

DOI: 10.22126/TBIH.2024.9567.1000



## Extended Abstract

### Introduction:

Persian, as a language with deep historical roots and contemporary relevance, holds the distinction of being the ninth most utilized language in digital spaces, surpassing Arabic, Turkish, and other regional languages in online content volume. This prominence underscores the urgent need for advanced linguistic tools to preserve and modernize classical Persian literature, which faces challenges in adapting to the digital era. Classical Persian texts—ranging from the poetry of Rumi and Hafez to historical prose—are integral to cultural identity but remain underrepresented in digital corpora. This study addresses this gap by investigating the role of citizen science, a collaborative approach that engages the public in academic research, to build and enrich linguistic corpora for classical Persian literature. The research aims to democratize access to cultural heritage while enhancing computational tools for language processing, thereby bridging traditional scholarship with modern technological demands.

### Method:

The study employed a seven-step meta-synthesis framework developed by Sandelowski and Barroso, designed to synthesize qualitative data from diverse sources. The process began with a systematic literature review of 85 peer-reviewed articles focused on citizen science, linguistic corpora, and Persian language processing. Key themes were identified through iterative coding, including machine learning integration, public participation models, and data quality assurance. Following this analysis, the research team designed an indigenous citizen science platform tailored for Persian literary corpus development. The platform's architecture incorporated: Crowdsourcing modules for text digitization and annotation, allowing users to transcribe manuscripts, tag linguistic features, and validate machine-generated outputs.

Machine learning algorithms to automate preliminary text analysis, including optical character recognition (OCR) for handwritten manuscripts and semantic tagging for archaic vocabulary. Gamification elements to sustain user engagement, such as progress badges and leaderboards for contributors.

Quality control mechanisms, such as peer-review workflows and AI-driven anomaly detection, to ensure corpus accuracy.

User testing involved 450 participants, including scholars, students, and heritage enthusiasts, who interacted with the platform over six months. Data collection focused on metrics like contribution volume, error rates, and user feedback.

### Results and Discussion:

The platform's implementation yielded transformative outcomes. Citizen science emerged as a cornerstone, with public contributors generating 62% of the initial corpus data, including rare manuscripts previously inaccessible to academic institutions. Machine learning integration reduced manual annotation time by 40%, though challenges persisted in recognizing cursive script variations. Key findings include:

Enhanced Data Diversity: Contributions from non-specialists introduced dialectal and stylistic variations often overlooked in traditional corpora, enriching the dataset's representativeness.

Cost Efficiency: Decentralized data collection lowered operational expenses by 35% compared to conventional methods.

Accuracy Improvements: Hybrid validation (human-AI) achieved a 92% accuracy rate in text transcription, surpassing purely algorithmic approaches.

Cultural Engagement: 78% of users reported increased interest in Persian literary heritage, highlighting the platform's role in fostering cultural stewardship.

Challenges included balancing technical accessibility for non-experts and maintaining scholarly rigor. For instance, contributors occasionally misinterpreted archaic grammatical structures, necessitating iterative revisions. However, the platform's collaborative design allowed these errors to be flagged and corrected through community feedback, demonstrating the resilience of participatory models.

### **Conclusion:**

This study establishes citizen science as a viable and impactful strategy for preserving classical Persian literature in the digital age. The platform's success lies in its dual capacity to harness public enthusiasm and leverage computational efficiency, creating a sustainable ecosystem for corpus development. By decentralizing expertise, the model not only amplifies data volume but also democratizes cultural preservation, inviting global participation in safeguarding Iran's literary legacy. Future directions include expanding multilingual support for comparative studies and integrating advanced NLP models to handle poetic meter and metaphor analysis. Policymakers and cultural institutions are urged to adopt such frameworks to mitigate the risks of linguistic erosion in an increasingly digitized world.

# واکاوی نقش علوم شهروندی در شکل‌گیری و توسعه پیکره‌های زبانی ادبیات سنتی فارسی

مریم ابوالقاسمی<sup>۱\*</sup> | نادر نقشینه<sup>۲</sup>

۱. نویسنده مسئول، پژوهشگر پسادکتری دانشگاه تهران، تهران، ایران. رایانامه: mary.abolghasemi@ut.ac.ir

۲. دانشیار گروه علم اطلاعات و دانش‌شناسی، دانشگاه تهران، تهران، ایران. رایانامه: nnaghsh@ut.ac.ir

## چکیده

## اطلاعات مقاله

نوع مقاله: مقاله پژوهشی

تاریخ دریافت: ۱۴۰۲/۰۶/۱۱

تاریخ بازنگری: ۱۴۰۳/۰۳/۰۷

تاریخ پذیرش: ۱۴۰۳/۰۳/۰۹

## واژه‌های کلیدی:

علوم شهروندی،

پیکره زبانی،

فضای مجازی،

محتوای فارسی،

بازایی اطلاعات.

زبان فارسی با توجه به رسمی بودن آن در ایران و سه کشور دیگر، به عنوان نهمین زبان پرکاربرد در محتوا و درونمایه بالاتر از زبان عربی و ترکی و سایر زبان‌های خاورمیانه در نظر گرفته می‌شود. از این رو، پردازش زبان فارسی یک نیاز ملی و بین‌المللی است. این پژوهش به دنبال بررسی نقش علوم شهروندی در شکل‌گیری و توسعه پیکره‌های زبانی ادبیات سنتی فارسی است. هدف از این مطالعه، بررسی این موضوع است که چگونه می‌توان از مشارکت عمومی برای ایجاد و غنی‌سازی پیکره‌های زبانی ادبیات سنتی فارسی استفاده کرد. در این پژوهش ابتدا با رویکرد کیفی و به روش فراترکیب با استفاده از روش هفت مرحله‌ای سندلوسمکی به بررسی مولفه‌های پیکره و مشخصه‌ها و کاربردها پرداخته شده است و پس انجام مطالعات کتابخانه‌ای اقدام به طراحی سامانه و پیاده‌سازی اولین سکوی بومی علوم شهروندی با رویکرد ادبیات فارسی سنتی شده است. یافته‌ها مبین این امر بود که کدهایی همانند علوم شهروندی، یادگیری ماشین و یادگیری عمیق، داده، اطلاعات، فضای مجازی و مواردی مشابه این، در مقالات مورد نظر مورد توجه بیشتری قرار گرفته‌اند. همچنین ۱۵ کد دیگر با بیشترین فراوانی از این مقالات استخراج گردید که منجر به طراحی سامانه‌ای برای سکوی بومی علوم شهروندی شد. نتایج حاصل از تعامل کاربران با این شبکه‌های مجازی نشان می‌دهد که علوم شهروندی می‌تواند ابزاری مفید و کارآمد برای توسعه ادبیات سنتی فارسی در فضای مجازی باشد. این ابزار می‌تواند به افزایش حجم و تنوع داده‌های پیکره، ارتقای دقت و کیفیت داده‌ها، کاهش هزینه‌های جمع‌آوری و پردازش داده‌ها و افزایش تعهد و مشارکت عمومی در حفظ و ترویج ادبیات سنتی فارسی کمک کند.

**استناد:** ابوالقاسمی، مریم؛ نقشینه، نادر (۱۴۰۳). واکاوی نقش علوم شهروندی در شکل‌گیری و توسعه پیکره‌های زبانی ادبیات سنتی فارسی.

کرسی‌های نظریه‌پردازی و نوآوری در علوم انسانی، ۱(۱)، ۱-۲۱.



© نویسنده گان.

ناشر: دانشگاه رازی

DOI: 10.22126/TBIH.2024.9567.1000

## ۱. پیشگفتار

### ۱-۱. تعریف موضوع

یکی از مهم‌ترین ابعاد شهروندی داشتن «دانش و آگاهی شهروندی» است. دانش شهروندی از عقاید و اطلاعات بنیادی تشکیل می‌شود که افراد باید بدانند و به کار ببرند تا شهروندان مؤثر و مسئولی باشند (ایزدی و عزیزی شمامی، ۱۳۸۸: ۲۵). دانش شهروندی عبارت است از کسب دانش و جست‌وجوی دانایی، مشارکت فعال، مسئولیت‌پذیری و داشتن آگاهی در زمینه موضوع‌های تاریخی، اجتماعی و فرهنگی در سطح محلی، ملی و بین‌المللی (ساعی ارسی، ۱۳۹۹: ۲۴۵).

کندی (۱۹۹۸: ۳۳) احساس تعهد، برابری، کنجکاوی، ترقی‌طلبی، اطاعت از قانون، توجه به نفع عمومی و مشارکت را از جمله مؤلفه‌های شهروندی می‌داند. هندای<sup>۱</sup> (۲۰۱۰: ۶۸۵) مشارکت، تعاون، همکاری و دانایی محوری، و سواسون‌ای و همکاران<sup>۲</sup> (۲۰۱۶: ۵۲۰) آراستگی شخصی، دانش محوری در سطح محلی، ملی و جهانی، سخت‌کوشی، وطن‌دوستی، تبعیت از قانون، نوع‌دوستی و مشارکت را مدنظر قرار می‌دهند (همات و استاهلی<sup>۳</sup>، ۲۰۱۱: ۱۴). دستیابی به هدف‌هایی چون: وفاداری، افزایش دانش، تسلیم شدن در برابر قانون و هنجارهای اجتماعی، اعتقاد به ارزش‌های اصلی جامعه مانند تساوی و برابری، علاقه به مشارکت سیاسی و مهارت در تجزیه و تحلیل اطلاعات و ارتباطات سیاسی، فلسفه اصلی تربیت شهروندی است. بدون تردید، یکی از مهم‌ترین رسالت‌ها و دلایل شکل‌گیری نظام آموزشی در طول تاریخ بشری و توسعه آن به شکل کنونی، آماده‌سازی شهروندانی متعهد، فعال و خلاق بوده است. در گزارش منتشر شده با موضوع «مطالعه تربیت شهروندی انجمن بین‌المللی ارزشیابی و پیشرفت تحصیلی» آمده است: تمامی جوامع معاصر دارای این نگرانی و دل‌مشغولی هستند که چگونه افراد را برای زندگی و شهروندی آماده کنند و راه و رسم مشارکت در مسائل اجتماعی را به آن‌ها بیاموزند (پورترا<sup>۴</sup>، ۲۰۱۸: ۱۱۷).

همچنین یکی از پیش‌نیازهای توسعه‌ی پایدار مشارکت گسترده‌ی مردم در تصمیم‌سازی است. سیاست‌های خوب از فرایندی حاصل می‌شود که در آن عموم مردم می‌توانند در شکل‌دهی به توسعه سیاست، همکاری کنند. گروه‌های ذی‌نفع اغلب دانش منحصر به فرد و پیچیده‌ای دارند که برای درک پویایی‌های سیستم‌های اجتماعی مفید است. تعاملات رو به رشد با مردم در مسائل علمی و پایه و

1. Hand, E.

2. Swanson, A. et al

3. Hamatt, D., and Staeheli, C.

4. Purta

زنجیره نوآوری، به عنوان پاسخی به کاهش اعتماد عموم به علم و فناوری مطرح می‌شود. علم شهروندی به عنوان مجموعه‌ای از فعالیت‌های مشترک بین دانشمندان مبتدی و افراد غیرحرفه‌ای (شهروندان)، می‌تواند جریان دو طرفه اطلاعات را بین عموم، دانشمندان، مدیران و سیاست‌گذاران ارتقا دهد. علم شهروندی پژوهش‌های علمی است که تمام یا بخشی از آن توسط دانشمندان آماتور (و یا غیرحرفه‌ای) انجام می‌شود. دانشوری شهروندی گاهی اوقات با عنوان «مشارکت عمومی در پژوهش‌های علمی»، «نظارت مشارکتی» و «اقدام پژوهی مشارکتی» توصیف می‌شود. شهروندان دانش محور در یکی از مهم‌ترین و پرهزینه‌ترین مراحل تولید علم، یعنی جمع‌آوری داده‌های محیطی، نقشی اثرگذار ایفا می‌کنند. شهروند علمی به شهروندی اطلاق می‌شود که اطلاعات علمی وی باعث می‌شود در برابر دولت برای خود حقوق و وظایف ویژه‌ای قائل شود و از جمله این حقوق؛ حق مشارکت علمی است. در نتیجه ما با شهروندانی فعال سروکار داریم که در تصمیم‌گیری‌ها و سیاست‌گذاری‌های علمی کاملاً مشارکت می‌کنند.

## ۱-۲. ضرورت، اهمیت و هدف

زبان فارسی یکی از زبان‌های هندواروپایی در شاخه زبان‌های ایرانی جنوب غربی است که در کشورهای ایران، افغانستان (حدود ۲۰ میلیون نفر)، تاجیکستان (حدود ۵ میلیون نفر) و ازبکستان (حدود ۷ میلیون نفر) به آن سخن می‌گویند. فارسی زبان رسمی کشورهای ایران و تاجیکستان و افغانستان است. زبان رسمی کشور هندوستان نیز تا پیش از ورود استعمار انگلیس، فارسی بود. پردازش زبان فارسی یک نیاز ملی و بین‌المللی است که به رغم تفاوت مشهود تلاش‌های صورت گرفته در ایران و کشورهای دیگر بر روی پردازش رایانه‌ای زبان فارسی، نیازهای بسیاری در این زمینه به چشم می‌خورد و با نیم‌نگاهی به ادبیات کلاسیک فارسی نیاز به این موضوع بیش از پیش احساس خواهد شد.

به‌طور کلی می‌توان گفت که پردازش زبان طبیعی یکی از زیرشاخه‌های با اهمیت در حوزه گسترده‌ی هوش مصنوعی، و نیز در دانش زبان‌شناسی است. با استفاده از این روش می‌توان به ترجمه‌ی زبان‌ها پرداخت، از صفحات وب و بانک‌های اطلاعاتی نوشتاری، جهت پاسخ دادن به پرسش‌ها استفاده کرد، یا با دستگاه‌ها مثلاً برای مشورت گرفتن به گفتگو پرداخت.

رویکرد اصلی این پروژه نیز بر مبنای همین موضوع و با تاکید بر زبان فارسی کلاسیک خواهد بود. اهمیت پرداختن به این موضوع از حیث کمک به رشد و توسعه شناساندن زبان فارسی کلاسیک در سطوح ملی و بین‌المللی، چه از منظر فرهنگی و چه از منظر غنای ادبی مورد توجه خواهد بود. از سویی دیگر علم شهروندی امروزه به عنوان یک ابزار بسیار مهم و به سرعت در حال توسعه شناخته می‌شود که صرفه‌جویی بسیاری در هزینه پروژه‌های علمی به همراه دارد و با استفاده از شهروندان معمولی یک منبع طبیعی انسانی برای بررسی داده‌های تحقیقاتی را فراهم می‌آید. بر همین اساس بهره‌گیری از علوم شهروندی به عنوان رویکرد اصلی در واکاوی و شناسایی هرچه بهتر پیکره‌های زبانی فارسی کلاسیک می‌تواند نقش موثری را در نیل به اهداف این پژوهش از جمله تبیین نقش علوم شهروندی به منظور توسعه ادبیات فارسی کلاسیک در فضای مجازی، مشخص نمودن مؤلفه‌های پیکره‌های زبانی فارسی کلاسیک در فضای مجازی با بهره‌گیری از علوم شهروندی، طراحی و استقرار نمونه اولیه نرم افزار بومی و تبیین میزان تطابق ادارک نمونه اولیه با ادارک متخصصان علوم شهروندی، ایفا کند.

### ۱-۳. پرسش‌های پژوهش

- مؤلفه‌های پیکره‌های زبانی فارسی کلاسیک در فضای مجازی مبتنی بر علوم شهروندی کدامند؟
- مدل‌سازی فیزیکی (طراحی پلت فرم) نمونه پروژه علوم شهروندی در توسعه ادبیات فارسی کلاسیک چگونه است؟
- دسته‌بندی‌ها و شناخت پلت فرم بومی طراحی شده از ادبیات کلاسیک فارسی تا چه حد با ادارک خبرگان علوم شهروندی مطابقت دارد؟

### ۱-۴. پیشینه پژوهش

عاصی در پردازش دستوری زبان فارسی با رایانه جهت تمایز قائل شدن خط و زبان در ابتدا به برخی کاربردهای رایانه در حوزه خط اشاره نمود و سپس به نقش آن در حوزه زبان به‌ویژه ساخت واژه پرداخته است. (عاصی، ۱۳۸۴: ۹۲) ارسطوپور و احمدی‌نسب در پژوهشی به آسیب‌شناسی زبان و خط فارسی در بازایی اطلاعات با نگاهی به موتورهای کاوش و پایگاه‌های برخط پرداخته‌اند. این نوشتار عمده‌ترین چالش‌های مطرح در زمینه زبان و خط فارسی را در سه گروه رسم الخط، مسائل صرفی و مسائل معنایی مورد اشاره قرار داده و پس از بیان نمونه‌هایی در پیوند با هریک از این چالش‌ها، و ارائه ساختارهای درختی از انواع مسائل مطرح در این سه گروه، با نگاهی به پژوهش‌های صورت گرفته در

زمینه زبان فارسی و عربی و ضمن معرفی پیکره‌های فارسی مختلفی همچون همشهری، محک، بیجن‌خان و دادگان زبان فارسی به ارائه راهکارهایی جهت حل مسائل آن حوزه پرداخته است (گل تاجی و همکاران، ۱۳۸۹: ۱۹۲).

دشتبانی و همکاران به طراحی و ساخت پیکره‌ی متنی برای حوزه تخصصی فاوا پرداختند و در این پژوهش تمرکز بر ساخت پیکره‌ی دوزبانه‌ی فارسی انگلیسی حوزه فاوا است. این پیکره به صورت خودکار ساخته شده و منابع آن اسناد تخصصی حوزه فاوا است و با ساخت این پیکره هزینه و مدت زمان ساخت پیکره را کاهش می‌دهد و قابلیت مدیریت پیکره را نیز برای کاربران فراهم می‌کند. (مشهور و همکاران، ۱۳۹۹: ۸۰). محمودی و فرهودی در پردازش زبان فارسی در محیط رایانه و وب به چالش‌هایی که در حوزه رسم‌الخط فارسی و فاصله بین صنعت این حوزه یک صنعت خودکفا وجود دارد، همچنین به معرفی فعالیت‌هایی که در این راستا در پژوهشگاه ارتباطات و فناوری اطلاعات انجام شده اشاره کرده است. (یاری، ۱۳۹۴: ۸۷)

امامی و همکاران ابزارهای پایه زبان فارسی تولیدشده در پژوهشگاه توسعه فناوری‌های پیشرفته خواجه نصیرالدین طوسی را مورد ارزیابی قرارداده یافته‌ها نشان داد که مجموعه ابزارها و پیکره‌های پایه تولیدشده در پژوهشگاه خواجه نصیرالدین طوسی از دقت کافی برای استفاده در سامانه‌های پردازش زبان طبیعی برخوردار بوده و با استفاده از این ابزارها می‌توان به نتایج مطلوب‌تری دست یافت. تولید پیکره‌های دقیق و با حجم مناسب از روش‌های بهبود این ابزارهاست که در کنار تولید الگوریتم‌های جدید در پژوهشگاه مدنظر قرار گرفته است. (یاری، ۱۳۹۴: ۸۷)

میروز<sup>۱</sup> (۱۳۷۰: ۲۰۱۵) در مقاله خود به پروژه فرانک در حوزه شکل‌گیری و توسعه پیکره‌های زبانی پرداخت. این مقاله به معرفی پیکره فرانک می‌پردازد که یک پیکره زبانی از زبان فرانسه برای تحقیقات در حوزه پردازش زبان طبیعی است. این پیکره شامل متن‌های مختلفی از جمله اخبار، مقالات علمی، رمان‌ها و وبلاگ‌ها است.

صادقی و همکاران (۱۳۹۰: ۲۰۲۱) درباره با پیکره بی‌جن-خان<sup>۲</sup> پژوهشی انجام دادند. این مقاله به معرفی پیکره بیجن‌خان می‌پردازد که یک پیکره زبانی از زبان فارسی برای تحقیقات در حوزه پردازش زبان طبیعی است. این پیکره شامل متن‌های مختلفی از جمله اخبار، مقالات علمی و رمان‌ها است.

1. Meurers

2. Bijankhan Corpus



### ۵-۱. روش پژوهش و چارچوب نظری

به دلیل اینکه نمونه‌ای فراگیر از رفتار انتخاب منابع اطلاعاتی کاربران همراه با جامعیتی از مولفه‌های تشکیل دهنده این رفتار بدست نیامد، در پژوهش پیش رو در جهت شناخت مولفه‌ها و ابعاد مختلف رفتار انتخاب منابع اطلاعاتی کاربران، از رویکرد کیفی و روش فراترکیب بهره گرفته شده است. فراترکیب شیوه‌ای برای بررسی و تحلیل پژوهش‌های کیفی است که پس از تلفیق و مقایسه یافته‌های مجموعه‌ای از پژوهش‌های کیفی، نتایج تفسیری را تولید می‌نماید (زیمِر<sup>۱</sup>، ۲۰۰۶: ۳۱۱). از آنجا که فراترکیب یک رویکرد منسجم برای تجزیه و تحلیل داده‌های مطالعات کیفی می‌باشد به پژوهشگران اجازه می‌دهد تا یک پرسش پژوهی خاص را شناسایی کنند و در مرحله بعد برای رسیدن به پاسخ، به جست و جو، ارزیابی، انتخاب، جمع بندی و ترکیب شواهد کیفی بپردازند. این فرآیند از روش‌های کیفی دقیق برای تحلیل مطالعات کیفی موجود برای ساخت معنای بیشتر از طریق یک فرآیند تفسیری بهره می‌جوید (یک، آلیس جی<sup>۲</sup>، ۲۰۱۳: ۳۷). فراترکیب به روش‌های مختلفی انجام می‌شود. در این پژوهش، از روش فلگ<sup>۳</sup> (۲۰۲۴: ۲۳) استفاده شده است که شامل هفت مرحله می‌باشد و برای انجام پژوهش‌ها با روش فراترکیب بیشتر از باقی روش‌ها مورد استفاده بوده است. مراحل اجرای این روش در ادامه و مطابق با شکل زیر شرح داده می‌شود:



شکل ۱. مراحل هفت‌گانه فراترکیب (سندلوسکی و باروسو، ۲۰۰۷)

1. Zimmer
2. Yick, Alice G.
3. Flage

## ۲. پردازش تحلیلی موضوع

### ۲-۱. علوم شهروندی

علم شهروندی اصطلاحی کلی برای تعریف پروژه‌هایی است که متخصصان و غیر متخصصان با هم در گردآوری، ارزیابی و یا محاسبه داده‌های علمی همکاری دارند. پژوهشگران و سازمان دهندگان پروژه‌های مشارکتی شهروندی تعاریف گوناگون و گاه متضادی را برای مشارکت شهروندان ارائه می‌کنند، برخی فرآیندهای تعامل را توصیف می‌کنند، برخی دیگر نتایج را مشخص می‌کنند، و گروهی دیگر که هر دو تعریف را به کار می‌برند. علم شهروندی مردم جامعه را تشویق می‌کند تا در گردآوری اطلاعات در هر جایی و در بازه‌های زمانی کوتاه یا طولانی با دانشمندان و پژوهشگران همراهی و همکاری داشته باشند. (رضایی و بهزادی، ۲۹۹۰:۱۴۰۲) در علوم شهروندی داوطلب یا گروهی از داوطلبان بدون نیاز به داشتن هرگونه اطلاعات تخصصی در این زمینه مسئولیت انجام بسیاری از محاسبات، اندازه‌گیری‌ها یا حتی جستجو درباره موضوعات خاصی را بر عهده می‌گیرند و به این ترتیب به دانشمندان کمک می‌کنند تا به اهداف مورد نظر در اجرای پروژه‌های تحقیقاتی دست یابند. (ایزدی و عزیزی شمایی، ۱۳۸۸: ۲۵). اگرچه اجرای چنین پروژه‌هایی همواره و از گذشته‌های دور مورد توجه بسیاری از دانشمندان و محققان بوده است، اما در سال‌های اخیر اجرای چنین طرح‌هایی با توجه به اهمیت مشکلات شهروندان در گسترش مرزهای علم و دانش بیش از پیش مورد توجه قرار گرفته است (سکارونی و همکاران<sup>۱</sup>، ۲۹:۲۰۱۹).

از آنجایی که در دنیای امروز فناوری‌های نوین نقش بسیار مهمی را در انجام تحقیقات علمی ایفا می‌کنند بنابراین مشارکت شهروندان عادی در اجرای پروژه‌های مبتنی بر علم شهروندی در مقایسه با گذشته که از این پروژه‌ها بیشتر به عنوان یک ابزار آموزشی برای علاقه‌مندان به حوزه‌های علمی مختلف استفاده می‌شد با افزایش چشم‌گیری مواجه بوده است. نخستین طرح از مجموعه طرح‌های علوم شهروندی از سال ۱۹۰۰ میلادی آغاز شد و از آن پس مجموعه‌های متنوعی از این طرح‌ها در زمینه‌های مختلف نظیر مدلسازی تغییرات آب و هوایی، پرنده‌نگری، رمزگشایی از پروتئین‌ها، گرمایش جهانی و غیره به اجرا درآمده است (روبرتس<sup>۲</sup>، ۲۰۱۹: ۴۲۳).

1. Ceccaroni et al.

2. Roberts

"مشارکت عمومی در علم"<sup>۱</sup> طی دو دهه گذشته به طور گسترده مورد بحث، شناسایی، تایید و پیاده‌سازی قرار گرفته است و از جمله تحت عنوان "علم شهروندی"<sup>۲</sup> به عنوان راهکاری برای شکل-دهی آینده‌ای متفاوت برای علم دنبال شده است. (هندای، ۲۰۱۰). طیف معانی که با مفهوم علم شهروندی در ارتباط است به دو دسته عمده قابل دسته‌بندی است: یکی مربوط به حالتی که شهروندان به عنوان ابزاری برای تقویت مشروعیت و قدرت نهاد علم بکار می‌روند، و دیگری مربوط به حالتی که شهروندان به عنوان منبعی نو و مکمل برای دانش و قدرت دیده می‌شوند تا پارادایم کنونی علم را به چالش بکشند. مفهوم علم شهروندی ابتدا متمرکز بر روابط عمومی علم و انتقال بهتر نتایج علمی به عموم مردم بود تا بدین وسیله مردم درک بهتری از علم پیدا کنند. این رویکرد اولیه در دسته اول جای می‌گیرد و باعث شد تا علم شهروندی به دنبال آموزش، آگاهی‌بخشی و افزایش اشتیاق عامه در شناخت علم باشد تا بدین وسیله سوءبرداشت‌ها و نقدهای عامه پیرامون علم را کاهش دهد. (سواسونای و همکاران، ۲۰۱۶: ۵۲۰). رویکرد اصلی دیگر در علم شهروندی به دنبال درگیر نمودن شهروندان در فرآیند علم و تولید دانش بوده است، و این کار را عمدتاً از طریق گردآوری داده از طریق اجتماعات و گروه‌های مردمی انجام داده است تا بدین وسیله بر محدودیت‌های زمانی، مالی و منابع انسانی جامعه علمی غلبه نماید. توسعه این سازوکار گردآوری و پایش محیط‌های طبیعی و انسانی هم‌راستا است با تعریف علم و نهاد علم همانگونه که هست.

مفهوم‌پردازی علم شهروندی در مکتب انگلستان که با نام آلن ایروین گره خورده است بر شهروندی علمی<sup>۳</sup> و سیاستگذاری شهروندمدارانه<sup>۴</sup> متمرکز است و در دسته دوم از دسته‌بندی دوگانه اخیر جای می‌گیرد. در این رویکرد، علم رویکردی گفتگویی (دیالوگی) و از پایین به بالا به خود می‌گیرد به طوری که مشارکت در سیاستگذاری علم توسط شهروندان به تغییرات در روابط قدرت و حکمرانی نهاد علم منجر می‌شود (سواسونای و همکاران، ۲۰۱۶: ۵۲۰). مبتنی بر نقدهای تاریخی، سیاسی، اجتماعی و نقد قدرت به نهاد علم، این دیدگاه سیاست‌محور به مفهوم علم شهروندی به دنبال نقشی فعال برای شهروندان در تصمیم‌گیری، مردم‌سالاری، امور سیاسی و روابط قدرت نهاد علم است. با این وجود، این تغییرات همگی متوجه نهاد قدرت هستند و به موضوع و ماهیت علم کاری ندارند. به

1. Public engagement with science

2. Citizen Science (CS)

3. Scientific citizenship

4. Citizen policy

عبارتی در این رویکرد نیز هنوز شهروندان به مثابه منبعی ارزشمند از دانشی که مکمل دانش علمی باشد دیده نمی‌شوند و صرفاً قرار است در تسهیم و تعدیل قدرت نهاد علم مشارکت داشته باشند بدون اینکه مشروعیت ماهوی داشته باشند. (پیترسون و نولز<sup>۱</sup>، ۲۰۰۹: ۴۰).

از سوی دیگر شاهد نقش فزاینده دانش ضمنی و تجربه‌محور کنشگران اجتماعی و محلی هستیم که در رویکرد کلاسیک به علم جایگاهی ندارند. اینکه آیا دانش ضمنی و تجربه‌محور نسبتی با تعریف علم شهروندی دارد یا خیر و آیا می‌تواند (باید) به بسط مفهومی آن منجر شود سوالی است که در این مقاله به دنبال پاسخ تجربی بدان است. چنین پروژه فکری و کاربردی می‌تواند منابع علم را چندگانه نموده و به غنی شدن فرآیند یادگیری اجتماعی در طی مسیر توسعه پایدار کمک نماید. همچنین این رویکرد همراستا با تلاش‌های جدید برای مفهوم‌پردازی تولید دانش مسئله‌محور و فرارشته‌ای است (ایزدی و همکاران، ۲۰۱۳: ۲۷).

علم شهروندی برگرفته از جنبش شهروند دانشور است. جنبش علم شهروندی باهدف ارتباط مردم با علم و پر کردن شکاف بین مردم و دانشمندان به وجود آورده است. (اروالکوا و جانسکوا<sup>۲</sup>، ۲۰۱۹: ۰۳) این امر به عموم مردم این امکان را می‌دهد که مشارکت مستقیمی در تحقیقات علمی داشته باشند (یانگ و دیگران<sup>۳</sup>، ۲۰۱۹: ۰۲).

به طور کلی، علم شهروندی به پژوهشی علمی گفته می‌شود که محققین آن را افراد عادی جامعه تشکیل می‌دهند. به عبارت دیگر، علوم جمعی به معنی مشارکت عموم مردم در یک پژوهش علمی می‌باشد. این مشارکت در سطوح مختلفی انجام می‌شود. از کوشش‌های ذهنی برای حل یک مسئله بزرگ تا استفاده از ابزارها و تجهیزات مشارکت‌کنندگان در پروژه، همگی نوعی مشارکت به حساب می‌آیند (تروژان و همکاران<sup>۴</sup>، ۲۰۱۹: ۲۵۴).

## ۲-۲. وضعیت کنونی علم شهروندی

امروزه شهروندان دانشور در پروژه‌های تغییرات اقلیمی، گونه‌های مهاجم، زیست‌شناسی حفاظت، بهسازی محیط، پایش کیفیت آب، جمعیت‌شناسی جانوران، بررسی وضعیت رشد گیاهان و بسیاری دیگر از فعالیت‌هایی شرکت می‌کنند که به پایش و مشاهده مستقیم در منابع طبیعی مربوط می‌شود. در

1. Peterson, A., & Knowles, C.

2. Urválková, E. S., & Janoušková, S.

3. Yang, D. et al.,

4. Trojan, J. et al.

این باره، تکنولوژی و به‌ویژه اینترنت، مهم‌ترین وظیفه را به عهده دارند؛ اما نحوه مشاهده و ثبت آن، اهمیت به‌مراتب بالاتری دارد (سکارونی و همکاران، ۲۰۱۹: ۱۱۷).

امروزه هر پروژه‌ای که به داده‌های انبوه زمینی نیاز داشته باشد، بی‌تردید باید از کمک شهروندان دانشور بهره‌مند شود. یکی از نخستین فعالیت‌ها در زمینه ثبت مشاهده جانوران، سیستمی است به نام سایبرترکر. وسیله‌ای شبیه جی‌پی‌اس‌های دستی که برای شکارچیان بیابان کالاهاری طراحی شد. کسانی که سواد خواندن و نوشتن ندارند، اما رزنان قابل و شکارچیان ماهری هستند. این سیستم توسعه پیدا کرد و در کشورهای دیگر مانند اسپانیا، آفریقای جنوبی و مغولستان نیز به کار گرفته شد. (لهجسک جی و همکاران<sup>۱</sup>، ۲۰۱۹: ۱۹).

دیده‌بان زمین یا «ارث‌واچ» مؤسسه‌ای است غیرانتفاعی که سالانه از داده‌های هزاران دانشور شهروند، در پروژه‌های مختلف بهره می‌برد؛ کسانی که برای تعطیلات به سفر می‌روند یا در شهرهایشان در تردد هستند، می‌توانند هم‌زمان به پروژه‌های تحقیقاتی دانشمندان یاری رسانند. (آلهومادا و همکاران<sup>۲</sup>، ۲۰۲۰: ۴۷). «مؤسسه ملی برای مطالعه گونه‌های مهاجم<sup>۳</sup>» در کلرادو آمریکا نیز برای بررسی وضعیت گونه‌های مهاجم، از داده‌های شهروندان استفاده می‌کند تا روند تغییرات و وضعیت این گونه‌های جانوری یا گیاهی را در سراسر کشور مورد ارزیابی قرار دهد (دلرمن و همکاران<sup>۴</sup>، ۲۰۱۹: ۱۵). «آی‌نچرالست<sup>۵</sup>» از دیگر نمونه‌های موفق دانشوری شهروندی است. یک شبکه اجتماعی که برای علاقه‌مندان به طبیعت راه‌اندازی شده تا هر گونه‌ای را که می‌بینند همراه با عکس و موقعیت جغرافیایی آن، از طریق تلفن همراه خود ثبت کنند. عکس و اطلاعات مربوط به گونه‌ها که تازه‌واردان به این شبکه ارسال می‌کنند، باید از سوی متخصصان یا کسانی که پیش‌تر صلاحیتشان در شناسایی گونه‌ها اثبات شده است، تأیید شود. این روش به جمع‌آوری همگانی داده‌ها معروف است (رودین<sup>۶</sup>، ۲۰۱۹: ۲۰۶).

در مقیاسی کوچک‌تر، در حومه تورنتو در کانادا نیز در سال ۲۰۱۶، پروژه‌ای با کمک شهروندان و عموماً دانش‌آموزان آغاز شد تا میزان تصادفات جاده‌ای لاک‌پشت، قورباغه، پستانداران، مارها و

1. Lehejcek, J. et al.,

2. Ahumada, J. et al.,

3. The National Institute of Invasive Species Science (NIISS)

4. Dellermann, D. et al.

5. INaturalist

6. Rudin

پرندگان در یک منطقه ثبت و ضبط شود. برای این کار، نخست ۲۲ داوطلب انتخاب شدند و مورد آموزش قرار گرفتند تا گونه‌ها را از هم تشخیص دهند و بتوانند با اپلیکیشنی که برای همین منظور طراحی شده بود، کار کنند. تبلتی به هر یک از این داوطلبان داده شد تا دو روز در هفته بین ساعت ۹ صبح تا ۱۲ ظهر، جانوران مرده را در یک کیلومتر و ۳۰۰ متر از چهار جاده نزدیک به هم، ثبت کنند. این داوطلبان بیش از ۶۰۰ مورد تصادف جانوران با خودرو را ثبت کردند. بالاترین میزان تصادفات به قورباغه و کمترین آن به مار اختصاص داشتند (آلهومادا و همکاران، ۱۴۰۲: ۴۷).

از سویی دیگر بهره‌گیری از دانش شهروندی در توسعه زبان‌های لاتین بسیار مورد توجه قرار گرفته که بررسی آثار شکسپیر آثار حماسی ایلید و ادیسه، نمونه‌ای از آن است. لذا اکنون نیاز به استفاده از این علوم در توسعه ادبیات فارسی کلاسیک به‌ویژه در فضای مجازی بیش از پیش احساس می‌شود، و بدون شک با توجه به موارد مذکور شاهنامه می‌تواند به‌عنوان یکی از بهترین منابع مورد مطالعه و مناسب برای این امر تلقی گردد.

با وجود سابقه درخشان ادبیات سنتی فارسی، به‌ویژه شاهنامه، در عصر حاضر کم‌توجهی به چنین گنجینه‌های ارزشمند بیشتر شده است. از سویی دیگر، علوم شهروندی به‌عنوان روشی نوین، فرصتی برای جلب مشارکت عمومی در حل مسائل مختلف، از جمله حفظ و نشر میراث فرهنگی، فراهم کرده است. با استفاده از علوم شهروندی می‌توان به‌طور فزاینده‌ای شاهد پویایی و بالندگی ادبیات سنتی فارسی، به‌ویژه شاهنامه، در فضای مجازی بود. می‌توان گفت که دانش شهروندان به‌عنوان یک منبع غنی از تجربیات و دیدگاه‌های مختلف، با ارتقاء آگاهی‌هایشان از ادبیات سنتی فارسی و به‌اشتراک‌گذاری آن در فضای مجازی، می‌تواند به ترویج و توسعه این ادبیات کمک کند. این افزایش آگاهی و فراهم آوردن یک بستر برای تبادل نظرها و اندیشه‌ها، می‌تواند باعث تحول در فهم و تفسیر ادبیات سنتی فارسی شده و به‌ویژه به تداوم جذب علاقه‌مندان درباره‌ی این ارزش‌های فرهنگی کمک کند.

سوال اصلی این پژوهش آن است که چگونه می‌توان از علوم شهروندی برای توسعه ادبیات سنتی فارسی، به‌ویژه طراحی سیستمی برای پلت فرم بومی علوم شهروندی استفاده کرد؟

### ۳-۲. شکاف نوآوری در تحقیق

از جمله نوآوری‌های مورد نظر در این پژوهش بهره‌گیری از علوم شهروندی در بررسی و واکاوی پیکره‌های زبانی ادبیات فارسی در فضای مجازی خواهد بود. استفاده از علوم شهروندی هم‌اکنون در

زمینه‌های مختلف علمی-پژوهشی در حال فراگیر شدن است و ورود آن به موضوع زبان و ادبیات فارسی می‌تواند دریچه‌های بسیار زیادی به سوی توسعه این امر در سطح ملی و بین‌المللی بگشاید. از طرفی باتوجه به بررسی انجام شده، پژوهش‌های گوناگونی در زبان‌های لاتین و عربی از حیث کتابت و فونت و باگ‌های کتابت سنگی و دستی انجام شده ولی در این پژوهش علاوه بر تاکید بر پیکره‌های زبان و ادبیات فارسی به مفاهیم و نحو و همچنین تاثیر علوم شهروندی بر آن‌ها توجه بیشتری خواهد. از طرفی با توجه به گسترش روز افزون ارتباطات و فناوری اطلاعات در میان افراد جامعه، نیاز به ایجاد پلت فرم‌های بومی مرتبط با موضوع پژوهش بسیار احساس می‌شود ولی متأسفانه تا کنون اقدامی در این خصوص صورت نگرفته است. لذا نتایج پژوهش حال حاضر نه تنها می‌تواند زمینه‌ساز توسعه علوم شهروندی در موضوعات مذکور و فراهم آوردن جایگاهی برای آن باشد، بلکه در نهایت با ایجاد اولین پلت فرم بومی در این حوزه گامی بسیار نو و اثربخش در این زمینه برداشته شده است.

## ۲-۴. روش تنظیم پرسش‌های پژوهش

پرسش‌های مبنایی پژوهش به شرح زیر تنظیم شدند.

### جدول ۱. پرسش‌های گام نخست فراترکیب

| پرسش‌ها                 | پاسخ                                                                                      |
|-------------------------|-------------------------------------------------------------------------------------------|
| چیستی کار (What)        | شناسایی مولفه‌های پیکره‌های زبان فارسی سنتی در فضای مجازی                                 |
| جامعه مورد مطالعه (Who) | آثار مختلف اعم از مقاله (در مجلات و کنفرانس‌ها) و پایان‌نامه‌ها                           |
| محدودیت زمانی (When)    | کلیه آثار موجود بین سال‌های ۲۰۱۴ تا ۲۰۲۱ میلادی و ۱۳۹۳ تا ۱۴۰۰ شمسی                       |
| چگونگی روش (How)        | بررسی موضوعی جستجوها و کدهای استخراج شده، تعیین مفاهیم، دسته‌بندی مفاهیم و ایجاد مقوله‌ها |

## ۲-۵. جست وجوی منابع

در این پژوهش، به جست و جو در میان مقالات و پایان‌نامه‌های انتشار یافته در پنج پایگاه اطلاعاتی فارسی و شش پایگاه اطلاعاتی انگلیسی بین سال‌های ۲۰۱۴ تا ۲۰۲۱ میلادی و ۱۹۳ تا ۱۴۰۰ شمسی پرداخته شده است. در این بازه زمانی به جهت تعدد پایگاه‌های مطرح و مهم مورد بررسی، دستیابی به آثار مناسب، بروز، جامع و متعدد در این حوزه میسر شد. برای این هدف طبق روش تحقیق و فرآیند پیشنهادی، خروجی فراترکیب ما در دو بخش «بررسی مولفه‌های پیکره‌های زبانی» و «مطالعه علوم شهروندی و کاربرد آن در پیکره‌های زبانی»، کلیدواژه‌های مدنظر برای جستجوی منابع طبق جداول

زیر در نظر گرفته شد تا بتوان به حداکثر بازیابی ممکن رسید. این جست‌وجوها در مرداد ۱۴۰۰ بروزرسانی شده است.

جدول ۲. کلیدواژه‌های مورد استفاده در جست‌وجوی منابع

| کلیدواژه‌های فارسی | کلیدواژه‌های لاتین  |
|--------------------|---------------------|
| علم شهروندی        | Citizen Science     |
| پیکره زبانی        | Linguistic Corpus   |
| فضای مجازی         | Cyber/virtual Space |
| پیکره زبانی        | Linguistic Corpus   |
| یادگیری ماشین      | Machine Learning    |
| یادگیری عمیق       | Deep learning       |
| تحلیل متن          | Text mining         |

کلیدواژه‌های نام برده در جدول همراه با ترکیبات مختلف آن‌ها، با توجه به شرایط موردنظر در پژوهش حاضر و راهبرد جست‌وجوی پیشرفته، در هر کدام از پایگاه‌های اطلاعاتی ذکر شده وارد گردید. پس از آن با حذف منابع تکراری بازیابی شده در هر پایگاه، تعداد منابع بازیابی شده در هریک از پایگاه‌های اطلاعاتی مورد جست‌وجو به شرح جدول ۳ بدست آمد.

جدول ۳. پایگاه‌های اطلاعاتی مورد جست‌وجو

| نام پایگاه فارسی   | تعداد مقاله | تعداد پایان‌نامه |
|--------------------|-------------|------------------|
| SID                | ۳۱          | -                |
| Magiran            | ۱۳۷         | -                |
| Ensani             | ۳۰          | -                |
| Noormag            | ۲۳۹         | -                |
| Civilica           | ۱۷          | -                |
| مجموع              | ۴۵۴         | -                |
| نام پایگاه انگلیسی | تعداد مقاله | تعداد پایان‌نامه |
| Emerald            | ۲۰۷۹        | -                |
| Scopus             | ۲۵          | -                |
| Taylor & Francis   | ۱۷۷         | -                |
| Science Direct     | ۱۱۳         | -                |
| JStore             | ۶۰          | -                |
| ProQuest           | ۲۰۴         | ۱۰۵              |
| مجموع              | ۲۶۵۸        | ۱۰۵              |



## کل مدارک بازاریابی شده: ۳۲۱۷

به‌طور کلی ۳۲۱۷ منبع برای غربالگری و بررسی بازاریابی شدند. از میان آنها ۹۱۴ منبع به زبان فارسی و ۲۳۰۳ منبع به زبان انگلیسی بودند.

## ۲-۶. انتخاب منابع مناسب

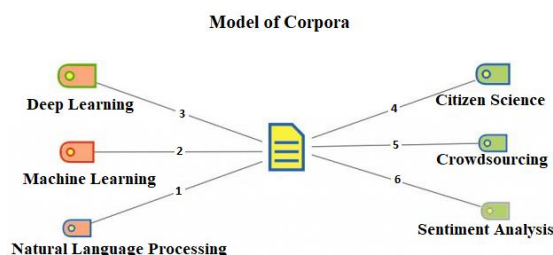
انتخاب منابع بر مبنای اصول ورود و خروج به فراترکیب و از راه غربالگری عنوان، محتوا، چکیده و روش‌شناسی صورت گرفت. پژوهش‌هایی معیارهای ورود به فراترکیب را داشتند که در آن‌ها به فرآیند، مولفه‌ها، ابعاد و عوامل اثرگذار بر مولفه‌های پیکره‌های زبان فارسی سنتی در فضای مجازی و علوم شهروندی پرداخته شده بود. در ارتباط با معیارهای خروج از فراترکیب، به پژوهش‌هایی پرداخته شد که اطلاعات کافی در زمینه اهداف این تحقیق گزارش نداده بودند، فاقد الگوی روش شناختی مناسب بودند، با اهداف و عناوین یکسان انجام شده بودند و یا کیفیت لازم علمی را نداشتند. تعدادی از منابع در هر مرحله از غربالگری حذف شدند.

در نهایت از بین ۳۲۱۷ منبع، با جست و جوی عمیق، ۳۹ منبع که همه آن‌ها به زبان انگلیسی و از مقاله‌های پژوهشی بودند تأیید شدند و به دلیل نبود وجود اطلاعات کافی در زمینه اهداف مورد نظر از میان منابع فارسی منبعی در انجام پژوهش‌ها تأیید نشد.

## ۲-۷. استخراج اطلاعات از متون

پس از انتخاب منابع مناسب، مجموع ۳۹ منبع منتخب تحت مطالعه ای چندباره قرار گرفتند و سپس اطلاعات آن‌ها شامل اطلاعات کتاب‌شناختی، اطلاعات روش شناختی و همچنین مولفه‌ها یا موضوعات و متغیرهای مهم مورد اشاره در منابع استخراج شد.

## ۲-۸. تجزیه و تحلیل و ترکیب یافته‌ها



شکل ۲. مقوله‌های استخراج شده از روش فراترکیب

باتوجه به خروجی نرم‌افزار در شکل ۳، در این مرحله، کدهای بدست آمده در مرحله قبل بر اساس مشابهت در گروهی با عنوان مفاهیم مشترک قرار گرفتند و مفاهیم مشابه نیز تحت عنوان یک مقوله دسته بندی شدند. ۳۹ مستند نهایی با کمک نرم افزار MaxQDA2020، مورد بررسی و تحلیل محتوا قرار گرفت و با استفاده از روش کدگذاری ۳۳ کد اصلی و ۶ مقوله اصلی بدست آمد. ۶ مقوله نام برده در تصویر زیر نشان داده شده است:



شکل ۳. مقوله‌های استخراج شده از روش فراترکیب

## ۹-۲. کنترل کیفیت کدهای استخراجی

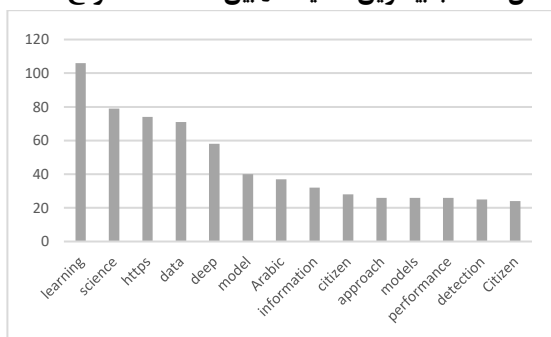
در این مرحله از شاخص کاپا<sup>۱</sup> برای کنترل کیفیت کدهای بدست آمده و حفظ پایانی مدل طراحی شده استفاده شده است. بر مبنای این روش، از شاخص دیگری که اطلاع از دسته‌بندی قبلی نداشت درخواست شد تا دسته بندی اطلاعات استخراج شده از متون را انجام دهد. سپس، دسته بندی ایشان با دسته بندی پژوهشگران مقایسه و شاخص کاپا اندازه گیری شد. اندازه ضریب کاپا بین ۱- و ۱+ است. میزان بدست آمده هرچه به ۱+ نزدیک تر باشد، نشان از توافق بیشتر بین دو کدگذار می‌دهد. اگر شاخص مورد نظر بیشتر از ۰/۶۱ باشد، از مطلوبیت و معتبر بودن برخوردار است. در این راستا برای کدگذاری، متون مورد بررسی در اختیار یکی از خبرگان حوزه هوش مصنوعی با تجربه در تحلیل متن قرار گرفت و با استفاده از ضریب کاپا ضریب توافق دو کدگذار در نرم افزار اس‌پی‌اس‌اس محاسبه شد. با توجه به سطح معناداری که برابر با ۰.۰۰۰ و کوچک تر از ۰.۵۳ است و میزان ضریب کاپا که ۰.۷۹۶ و بیشتر از ۰/۶ بود، بین نظرات دو گروه کدگذار رابطه‌ای معنادار وجود داشته و در نتیجه استخراج کدها از پایایی مناسبی برخوردار می‌باشد.

۱. ضریب کاپا یک معیار آماری است که برای اندازه‌گیری توافقی بین دو مشاهده گر یا دو متغیر دودویی استفاده می‌شود. این ضریب از ۰ تا ۱ ارزیابی می‌شود، که مقدار ۱ نشان‌دهنده توافقی کامل بین دو مشاهده گر و ۰ نشان‌دهنده توافقی تصادفی است. این معیار معمولاً در زمینه‌هایی مانند علوم انسانی، علوم اجتماعی و مطالعات روان‌شناسی برای ارزیابی توافقی بین مشاهده‌گران یا دقت مدل‌های طبقه‌بندی استفاده می‌شود.

## ۱۰-۲. ارائه یافته‌ها

در مجموع مشاهده شد کدهایی همانند علوم شهروندی، یادگیری ماشین و یادگیری عمیق، داده، اطلاعات، فضای مجازی و مواردی مشابه این، در مقالات مورد نظر مورد توجه بیشتری قرار گرفته‌اند. ۱۵ کد با بیشترین فراوانی استفاده در شکل زیر نمایش داده شده‌اند.

شکل ۴. کد با بیشترین اهمیت از بین ۹۵ کد استخراج شده



همچنین ارتباط میان کدها و مقوله‌ها در شکل زیر نمایش داده شده است:



شکل ۵. ارتباط میان کدها و مقوله‌های استخراج شده

ورودی‌های طراحی سیستم علوم شهروندی بومی مورد نظر، بر اساس مقوله و کدهای استخراج شده مشخص شد. بر همین اساس چهار مولفه یادگیری عمیق، یادگیری ماشین، پردازش زبان طبیعی و تحلیل احساسات به عنوان عناصر اصلی تشکیل دهنده سیستم مورد نظر مشخص شدند. به علاوه می‌توان به عنوان رکن اصلی این سیستم مولفه‌های علوم شهروندی و جمع‌سپاری را مد نظر قرار داد. همچنین به پیکره‌های زبان و ادبیات فارسی سنتی از دیگر ورودی‌های این سیستم می‌توان اشاره نمود. بر همین اساس سیستم طراحی شده بر مبنای یافته‌های مراحل مطالعاتی و کتابخانه‌ای پژوهش، و همچنین بررسی نمونه‌های داخلی و بین‌المللی، در ادامه نمایش داده شده است:



شکل ۶. سیستم طراحی شده برای پلت فرم بومی علوم شهروندی

نتایج خروجی حاصل از این سیستم، با مشارکت شهروندان و دریافت تعامل ایشان، به ارزیابی نمونه‌های اولیه و تحلیل داده‌ها می‌انجامد و سیستم براساس بازخوردهای حاصل قابلیت توسعه خواهد داشت.

### ۳. نتیجه‌گیری

این پژوهش با هدف واکاوی نقش علوم شهروندی در شکل‌گیری و توسعه پیکره‌های زبانی ادبیات کلاسیک فارسی در فضای مجازی صورت گرفت. در ابتدا با بهره‌گیری از روش فراترکیب به بررسی مولفه‌های پیکره و مشخصه‌ها و کاربردها پرداخته شد. لذا همانطور که اشاره شده است، پژوهش‌های

گونگونگی در حوزه‌های مرتبط با یادگیری ماشین، مدل‌سازی موضوعی، پیکره‌های زبانی و حتی علوم شهروندی به تفکیک انجام شده است اما متأسفانه در هیچ موردی پژوهش‌های معتبری در قالب مقالات علمی در خصوص نقش و کاربرد علوم شهروندی در توسعه ادبیات سنتی فارسی صورت نپذیرفته بود. جزو معدود فعالیت اجرایی در این حوزه در حال حاضر می‌توان به تکمیل، شناسایی و اصلاح آثار شکسپیر به‌عنوان یکی از پروژه‌های عملیاتی دانشگاه زوریخ اشاره نمود که منتج به ارائه برخی گزارش‌ها بصورت آنلاین در سایت دانشگاه مذکور شده است.

از طرفی با گسترش روزافزون فضای مجازی متأسفانه به نظر می‌رسد که پرداختن به ادبیات فارسی سنتی به شیوه‌های نوین امروزی در پژوهش‌های محققان گرامی اندکی مغفول مانده فلذا عملاً هیچ پژوهشی مشابهی در خصوص نقش علوم شهروندی در شکل‌گیری و توسعه پیکره‌های زبانی ادبیات کلاسیک فارسی وجود ندارد.

علی‌رغم تمامی نقیصه‌های موجود در پیشینه تحقیق، پس از انجام مطالعات کتابخانه‌ای اقدام به طراحی سامانه و اجرای اولین سکوی بومی علوم شهروندی با رویکرد ادبیات فارسی سنتی شد و بررسی‌های اولیه حاصل از تعامل کاربران، موفقیت اولیه و کارایی مناسب سیستم در مراحل تست و اجرای سکوی صورت گرفت.

با این که برخی افراد با نگرشی منفی پروژه‌های علم شهروندی را سطحی و نه چندان عمیق می‌دانند، اما واقعیت آن است که با طراحی درست، داشتن ابزار مناسب و نیز آموزش مطلوب، می‌توان به نتایجی شگفت‌انگیز از این پروژه‌ها دست یافت. نتایج چنین پروژه‌های در حوزه‌ها و علوم مختلف به خودی خود مویید این موضوع است. اما به هر روی با توجه به نقص‌های پژوهشی حال حاضر در زمینه تاثیر علوم شهروندی بر روی ادبیات فارسی کلاسیک و حتی زبان و ادبیات فارسی به‌شیوه مورد استفاده امروزی در فضای، به نظر می‌رسد که باید توجه ویژه‌تری به این امر داشت. همچنین بر اساس نتایج حاصل از تعامل کاربران در این پژوهش، این موضوع انکارناپذیر است که در صورت بهره‌گیری از پروژه‌های علوم شهروندی، می‌توان افراد بیشتری را به آشنایی و بررسی آثار ادبیات فارسی کلاسیک ترغیب نمود.

از طرفی شبکه‌های بومی تهیه شده در حال حاضر به‌عنوان اولین و تنها شبکه بومی علوم شهروندی به زبان فارسی (طراحی و تولید در دانشگاه تهران)، می‌تواند ضمن فراهم آوردن بستری مناسب برای پژوهش‌های آتی این امکان را برای علاقه‌مندان و متخصصان امر فراهم آورد که با توسعه بستر

نرم‌افزاری آن گام‌های مثبتی در جهت شناساندن و اعتلای زبان شیرین پارسی چه در سطح ملی و چه در سطح بین‌المللی بردارند.

## منابع

- از پیکره زبانی تا زبان شناسی پیکره‌ای: عاصی، مصطفی (۱۳۸۲). مجموعه مقالات پنجمین کنفرانس زبان شناسی کشور، تهران، دانشگاه علامه طباطبائی، شماره ۱۴، ۷۵-۹۲.
- ایزدی، صمد، و عزیزی شمامی، مصطفی (۱۳۸۸). «ضرورت برخورداری دانشجویان از دانش شهروندی و مسئولیت‌پذیری اجتماعی در فرایند جهانی شدن». فصل‌نامه مطالعات برنامه درسی، شماره ۱۵، ۲۵-۳۲.
- فرهادی، الهام، و کسرای، محمد سالار، و توسلی رکن آبادی، مجید (۱۴۰۲). آموزش شهروندی در سند تحول بنیادین آموزش و پرورش جمهوری اسلامی ایران. جامعه شناسی سیاسی انقلاب اسلامی، ۴(۲)، ۲۱۱-۲۳۱.
- DOI: 20.1001.1.27835200.1402.4.2.7.4
- فناوری نانو و علم شهروندی: رضایی، آرزو، و بهزادی، آیدا، (۱۴۰۲). مجموعه مقالات هفتمین کنگره سالانه بین‌المللی توسعه کشاورزی، منابع طبیعی، محیط زیست و گردشگری ایران، تبریز، ۲۹۸-۳۱۵.
- گل تاجی، محمد، و بذرگر، سارا (۱۳۸۹). بررسی مشکلات ریخت شناسی زبان فارسی در سه پایگاه اطلاعاتی مرکز منطقه‌ای اطلاع‌رسانی علوم و فناوری. پژوهشگاه اطلاعات و مدارک علمی ایران و جهاد دانشگاهی، کتابداری و اطلاع‌رسانی، ۱۳ (۲): ۱۹۱-۲۱۴.
- مشهور، پروین دخت، و فقیری، غلام محمد (۱۳۹۱). «بررسی و تحلیل انواع زمان در دفتر اول، دوم و سوم مثنوی مولوی با رویکرد سبک شناسی رایانشی - پیکره‌ای». مجله زبان‌شناسی و گویش‌های خراسان، دانشگاه فردوسی مشهد، ۶، ۹۹-۷۹.
- یاری، علیرضا (۱۳۹۴). بررسی پیکره‌ها و ابزارهای پردازش زبان طبیعی. گزارش طرح پژوهشکده ارتباطات و فناوری اطلاعات، شماره ۱۵، ۸۷-۹۷.

## References

- Ahumada, J. A., Fegraus, E., Birch, T., Fores, N., Kays, R., O'Brien, T. G., et al. (2020). Wildlife insights: A platform to maximize the potential of camera trap and other passive sensor wildlife data for the planet. *Environmental Conservation*, 47(1).
- Ceccaroni, L., Bibby, J., Roger, E., Flemons, P., Michael, K., Fagan, L., & Oliver, J. L. (2019). Opportunities and risks for citizen science in the age of artificial intelligence. *Citizen Science: Theory and Practice*, 4(1), 29.
- Meurers, D. (2015). *Learner corpora and natural language processing*. The Cambridge handbook of learner corpus research, 537-566.
- Dellermann, D., Calma, A., Lipusch, N., Weber, T., Weigel, S., & Ebel, P. (2019). The future of human-AI collaboration: A taxonomy of design knowledge for hybrid

- intelligence systems. In T. Bui (Ed.), *Proceedings of the Hawaii International Conference on System Sciences (HICSS)*. (3):15-19
- Flage, A. (2024). Taking games: a meta-analysis. *Journal of the Economic Science Association*, 1-24.
- Hamatt, D, Staeheli, C (2011). Respect and responsibility: Teaching citizenship in South African high schools International. *Journal of Educational Development*, 31 (3). p:14-27
- Hand, E. (2010). "Citizen science: People power". *Nature*. 466 (7307): 685–687.
- Kennedy, Graeme (1998). *An Introduction to CorPus Linguistics*. London: Longman. 13-85.
- Lehejcek, J., Adam, M., Tomasek, P., & Trojan, J. (2019). *Informacni system pro spravu fotopasti* (National database of photo trap records).
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215.
- Peterson, Andrew & Knowles, Catherine (2009). Active Citizenship: A Preliminary Study into Student Teacher Understandings. *Journal of Educational Research*, Vol 51 No 1, PP 39-59.
- Sandelowski, M., Docherty, S., & Emden, C. (1997). Focus on qualitative methods Qualitative metasynthesis: issues and techniques. *Research in nursing and health*, 20, 365-372
- Swanson, A., Kosmala, M., Lintott, C., & Packer, C. (2016). A generalized approach for producing, quantifying, and validating citizen science data from wildlife images. *Conservation Biology*, 30 (3), 520–531.
- Purta, J (2018). *Civic Education*, In: *International Encycloped Curriculum*. Dergamon press, v (9):117-132
- Trojan, J., Schade, S., Lemmens, R., & Frantál, B. (2019). Citizen science as a new approach in geography and beyond: Review and reflections. *Moravian Geographical Reports*, 27(4), 254–264.
- Yick, Alice G. (2013), "A Meta synthesis of Qualitative Findings on the Role of Spirituality and Religiosity Among Culturally Diverse Domestic Violence Survivors". *Health Policy & Services*, 37 out of 70.
- Zimmer L. (2006), "Qualitative meta-synthesis: a question of dialoguing with texts", *Journal of Advanced Nursing*. 53(3): 311-318.
- Sadeghi, S. S., Khotanlou, H., & Rasekh Mahand, M. (2021). Automatic Persian text emotion detection using cognitive linguistic and deep learning. *Journal of AI and Data Mining*, 9(2), 169-179.
- Urváľková, E. S., & Janoušková, S. (2019). Citizen science–bridging the gap between scientists and amateurs. *Chemistry Teacher International*, 1(2), 20180032

Yang, D., Wan, H. Y., Huang, T. K., & Liu, J. (2019). *The role of citizen science in conservation under the telecoupling framework*. Sustainability, 11(4), 1108.



