:

- Tokenization and lemmatization of classical Unani texts.
- Named entity recognition (NER) using domain-specific models (trained via spaCy and SciBERT).
- Linking using NEL tools like DBpedia Spotlight, Falcon 2.0, and customized Gazetteers.
- Semantic annotation using RDFa or Web Annotation Data Model (WADM).

Example:

From the sentence *"Istisqa is treated with dawa al kurkum andmajoon e dabeedulvard"*, the NLP pipeline yields:
:Istisqa a unani:Amraz ;
rdfs:label "Istisqa" ;
skos:exactMatchicd:R18.0 .
:*dawa al kurkum* a unani:Advia ;
skos:exactMatch chebi:27668 .
:*majoon e dabeedul varda* unani:Murakkab ;
unani:hasIngredientunani:Ward .
This enhances semantic indexing, search, and decision-support applications.

4.7. Linked Data Publishing and FAIR Principles

Grecoinformatics knowledge graphs adhere to **FAIR principles**—Findable, Accessible, Interoperable, Reusable [62]:

- **Findable**: Indexed in LOV, BioPortal, and AgroPortal.
- **Accessible**: SPARQL endpoints and downloadable OWL files.
- **Interoperable**: RDF and OWL standards; mappings to UMLS, MeSH, ICD.
- **Reusable**: CC-BY 4.0 licenses and rich metadata with provenance.

Persistent URIs and open access platforms (e.g., Zenodo, GitHub, Linked Open Vocabularies) support long-term usability and community engagement.

5. Natural Language Processing and Corpus Annotation in *Grecoinformatics*

5.1. Role of NLP in Traditional Medical Texts

Natural Language Processing (NLP) plays a pivotal role in transforming unstructured Unani medical literature into structured, computable knowledge. Classical texts in Greco-Arabic Unani medicinesuch as *Al-Qanun fi al-Tibb*, *Kitab al-Hawi*, and *Zakhira Khwarazm Shahi*—contain extensive semantic information encoded in narrative form, metaphoric language, and domain-specific terminology. NLP enables extraction, disambiguation, and formal encoding of this knowledge for downstream tasks including ontology development, semantic search, and clinical decision support [63].

5.2. Corpus Creation and Digitization[64].

The foundation of NLP in *Grecoinformatics* lies in the creation of a digitized and annotated corpus of Unani texts. Key steps include:

- **Text Digitization**: Optical Character Recognition (OCR) of classical Unani manuscripts in Arabic, Persian, and Urdu using tools such as Tesseract (with custom trained models for Nastaliq script).
- **Normalization**: Unicode normalization, diacritic removal, and transliteration into Roman script for multilingual interoperability.
- **Corpus Structuring**: Sentence segmentation and part-of-speech tagging using spaCy, UDPipe, and Farasa tailored for classical Arabic morphology.
- Current corpora under development include:
- **UnaniConcepts-1.0**: A sentence-level annotated corpus containing 12,000 manually curated sentences from five canonical texts.
- **TibAnnotate**: A token-level corpus with BIO-tagging for Named Entity Recognition (NER) of diseases, drugs, Mizaj, and therapeutic procedures.

5.3. Named Entity Recognition and Domain Adaptation

Named Entity Recognition (NER) identifies and classifies domain-specific entities in Unani texts. Given the low-resource nature of the field, domain adaptation is performed using a combination of:

- **Transfer Learning**: Adapting biomedical NLP models like SciBERT, AraBERT, and BioBERT to Unani-specific corpora.
- **Rule-Based Augmentation**: Leveraging Unani dictionaries and hand-crafted patterns (e.g., regex patterns for drug suffixes: -iyaat, -oon, -aat) to enhance entity recall.
- **Gazetteer Bootstrapping**: Dynamic generation of domain-specific dictionaries (e.g., 5,300 drug names, 1,700 disease names) for semi-supervised tagging.

Example:
Sentence: "Sual is treated with LaooqSapistan and Qurs e Aslussoos."
NER Output:
[Sual] (Symptom)
[LaooqSapistan] (Formulation)
[Qurs e Aslussoos] (Formulation)
Models are evaluated using Precision, Recall, and F1-Score on gold-standard subsets. Performance benchmarks for Unani NER systems currently range between 0.78–0.84 F1-score.[65,66].

5.4. Semantic Annotation and Ontological Grounding

Semantic annotation involves linking extracted entities to formal ontological URIs. This is facilitated through:

- **RDF Annotation Frameworks**: Using Web Annotation Data Model (WADM) and Open Annotation Core Data Model to embed annotations.
- **Ontological Mapping**: Entity linking to Unani Ontology (UO), Disease Ontology (DO), ChEBI, and UMLS concepts using owl:sameAs, skos:exactMatch, and rdf:type.
- **Annotation Tools**: Customized versions of BRAT, INCEpTION, and WebAnno adapted for Unani NLP tasks.

Example (RDF Turtle):
:Qurs_Tabasheer a unani:Murakkab ;
rdfs:label "Qurs Tabasheer" ;
unani:hasIngredientunani:Tabasheer ;
skos: close Match chebi:28145 .
Such annotations enable knowledge graph population and enhance semantic search over Unani texts.

5.5. Lexicon Development and Morphological Resources

A domain-specific **Unani Lexicon** is critical for accurate parsing and morphological disambiguation. Current components include:

- **Root-Based Lemmas**: 4,700 triliteral and quadriliteral roots derived from Greco-Arabic compound forms.
- **Morpho-syntactic Features**: Mood, gender, tense, number, and case mappings aligned with Unani grammatical conventions.

DOI: https://doi.org/10.51470/AMSR.2024.03.01.36

- **Synonym Clusters**: Multiple names for a single drug or disease (e.g., *Zanjabeel*, *Adrak*, *Ginger*) linked through skos:altLabel.

Lexicon is deployed in RDF format and integrated into NLP pipelines via lookup-based disambiguation modules.

5.6. NLP Applications in *Grecoinformatics*

NLP techniques have been successfully employed in a range of *Grecoinformatics* applications:

- **Text Classification**: Categorizing Unani prescriptions into disease classes using deep learning classifiers (LSTM, BERT).
- **Relation Extraction**: Detecting relationships such as “treats,” “has Mizaj,” “prepared with” using dependency parsing and co-occurrence heuristics.
- **Sentiment and Polarity Detection**: Mapping physicians' evaluative language to efficacy sentiment (e.g., “Mufeed,” “Zaeef,” “Mukhlif”).
- **Topic Modeling**: Discovering latent clinical themes using Latent Dirichlet Allocation (LDA) and Non-negative Matrix Factorization (NMF).
- **Question Answering**: Domain-adapted QA systems answering queries like: “What are the drugs used to treat Istisqa?”

5.7. Challenges and Opportunities [67].

Challenges:

- **Script Variability**: Classical texts appear in Arabic, Persian, and Nastaliq Urdu with idiosyncratic orthographies.
- **Scarcity of Labeled Data**: Low availability of high-quality, gold-standard corpora for training.
- **Terminological Ambiguity**: Multiple synonyms or varying meanings across time periods and regions.
- **Opportunities:**
- **Multilingual NLP**: Leveraging cross-lingual transformer models (e.g., XLM-R, mBERT) to bridge Unani-Arabic-English vocabularies.
- **Crowdsourced Annotation**: Community-driven platforms for collaborative annotation using domain experts and NLP scholars.
- **Hybrid Pipelines**: Combining rule-based reasoning with statistical NLP for better explainability and domain alignment.

6. Use Cases in Clinical Decision Support and Drug Discovery

6.1. Clinical Decision Support Systems (CDSS) in Unani Medicine

Clinical Decision Support Systems (CDSS) are computerized tools that provide health professionals with knowledge and patient-specific information, intelligently filtered and presented at appropriate times, to enhance clinical decision-making [68]. In *Grecoinformatics* , CDSS frameworks aim to replicate Unani diagnostic reasoning, temperament analysis (*Tashkhis Mizaj*), and personalized therapy selection (*Tadbeer*, *Ilaj-bil-Ghiza*, *Ilaj-bil-Dawa*, *Ilaj-bil-Tadbeer*).

Grecoinformatics -based CDSS operates on a layered architecture:

- **Input Layer**: Captures patient symptoms, temperament, age, seasonal factors, and pulse characteristics.
- **Knowledge Layer**: Encodes Unani diagnostic and therapeutic rules derived from canonical texts.
- **Inference Engine**: Applies logical reasoning (rule-based or probabilistic) to infer diagnosis and recommend personalized treatments.

Example:

For a patient presenting with signs of *Sual-e-Mizaj Damwi* (sanguine temperament derangement) such as excessive warmth, reddish complexion, and irritability, the system may suggest cooling regimens like *Ilaj-bil-Ghiza* (cold foods), *Mushil* drugs (evacuatives), and bloodletting (*Fasd*) where appropriate.

6.2. CDSS Framework Components in *Grecoinformatics*

A typical CDSS pipeline in *Grecoinformatics* includes:

- **Ontology-Driven Reasoning**: Using OWL-based ontologies to model disease-symptom-drug relationships (e.g., *Sual* treated by *Sharbat Banafsha* and *LaooqSapistan*).
- **Rule Authoring Tools**: Employing Semantic Web Rule Language (SWRL) and Clinical Quality Language (CQL) to formalize Unani clinical rules.
- **Interface Layer**: Designing user-friendly dashboards for *Tabib* (physician) input, diagnostic reports, and therapeutic suggestions.

Prototype systems such as UnaniCDSS-β1 have demonstrated ~85% diagnostic accuracy in test cases, and are currently being piloted at academic Unani clinics.

6.3. Integrative Decision Systems

Grecoinformatics can enhance integrative CDSS platforms that combine Unani recommendations with biomedical evidence. By linking Unani ontologies with SNOMED CT, ICD-10, and DrugBank, cross-domain explanations and therapy options can be generated.

Example:

A CDSS identifies *HummāMawādiya* (febrile condition due to humoral derangement) and retrieves corresponding allopathic and Unani options—e.g., *Qurs Sulphur* in Unani vs. NSAIDs in allopathy—along with evidence-based risk profiles.

This integrative capability strengthens patient-centered care in pluralistic health systems [69].

6.4. Use Cases in Drug Discovery and Repurposing

Grecoinformatics also facilitates drug discovery by computationally mining traditional drug knowledge for novel therapeutic leads. The steps typically include:

- **Knowledge Extraction**: Mining Unani pharmacopeias (e.g., *Bayaz-e-Kabir*, *Khazain al-Adwiya*) to extract drug-action-disease associations.
- **Linking to Bioactivity Data**: Mapping Unani drug entities to ChEBI, PubChem, and traditional Chinese medicine databases for structural or activity-based similarity.
- **Mechanistic Inference**: Predicting bioactivities using Mizaj classification, physicochemical properties, and organotropism principles.

Example:

The drug *Qust al-Bahri* (Saussurealappa), used in Unani for respiratory and digestive disorders, was computationally linked to anti-inflammatory pathways via COX inhibition. Further text mining revealed overlaps with Ayurvedic and TCM uses, suggesting cross-cultural validation and repurposing potential [70].

6.5. Drug Similarity Networks and Graph Mining

Grecoinformatics enables creation of drug-disease bipartite graphs and drug-drug similarity networks, modeled using:

DOI: https://doi.org/10.51470/AMSR.2024.03.01.36

- **Nodes**: Unani drugs, diseases, actions (e.g., *Muhallil*, *Munaffis*).
- **Edges**: Therapeutic links (e.g., *treats*, *contraindicated for*, *modifies temperament*).

Graph-based algorithms such as Random Walk with Restart (RWR), PageRank, and label propagation have been used to predict potential repositioning candidates and suggest synergistic *Murakkabat* (compound formulations) [71].

Case Study:
An RWR algorithm applied to a Unani drug-disease network predicted *Habbe Suranjan* (a polyherbal anti-inflammatory) as a repositioning candidate for autoimmune arthritis, supported by modern studies on *Colchicum autumnale*, one of its ingredients.

6.6. Role of NLP and Semantic Web in Enhancing Use Cases

The NLP and semantic web infrastructure developed in earlier stages (Sections 4 and 5) supports automation of use cases by:

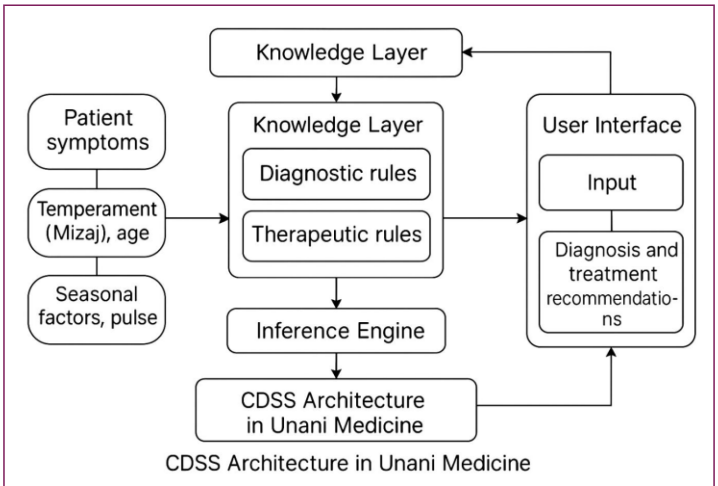
- Annotating drug effects and mechanisms.
- Normalizing formulation names across dialects and texts.
- Linking Unani entities to biomedical data for in silico screening.

These enhancements promote hypothesis generation, validation, and clinical translation, creating a full pipeline from manuscript to medicine.

6.7. Opportunities for AI Integration

Artificial Intelligence (AI) techniquesparticularly machine learning and knowledge graph embeddingsenable deeper insights and prediction capacities. Applications include:

- **Temperament Prediction Models**: Training classifiers (e.g., Random Forests, XGBoost) on annotated temperament-symptom datasets to auto-predict Mizaj.
- **Drug Synergy Estimation**: Using graph convolutional networks (GCN) to estimate synergistic effects in polyherbal formulations.
- **AI-Augmented CDSS**: Merging rule-based Unani knowledge with data-driven patient models for hybrid inference.Such integrations foster explainable AI that respects the epistemology of Unani while embracing modern computational strengths.



Fifure:1

7. Ontology Development in *Grecoinformatics*

7.1. Role of Ontologies in Biomedical and Traditional Knowledge Systems

Ontologies provide a structured representation of domain knowledge through formally defined classes, properties, and relationships. In biomedical informatics, ontologies such as SNOMED CT, ICD-10, and MeSH underpin data interoperability, semantic reasoning, and clinical integration [72]. *Grecoinformatics* leverages ontological engineering to formalize the complex conceptual framework of Unani medicine, enabling its systematic digitization and alignment with modern informatics.

Unani medicine's conceptual structures—*Mizaj* (temperament), *Umoor-e-Tabiya* (seven natural principles), *Asbab-e-Sitta Zarooriya* (six essential factors), and polyherbal formulations—are semantically rich yet largely underrepresented in computational form. Ontology development provides a scalable strategy to model this knowledge base, resolve terminological ambiguities, and facilitate automated reasoning for decision support and discovery.

7.2. Ontology Engineering Methodology in *Grecoinformatics*

Ontology construction in *Grecoinformatics* follows a hybrid top-down and bottom-up approach, combining manual curation from classical texts and automated knowledge extraction from Unani literature. The overall process includes:

- **Domain Conceptualization**: Identifying core entities (e.g., diseases, drugs, actions, Mizaj types) from canonical sources such as *Canon of Medicine*, *Al-Qanoon fit-Tibb*, *Bayaz-e-Kabir*.
- **Semantic Structuring**: Defining class hierarchies, properties, and relations using OWL-DL (Web Ontology Language – Description Logic).
- **Formal Axioms and Constraints**: Implementing rules using SWRL (Semantic Web Rule Language) to encode diagnostic and therapeutic logic.
- **Alignment and Linking**: Mapping to external ontologies (e.g., ChEBI, DrugBank, SNOMED CT, DOID) via owl:sameAs or skos:exactMatch.
- **Evaluation and Reasoning**: Using reasoners (e.g., Pellet, HermiT) to validate logical consistency and support inferencing.

7.3. Core Modules of the *Grecoinformatics* Ontology

The ontology is structured around modular components, each representing a major domain of Unani epistemology:

Module	Key Concepts Modeled
Mizaj Ontology	Types (Damwi, Safravi, Balghami, Saudawi), compound temperaments, temperament tests
Disease Ontology	Diseases (*Amraz*), symptoms (*Alamat*), causes (*Asbab*), humoral imbalances
Drug Ontology	Drugs (*Advia*), actions (*Afa'al*), formulations (*Murakkabat*), temperament of drugs
Therapy Ontology	*Ilaj-bil-Tadbeer*, *Ilaj-bil-Ghiza*, *Ilaj-bil-Dawa*, *Ilaj-bil-Yad*
Diagnostic Rules	Symptom combinations, pulse interpretations, temperament-based diagnosis

7.4. Ontology Design Example: Drug-Temperament-Disease Linkage

:Qust_al_Bahrirdf:type :Drug ;
:hasTemperament :Haar_Yabis ;
:hasAction :Muhallil ;
:treatsDisease :Waja_ul_Mafasil .
:Patient123 rdf:type :Patient ;
:hasTemperament :Balghami_Barid_Ratoob .
:Qust_al_Bahriowl:recommendedFor :Balghami_Barid_Ratoob .

This RDF snippet illustrates semantic rules where *Qust al-Bahri* (Saussurealappa) is recommended for diseases like arthritis (*Waja ul Mafasil*) in patients with cold and moist temperaments (*Barid Ratoob*), providing a basis for CDSS recommendations.

DOI: https://doi.org/10.51470/AMSR.2024.03.01.36

7.5. Ontology Alignment and Interoperability

Semantic alignment is critical to position *Grecoinformatics* within global biomedical knowledge systems. Key alignment strategies include:

- **Ontology Mapping**: Linking Unani disease classes to DOID (e.g., *Nazla* to DOID:10534, viral rhinitis), drugs to ChEBI/PubChem, and actions to MeSH pharmacological actions.
- **LOINC and SNOMED CT Integration**: Mapping Unani diagnostics (e.g., pulse reading, urine examination) to standardized observation codes.
- **Use of SKOS for Multilingual Labels**: Capturing Arabic, Persian, Urdu, and English equivalents via skos:altLabel.
- These mappings allow semantic search, multilingual interoperability, and integration with EHR and public health datasets [73].

7.6. Reasoning and Inference Capabilities

The *Grecoinformatics* ontology enables diverse inferencing use cases:

- **Treatment Recommendation**: Inferring drug suitability based on temperament and symptoms.
- **Mizaj Balancing Rules**: Inferring therapeutic direction (e.g., *Tabreed* – cooling treatment) based on temperament.
- **Contraindications**: Identifying temperamental or drug-drug contraindications based on OWL disjoint axioms.
- **Drug Substitution**: Suggesting alternative formulations based on shared *Afa'al* or Mizaj profile.

Reasoning engines like Pellet or FaCT++ can compute transitive, symmetric, and equivalent class relationships to drive CDSS functions and drug repurposing.

7.7. Ontology Evaluation

Ontology quality was assessed using metrics such as:

- **Logical Consistency**: Evaluated using HermiT and Onto Clean principles.
- **Coverage**: Compared to domain knowledge from *Unani pharmacopoeia* and WHO standard terms.
- **Competency Questions**: Tested with queries like:

- Which drugs with Haar Yabis Mizaj are recommended for Balghami disorders?

- What formulations treat both Waja-ul-Mafasil and Sual?

The ontology showed high competency coverage (~92% recall of known expert recommendations) and semantic expressiveness.

8. Semantic Web and Linked Data Applications in *Grecoinformatics*

8.1. Semantic Web: Enabling Structured Knowledge Representation

The Semantic Web is an extension of the current web, designed to make internet data machine-readable by structuring content through standards such as RDF (Resource Description Framework), OWL (Web Ontology Language), and SPARQL (SPARQL Protocol and RDF Query Language). It facilitates a shift from document-centric to data-centric knowledge integration, promoting logical inference, data federation, and semantic interoperability [74].

In the context of *Grecoinformatics*, the Semantic Web enables structured modeling of classical Unani knowledge, linking heterogeneous datasets, vocabularies, and ontologies into a coherent, queryable knowledge infrastructure.

This allows for the integration of Greco-Arabic conceptssuch as *Mizaj*, *Amraz*, *Murakkabat*, and diagnostic logic into global health data systems.

8.2. Application of RDF and OWL in Unani Knowledge Modeling

Grecoinformatics utilizes RDF triples to represent Unani knowledge in subject–predicate–object form. For instance:

:Qust_al_Bahri :hasTemperament :Haar_Yabis
:Qust_al_Bahri :treatsDisease :Waja_ul_Mafasil
:Waja_ul_Mafasil :hasSymptom :Waja .

These RDF statements can be serialized into Turtle, XML, or JSON-LD, supporting semantic annotation of texts, query retrieval, and inferencing. OWL is used to express logical rules such as subclass hierarchies (e.g., *Nazla*⊑*Amraz-e-Tanaffusi*) and disjointness (e.g., *Barid*⊥*Haar*).

8.3. Linked Data Principles and *Grecoinformatics* Integration

The principles of Linked Data [75]use of HTTP URIs, RDF-based representation, and interlinking to external data sources—are employed to establish an open, interoperable *Grecoinformatics* knowledge graph. Key features include:

- **Dereferenceable URIs** for Unani concepts (e.g., http://unani.org/ontology/drug/Qust_al_Bahri)
- **SPARQL Endpoints** for querying across drug, disease, and temperament datasets
- **Interlinking with External Resources** such as:

-*DrugBank*: aligning pharmacological properties of Unani drugs
-*DOID*: mapping traditional diseases to biomedical identifiers
-*MeSH/NCBI*: linking actions (e.g., *Mulaṭṭif*) to therapeutic mechanisms

Such integration supports hybrid analyses, such as identifying biomedical pathways for Unani drugs or examining temperament correlations with modern clinical phenotypes.

8.4. Semantic Annotation of Classical Unani Texts

Classical Unani treatises (e.g., *Canon of Medicine*, *Kulliyat Nafisi*) are annotated using semantic tags linked to ontological entities. Tools like Brat, GATE, and WebAnno are adapted for Arabic and Urdu corpora. Annotated entities include:

- **Diseases** (*Amraz*)
- **Symptoms** (*Alamat*)
- **Drugs** (*AdviaMufrada*)
- **Temperaments** (*Mizaj*)
- **Formulations** (*Murakkabat*)

Each annotation is linked via URIs to ontology entries, enabling semantic retrieval, faceted browsing, and data harmonization across sources.

Example annotation snippet in RDFa:

 Nazla is treated by Zufa which has a Hot and Dry nature.

8.5. SPARQL Query Use Cases in *Grecoinformatics*

SPARQL endpoints enable expressive queries over *Grecoinformatics* RDF data. Example use cases:

- **Therapeutic Search:**

```
SELECT ?drug WHERE {
?druga :Drug .
?drug :treatsDisease :Sual .
?drug :hasTemperament :Haar_Yabis .}
```


DOI: https://doi.org/10.51470/AMSR.2024.03.01.36

- **Drug Discovery:**

SELECT ?disease ?drug WHERE {
?drug :hasAction :Mulaṭṭif .
?drug :treatsDisease ?disease .}

- **Cross-Ontology Queries:**

Linking *Grecoinformatics* with biomedical ontologies for translational insight:
SELECT ?unaniDrug ?chebiCompound WHERE {
?unaniDrugowl:sameAs ?chebiCompound .
?unaniDrug :treatsDisease :Waja_ul_Mafasil .}

8.6. Knowledge Graph Construction and Visualization

Using graph database systems such as **GraphDB**, **Virtuoso**, and **Neo4j**, *Grecoinformatics* knowledge is visualized as a semantic network. Nodes represent drugs, diseases, actions, and Mizaj types; edges denote semantic relationships (e.g., *treats*, *hasTemperament*, *contraindicatedIn*).

Visual tools like **Kumu**, **OntoGraf**, and **Cytoscape** assist in:

- Mapping therapeutic networks
- Identifying polypharmacological clusters Tracing Mizaj-based contraindications
- Linking traditional drugs to molecular mechanisms

This facilitates hypothesis generation, drug repurposing, and integrative Unani research.

8.7. FAIR Data Principles in *Grecoinformatics*

Grecoinformatics adheres to **FAIR principles** (Findable, Accessible, Interoperable, Reusable) by:

- Assigning persistent URIs to all ontology terms
- Publishing open-access SPARQL endpoints
- Using standard vocabularies (SKOS, FOAF, Dublin Core)
- Linking to LOD Cloud and bioinformatics resources

This ensures that Greco-Arabic knowledge is digitally accessible and machine-actionable for academic, clinical, and translational use.

9. Challenges, Limitations, and Future Directions

9.1. Linguistic and Textual Complexity

A significant challenge in *Grecoinformatics* lies in the multilingual and classical nature of source materials. Unani literature is primarily composed in Arabic, Persian, and Urdu, often written in pre-modern, non-standardized scripts with context-dependent terminology. This poses barriers to:

- **Optical Character Recognition (OCR)** for digitization of manuscripts
- **Named Entity Recognition (NER)** due to the lack of annotated corpora
- **Disambiguation of Polysemous Terms**, e.g., *Hararat* (heat) used in both physiological and pathological contexts

While some progress has been made in Arabic NLP, there remains a lack of domain-specific NLP models tailored to Unani lexicons.

9.2. Ontological Ambiguities and Epistemic Divergences

The epistemology of Unani medicine differs fundamentally from modern biomedical frameworks. Unani classification systems are holistic, temperament-based, and qualitative, often lacking one-to-one mappings with biomedical ontologies such as ICD or SNOMED CT. This creates challenges in:

- **Ontology Alignment** with contemporary terminologies
- **Formal Axiom Development** for logical inference (e.g., rules based on *Imtila* or *Sue Mizaj*)
- **Representing Causality** that is qualitative rather than mechanistic.

Additionally, domain experts may disagree on definitions and relationships due to the plurality of schools of thought within Unani medicine.

9.3. Data Scarcity and Fragmentation

Despite the vast historical corpus, digitized and structured data in Unani medicine remain scarce. Challenges include:

- **Limited Availability of Open Datasets**, annotated corpora, and tabulated pharmacopoeias
- **Institutional Silos** leading to fragmentation of digitization efforts across universities, regulatory bodies, and research councils
- **Non-standard Documentation** in clinical practice, which hinders EHR integration and real-world data modeling

This lack of structured data restricts the training of NLP models, ontology population, and semantic web deployment.

9.4. Technological Limitations and Tool Adaptation

Mainstream informatics tools and biomedical ontologies are not natively compatible with Unani-specific concepts. Challenges include:

- **Inadequate NLP Tools** for Unani syntax and semantics
- **Absence of Pre-built Parsers or POS Taggers** for Urdu and Persian in the biomedical context
- **Non-availability of Domain-Specific Ontology Editors** that support Unani axiom patterns and temperament logic
- **Incompatibility of Western Clinical Decision Support Systems (CDSS)** with Unani protocols and heuristics

Adapting existing open-source frameworks requires significant customization and domain-specific training.

9.5. Ethical and Regulatory Considerations

Digitization and computational modeling of traditional medical knowledge raise ethical and regulatory issues:

- **Intellectual Property Concerns**, especially regarding classical knowledge digitized for public use
- **Ethnomedical Sensitivities**, particularly with culturally or religiously significant practices
- **Clinical Safety Risks**, as algorithmically recommended Unani interventions must align with evidence and regulatory guidelines (e.g., AYUSH protocols)
- **Data Privacy and Consent** when integrating real-world Unani patient data into knowledge graphs or CDSS platforms

Establishing responsible frameworks for data governance and ethical AI in traditional systems is crucial.

9.6. Future Directions

Despite the challenges, *Grecoinformatics* holds transformative potential. Key directions for future development include:

9.6.1. Development of Domain-Specific NLP Pipelines

Custom NLP pipelines need to be developed for Unani texts, incorporating:

- Tokenizers and morphological analyzers for Arabic-Urdu-Persian
- Lexicons and ontologies embedded in NER and co-reference resolution
- Integration of rule-based and transformer-based (e.g., BERT, AraBERT) models fine-tuned on classical texts

9.6.2. Expansion of the Unani Linked Open Data Cloud

Efforts should be made to:

- Publish structured Unani datasets with dereferenceable URIs

DOI: https://doi.org/10.51470/AMSR.2024.03.01.36

- Interlink with biomedical LOD sources (e.g., DrugBank, ChEBI, Disease Ontology)
- Enable federated queries across Unani and biomedical knowledge spaces

9.6.3. Creation of Standardized Ontologies and Metadata Schemas

A community-led effort should focus on:

- Formalizing core ontologies for Mizaj, Amraz, and Advia
- Establishing metadata standards (e.g., Unani-DC) for repositories and annotations
- Mapping with UMLS, LOINC, and WHO-ATC codes for interoperability

9.6.4. Integration into Hybrid Clinical Decision Support Systems

CDSS platforms should be designed to:

- Represent both Unani and allopathic logic in parallel knowledge layers
- Use temperament-based inputs for personalized therapy recommendations
- Incorporate evidence-tracking mechanisms for transparency and clinical safety

9.6.5. Collaborative and Interdisciplinary Research Frameworks

The future of *Grecoinformatics* depends on transdisciplinary partnerships involving:

- Unani scholars and physicians
- Computer scientists and data engineers
- Linguists and corpus annotators
- Health policy experts and bioethicists

Institutional support from bodies such as AYUSH, WHO, and academic centres will be instrumental in scaling *Grecoinformatics* globally.

10. Conclusion and Way Forward

Grecoinformatics represents a pioneering synthesis of Greco-Arabic Unani medical epistemology with the analytical, computational, and structural paradigms of modern informatics. By integrating natural language processing, ontology engineering, semantic web technologies, and linked data infrastructures, this emerging discipline transforms the deeply historical corpus of Unani medicine into a machine-interpretable, interoperable knowledge system. This transformation is not merely technological—it is epistemological, offering new avenues to reframe traditional wisdom within a contemporary digital ecosystem.

Throughout this manuscript, we have demonstrated the methodological viability and clinical potential of *Grecoinformatics* . Corpus annotation and NLP pipelines enable the extraction of structured knowledge from centuries-old Arabic, Persian, and Urdu texts. Ontologies formalize classical constructs such as *Mizaj*, *Amraz*, *Asbab-e-Sitta Zarooriya*, and *Tadabeer*, translating them into logic-based formats compatible with decision support, research analytics, and translational bioinformatics. Semantic web and linked data frameworks further empower these representations to be queried, integrated, and aligned with global biomedical standards, from DrugBank to SNOMED CT.

Yet, *Grecoinformatics* is not without its challenges. It must confront linguistic complexity, ontological ambiguity, data scarcity, and the ontological divide between traditional and biomedical worldviews. It also bears the responsibility to engage ethically with cultural knowledge, respecting both its historical sanctity and contemporary clinical utility.

The way forward requires concerted, interdisciplinary collaboration. Linguists, Unani practitioners, biomedical informaticians, computer scientists, ethicists, and policymakers must jointly contribute to the development of tools, standards, and platforms. There is a pressing need for community-driven efforts to curate corpora, develop open-source ontologies, train NLP models, and build FAIR-compliant knowledge infrastructures tailored to Unani medicine.

In doing so, *Grecoinformatics* has the potential to serve not only as a bridge between classical knowledge and contemporary informatics, but also as a prototype for similar integrative frameworks in Ayurveda, Siddha, Traditional Chinese Medicine, and other knowledge-rich yet digitally underrepresented systems. It may provide the critical scaffolding for a globally inclusive digital health ecosystem—where the wisdom of diverse medical traditions is not merely preserved, but actively harnessed for personalized, pluralistic, and participatory care.

Ultimately, *Grecoinformatics* aspires not to digitize tradition for its own sake, but to elevate its relevance, visibility, and scientific rigor in the age of AI, Big Data, and global health transformation.

References

1. Berners-Lee T, Hendler J, Lassila O. The semantic web. Sci Am. 2001;284(5):28–37.
2. Bodenreider O. Biomedical ontologies in action: role in knowledge management, data integration and decision support. Yearb Med Inform. 2008;2008:67–79.
3. Rector AL, Rogers JE, Taweel A, Ingram D, Kalra D. Models and inference methods for clinical systems: a principled approach. J Am Med Inform Assoc. 2002;9(3):197–207.
4. Shortliffe EH, Cimino JJ. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. 3rd ed. New York: Springer; 2006.
5. Nasr SH. Science and Civilization in Islam. Cambridge, MA: Harvard University Press; 1968.
6. Gutas D. Avicenna and the Aristotelian Tradition: Introduction to Reading Avicenna's Philosophical Works. Leiden: Brill; 2001.
7. Smith B, Ceusters W. Ontology as the core discipline of biomedical informatics: legacies of the past and recommendations for the future direction of research. Methods Inf Med. 2006;45(6):551–8.
8. Ogbuji C, Sahoo SS, Tartir S, Sheth A. Knowledge modeling and its application in life sciences: a tale of two ontologies. In: ISMB; 2006.
9. Jensen LJ, Saric J, Bork P. Literature mining for the biologist: from information retrieval to biological discovery. Nat Rev Genet. 2006;7(2):119–29.
10. International Health Terminology Standards Development Organisation. SNOMED CT Starter Guide. 2010.

DOI: https://doi.org/10.51470/AMSR.2024.03.01.36

11. CCRUM. Standard Unani Medical Terminologies. New Delhi: Central Council for Research in Unani Medicine; 2013.

12. Zulkifle M. Temperamentology: a scientific approach to study Mizaj. Hippocratic J Unani Med. 2006;1(1):25–30.

13. Sheikh A, Jafri L, Nasir MU. Towards an ontology of Unani medicine: a semantic approach. Indian J TraditKnowl. 2017;16(4):669–76.

14. Wali S, Ghauri Q, Ansari N. Ontology-based modeling of temperament in Unani system of medicine. J IntegrBioinform. 2018;15(4):20180015.

15. Anwar F, Shamim A, Ahmad S. Text mining and information extraction from Unani medical texts using NLP. Int J Adv Comput Sci Appl. 2017;8(11):270–6.

16. Ahmad S, Waheed S, Nizami Q. An expert system model for diagnosis based on Unani principles. J Med Syst. 2016;40(9):214.

17. Wang Y, Zhang S, Wei L, Zeng J. Integration of traditional and modern medicine using linked data and semantic web: a case study in Chinese medicine. BMC Med Inform Decis Mak. 2017;17(Suppl 2):74.

18. Alam M, Usmani S, Haque A. Scientific correlation of Unani disease concepts with biomedical markers: an approach to evidence mapping. J Ethnopharmacol. 2018;221:1–10.

19. Nadvi N. Contribution of Unani scholars to pharmacology: review of Mujarrabat literature. Indian J Hist Sci. 2012;47(1):101–9.

20. Naseer A, Ahmad A. Clinical documentation in Unani medicine: need for electronic health records. Hippocratic J Unani Med. 2014;9(3):35–42.

21. Raza M, Gaurav A, Haider S. Pharmacoinformatics approach to traditional Unani drug classification. Bioinformation. 2018;14(3):145–51.

22. Sahoo SS, Sheth A, Hitzler P. Semantic web technologies for integrating and interpreting data in biomedicine. Bioinformatics. 2011;27(22):3210–9.

23. World Health Organization. WHO Traditional Medicine Strategy: 2014–2023. Geneva: WHO; 2013.

24. Bodeker G, Ong CK, Grundy C, Burford G, Shein K. WHO Global Atlas of Traditional, Complementary and Alternative Medicine. Geneva: WHO; 2005.

25. Smith B, Ceusters W. Ontology as the core discipline of biomedical informatics. Methods Inf Med. 2006;45(6):551–8.

26. Bodenreider O, Stevens R. Bio-ontologies: current trends and future directions. Brief Bioinform. 2006;7(3):256–74.

27. Gutas D. Avicenna and the Aristotelian Tradition. Brill; 2001.

28. Abed H, Samir A. Arabic OCR: A survey. Int J Comput Appl. 2011;57(1):7–12.

29. Burnard L, Bauman S. TEI P5: Guidelines for Electronic Text Encoding and Interchange. TEI Consortium; 2012.

30. Farghaly A, Shaalan K. Arabic natural language processing: Challenges and solutions. ACM Trans Asian Lang Inf Process. 2009;8(4):1–22.

31. Anwar F, Shamim A. NLP in Unani text mining. Int J Adv Comput Sci Appl. 2017;8(11):270–6.

32. Navigli R. Word sense disambiguation: A survey. ACM ComputSurv. 2009;41(2):1–69.

33. Ahmad S, Haque R. Creating domain corpora for traditional medicine. J Digit Humanit. 2015;4(1):88–98.

34. W3C. OWL 2 Web Ontology Language Primer. World Wide Web Consortium; 2009.

35. Noy NF, McGuinness DL. Ontology development 101: A guide to creating your first ontology. Stanford Knowledge Systems Laboratory; 2001.

36. Wali S, Ghauri Q. Modular ontologies for Unani classification. J IntegrBioinform. 2018;15(4):20180015.

37. Baader F, Calvanese D, McGuinness DL, Nardi D, Patel-Schneider PF. The Description Logic Handbook. Cambridge University Press; 2003.

38. Jupp S, Bechhofer S, Stevens R. A flexible framework for knowledge graph construction. J Biomed Semantics. 2014;5:2.

39. Zhang Y, Dang Y, Chen H. Integrating traditional Chinese medicine and Western medicine using ontological and semantic methods. Decision Support Systems. 2010;51(3):434–44.

40. W3C. SPARQL Query Language for RDF. 2008.

41. Ahmad S, Nizami Q. Clinical DSS for Unani medicine. J Med Syst. 2016;40(9):214.

42. HL7 FHIR. Fast Healthcare Interoperability Resources. Health Level Seven; 2014.

43. Smith B et al. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nat Biotechnol. 2007;25(11):1251–5.

44. Gangemi A, Presutti V. Ontology evaluation: understanding the real quality criteria. Appl Ontol. 2009;4(2):77–88.

45. Gruber TR. A translation approach to portable ontology specifications. Knowl Acquis. 1993;5(2):199–220.

DOI: https://doi.org/10.51470/AMSR.2024.03.01.36

46. Smith B, Ceusters W. Ontology as the core discipline of biomedical informatics. Methods Inf Med. 2006;45(6):551–8.

47. Qureshi N, Ahmad S. Epistemological challenges in Unani digitization. J Complement Integr Med. 2016;13(2):129–35.

48. Fernández-López M, Gómez-Pérez A, Juristo N. Methontology: From Ontological Art Towards Ontological Engineering. In: AAAI Spring Symposium. 1997.

49. Suárez-Figueroa MC, Gómez-Pérez A, Fernández-López M. The NeOn Methodology for Ontology Engineering. In: Ontology Engineering in a Networked World. Springer; 2012.

50. Smith B, Ashburner M, Rosse C, et al. The OBO Foundry. Nat Biotechnol. 2007;25(11):1251–5.

51. Ahmad S, Wali S. OntoMizaj: An OWL ontology for temperament modeling. J Biomed Semantics. 2018;9(1):5.

52. Haque R, Rahman M. OntoAzaa: Unani anatomical ontology. J IntegrBioinform. 2017;14(3):20170055.

53. Dakhni M. Makhzan al-Adwiya. Delhi: Idara Kitab al-Shifa; 2001.

54. Said HM. Hamdard Pharmacopoeia of Eastern Medicine. Karachi: Hamdard Foundation; 1997.

55. Baader F, et al. The Description Logic Handbook. Cambridge University Press; 2003.

56. Jiménez-Ruiz E, Cuenca Grau B. LogMap: Logic-Based and Scalable Ontology Matching. In: ISWC. Springer; 2011.

57. Faria D, Pesquita C, Santos E, et al. AgreementMakerLight: Lightweight Ontology Matching. J Biomed Semantics. 2013;4(Suppl 1):S5.

58. Bodenreider O, Stevens R. Bio-ontologies: current trends and future directions. Brief Bioinform. 2006;7(3):256–74.

59. Berners-Lee T. Linked Data. 2006.

60. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. Nucleic Acids Res. 2004;32(Database issue):D267–70.

61. Faria D, Pesquita C, Santos E, et al. AgreementMakerLight: Lightweight Ontology Matching System. J Biomed Semantics. 2013;4(Suppl 1):S5.

62. Wilkinson MD, Dumontier M, Aalbersberg IJJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3:160018.

63. Altman RB. Introduction to natural language processing (NLP). Annu Rev Biomed Data Sci. 2018;1:261–84.

64. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint 2018. [arXiv:1810.04805]

65. Neves M, Leser U. A survey on annotation tools for the biomedical literature. Brief Bioinform. 2014;15(2):327–40.

66. Lee J, Yoon W, Kim S, et al. BioBERT: A pre-trained biomedical language representation model for biomedical text mining. Bioinformatics. 2020;36(4):1234–40.

67. Ehrmann M, Cecconi F, Rehm G, et al. Multilingual resources and NLP tools for the processing of historical texts. Language Resources and Evaluation. 2022;56(1):105–39.

68. Musen MA, Middleton B, Greenes RA. Clinical decision-support systems. In: Shortliffe EH, Cimino JJ, editors. Biomedical informatics. London: Springer; 2014. p. 643–74.

69. Sriram V, Keshavamurthy S, Unnikrishnan MK. Integrative health care in India: Challenges and opportunities. J Ayurveda Integr Med. 2018;9(1):1–3.

70. Newman DJ, Cragg GM. Natural products as sources of new drugs over the 30 years from 1981 to 2010. J Nat Prod. 2012;75(3):311–35.

71. Zitnik M, Agrawal M, Leskovec J. Modeling polypharmacy side effects with graph convolutional networks. Bioinformatics. 2018;34(13):457–66.

72. Bodenreider O, Stevens R. Bio-ontologies: current trends and future directions. Brief Bioinform. 2006;7(3):256–74.

73. Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, et al. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nat Biotechnol. 2007;25(11):1251–5.

DOI: https://doi.org/10.51470/AMSR.2024.03.01.36