نقش نشانه M157 در ساختاردهی متون آغازعیلامی :

رویکردی محاسباتی بر پایه مدل‌های آماری و یادگیری ماشین

چکیده

نشانه M157 ( ) در متون آغازعیلامی از جمله پرکاربردترین نشانه‌هایی است که غالباً در جایگاه نشانه آغازین (Header) ظاهر می‌شود و به‌نظر می‌رسد نقشی ساختاری در سازمان‌دهی لوح‌ها ایفا می‌کند. پژوهش‌های اولیه، به‌ویژه توسط شایل، احتمال داده‌اند که این نشانه، معادل با نشانه بابلی (DUB) به‌معنای «لوح/گل‌نبشته» باشد. در این مقاله، با بهره‌گیری از داده‌های CDLI، فهرست نشانه‌های موجود و مجموعه‌ای از متون منتشرشده، نقش (M157) در چارچوب هر دو رویکرد سنتی و محاسباتی بررسی شده است. از منظر آماری، تحلیل فراوانی و هم‌رخدادی نشانه‌ها نشان داد که (M157) در پیوند با نشانه‌های شمارشی و واژگان مرتبط با کالاها و واحدهای اداری بیشترین حضور را دارد. تحلیل شبکه‌ای نیز جایگاه مرکزی (M157) را در خوشه‌های معنایی-ساختاری آشکار ساخت. افزون بر این، مدل‌های ماشین بُردار پشتیبانی برای طبقه‌بندی نوع لوح‌ها و مدل شبکه‌های عصبی بازگشتی برای پیش‌بینی نشانه‌های گمشده، نقش تعیین‌کننده (M157) را در شناسایی ساختار متون تأیید کردند. تفسیر این داده‌ها در پرتو شواهد باستان‌شناختی و تاریخی نشان می‌دهد که (M157) احتمالاً نه‌تنها برچسبی نهادی، بلکه نشانه‌ای **فرامتنی** برای تعریف و استانداردسازی اسناد اداری-اقتصادی بوده است؛ کارکردی که به‌ویژه در مقایسه با نشانه‌های مشابه اهمیت بیشتری پیدا می‌کند. در نتیجه، می‌توان فرضیهٔ شایل را بازاندیشی کرد و پیشنهاد داد که (M157) به‌معنای «گل‌نبشته/سند» یا نشانه‌ای مرتبط با یک خانوار یا نهاد خاص به‌کار رفته است. این پژوهش نشان می‌دهد که رویکردهای محاسباتی می‌توانند در بازسازی کارکردهای نشانه‌ای و ساختار متون باستانی ابزاری حیاتی باشند. یافته‌های این پژوهش از جمله داده‌های فراوانی، هم‌آیی و خوشه‌بندی نشان می‌دهند که (M157) یکی از نشانه‌های ساختاری در زبان‌شناسی آغازعیلامی، همراه با نقش سازمان‌دهنده‌اش در شبکه‌های نشانه‌ای بوده و بیانگر آن است که (M157) کارکردی فراتر از یک نشانه صرفاً صوری داشته و احتمالاً به‌عنوان شاخصی برای طبقه‌بندی و هویت‌بخشی اسناد به کار می‌رفته است. همچنین توصیه می‌شود در پژوهش‌های آینده، مقایسه تطبیقی (M157) با متون آغازمیخیِ سومری و تحلیل تغییرات زمانی جایگاه این نشانه در چارچوب تحول نظام‌های نوشتاری مورد توجه قرار گیرد.

واژگان کلیدی: آغازعیلامی، نشانه M157، نشانه آغازین (Header)، تحلیل آماری و شبکه‌ای، باستان‌شناسی دیجیتال.

ACCEPTED MANUSCRIPT

مقدمه

آغازعیلامی دوره‌ای است که در چند دههٔ اخیر بیش از پیش شناخته و مورد توجه قرار گرفته است. نخستین بار، ونسان شیل، پژوهشگر خطوط باستانی و عضو هیئت کاوش ژاک دمورگان، اصطلاح «آغازعیلامی » را در سال ۱۹۰۵ برای گل‌نبشته‌هایی با تصاویر اندیشه‌نگار به کار برد و این متون را به‌اشتباه متعلق به پیش‌ازتاریخ ایران و نظامی بومی با زبانی فرضی به نام «انشانی» دانست (Scheil 1905 : 60; Dahl 2009 : 31). پس از آن، با اتمام کاوش‌های دمورگان و دومکنم در شوش و کشف متونی مشابه در سیلک، پرسش‌های تازه‌ای دربارهٔ گستره و ماهیت این نظام نوشتاری پدید آمد. گیرشمن وجود گل‌نبشته‌های آغازعیلامی در سیلک را نتیجهٔ کشورگشایی با اهداف سیاسی و اقتصادی از جنوب‌غرب ایران دانست (گیرشمن، ۱۳۷۹: ۷۱). در ادامه، پژوهشگرانی همچون مک‌کان (1949) و کارلوفسکی (1978) مفهوم آغازعیلامی را به‌عنوان یک هویت فرهنگی یا واحد سیاسی به کار بردند. بررسی‌های بعدی نیز دیدگاه‌هایی نوین عرضه کردند: هالی پیتمن این دوره را بر اساس نشانه‌های گوناگونی همچون نظام نوشتاری، هنر مهرسازی و شواهد مدیریتی، پدیده‌ای مستقل از بین‌النهرین معرفی کرد (Pittman, 1997 : 135) و باربارا هلوینگ پس از کاوش‌های اریسمان و سیلک، احتمال وجود یک هویت فرهنگی یا ساختار سیاسی منسجم را مطرح ساخت (هلوینگ، ۱۳۸۳: ۱۵۱).

با وجود این تلاش‌ها، اطلاعات ما دربارهٔ ساختار دقیق این متون همچنان ناقص است. گزارش‌های اولیهٔ کاوش‌ها اغلب فاقد مستندنگاری علمی دقیق بوده و رمزگشایی کامل نظام نوشتاری آغازعیلامی هنوز به سرانجام نرسیده است (Englund, 1998; Damerow, 1999; Dahl, 2009; Born 2023). این نظام هنوز به‌طور کامل رمزگشایی نشده، اما ساختار آن به‌خوبی نشان می‌دهد که هدف اصلی آن سامان‌دهی امور اقتصادی و اداری بوده است. نوشتار آغازعیلامی برخلاف خط میخی متأخر فاقد ارزش‌های آوایی تثبیت‌شده است و بیشتر از نشانه‌های تصویری-مفهومی و عددی تشکیل شده که در قالبی مشخص میان نشانه‌های آغازین و بدنهٔ متنی توزیع می‌شوند. یکی از ویژگی‌های شاخص این متون، استفادهٔ گسترده از «سرتیترها (headers)» یا سرنشانه‌های آغازین است که در ابتدای لوح‌ها قرار می‌گیرند و به‌عنوان برچسب یا رده‌بندی‌کنندهٔ محتوای متنی عمل می‌کنند. نشانه‌های آغازین که در بخش آغازین متن را حضور دارند، کارکردی شبیه به «فراداده» داشته و به‌عنوان برچسب نهادی یا موضوعی برای دسته‌بندی اسناد عمل می‌کرده‌اند (Kelley 2024 : 86). افزون بر این، پژوهش‌ها بر نقش اجتماعی-نمادین برخی نشانه‌ها در بازتاب تقسیم کار، جنسیت و سن تأکید کرده‌اند (Kelley 2018, 2019)، و نیز بر ماهیت تصویری-آیکونوگرافیک نشانه‌های آغازین به‌عنوان بازمانده‌ای از سنت‌های مهری و تصویری پیش از نوشتار (Kelley 2024 : 89). بنابراین، هرچند نظام نوشتاری آغازعیلامی هنوز به‌طور کامل خوانده نشده، اما ساختار منسجم و قابل تحلیل آن نشان می‌دهد که می‌توان آن را نه یک زبان نوشتاری واجی، بلکه یک نظام اداری-مفهومی دانست که در نقطهٔ گذار از تصویر به متن جای گرفته است.

در میان نشانه‌های آغازین مختلف در متون آغازعیلامی، نشانهٔ (M157) جایگاهی ویژه دارد؛ زیرا در طیف وسیعی از متون تکرار می‌شود و نقشی بنیادین در سازمان‌دهی و طبقه‌بندی اطلاعات به نظر می‌رسد. پژوهش‌های پیشین دربارهٔ نوشتار آغازعیلامی عمدتاً بر گردآوری، فهرست‌برداری و تحلیل توصیفی نشانه‌ها متمرکز بوده‌اند. کارهای کلاسیکی همچون تلاش‌های (Damerow 1999; Dahl, 2002; Dahl et al. 2013; Englund 2001) با تمرکز بر تحلیل‌های تاریخی–مقایسه‌ای و بررسی نظام حسابداری نخستین انجام شده، و یا بر خوانش‌های موضوعی مبتنی بوده است (Dahl 2005, 2016; Dahl et al. 2018; Kelley 2019; Yousefi Zoshk, et al. 2022). با روی کار آمدن روش‌های جدید مبتنی بر مدل‌های بزرگ زبان طبیعی[1] که به‌صورت ویژه‌ای بر پیوند میان تصویر و متن تاکید می‌کند و بر نقش هوش مصنوعی در تحلیل نشانه‌ها مبتنی است، چشم‌اندازهای نوینی در زمینه تحلیل محاسباتی و آماری و رمزگشایی متون باستانی فراهم کرده است. این فعالیت‌ها با به‌کارگیری روش‌های برنامه‌نویسی و استفاده از تحلیل‌های محاسباتی در زبان آغازمیخی شروع شد (Zadworny & Gordin 2023) و تلاش‌های کاترین کلی (Kelley & Wood, 2018; Kelley et al. 2022; Kelley 2024) و لوگان بورن (Born et al. 2019; Born et al. 2021; Born

[1] LLM (Large Language Model)

ACCEPTED MANUSCRIPT

2023; Born et al. 2023) تاکنون همگی مبتنی بر معرفی روش‌های پردازش زبان طبیعی و تحلیل محاسباتی در حوزه متون آغازعیلامی بوده و مسیرهای تازه‌ای را در این حوزه گشوده است. با وجود این دستاوردها، هنوز یک خلأ آشکار در پژوهش‌ها وجود دارد: نبود یک تحلیل محاسباتی جامع و متمرکز بر روی نشانه‌های خاص و جایگاه آن در ساختار نشانه‌های آغازین. بیشتر مطالعات یا به سطح کلی بررسی نشانه‌ها بسنده کرده‌اند یا تمرکز آن‌ها بر کل سیستم نوشتاری بوده است، بی‌آنکه به نقش یک نشانهٔ کلیدی در سازمان‌دهی متون پرداخته شود.

با وجود مطالعات متعدد، تاکنون نقش ساختاری یک نشانهٔ خاص (M157) در جایگاه نشانه‌های آغازین بررسی نظام‌مند نشده است. به‌عنوان مثال، یکی از تفاسیر اولیه ونسان شایل، شناسایی (M157) به‌عنوان معادل نشانه بابلی (DUB) (لوح/گل‌نبشته) بود (Scheil, 1923 : 1). این فرضیه در پرتو تحلیل‌های آماری و شبکه‌ای امروز می‌تواند بازخوانی شود: آیا بسامد و جایگاه ساختاری (M157) نشان می‌دهد که این نشانه نه صرفاً برچسب نهادی، بلکه نشانه‌ای فرا-متنی به‌معنای "لوح" یا "سند" بوده است؟ پژوهش حاضر می‌کوشد این خلأ را با تلفیق روش‌های سنتی و رویکردهای محاسباتی پرکند. سوالات این پژوهش با هدف بررسی نقش نشانه (M157) در ساختاردهی متون آغازعیلامی طراحی‌شده و تلاش می‌کند از سه منظر به دنبال پاسخ به سوالات باشد. ابتدا، از منظر باستان‌شناختی، پرسش اصلی این است که نشانه (M157) در نظام نوشتاری آغازعیلامی چه نقشی ایفا می‌کند و آیا می‌توان آن را به‌عنوان یک «نشانه آغازین سازمانی» با کارکرد اداری ـ اقتصادی در ساختار لوح‌ها شناسایی و تحلیل کرد. از منظر آماری، این پرسش مطرح می‌شود که چه الگوهای هم‌رُخدادی و توزیعی میان (M157) و دیگر نشانه‌های پرتکرار (اعم از نشانه‌های عددی و غیرعددی) وجود دارد و این الگوها چه تصویری از سازمان‌دهی و طبقه‌بندی متون آغازعیلامی ارائه می‌دهند. از منظر یادگیری ماشین و روش‌های محاسباتی، پرسش کلیدی آن است که آیا می‌توان با بهره‌گیری از مدل‌های تحلیلی نوین مانند بُردار معنایی (Word2Vec)[۲]، بُردار عددی (Node2Vec)[۳]، مدل‌های مارکوف[۴] و شبکه‌های عصبی LSTM[۵]، نشانه‌های پس از (M157) را پیش‌بینی کرد و از این رهگذر به بازسازی بخش‌های ناقص متون آغازعیلامی دست یافت. در این پژوهش سعی بر این است که مطالعه‌ای نظام‌مند و مبتنی بر روش‌های محاسباتی دربارهٔ نقش ساختاری (M157) چه به عنوان نشانه آغازین و چه در متن و در ارتباط آن با سایر نشانه‌ها انجام شود. این مقاله می‌کوشد با بهره‌گیری از رویکردهای نوین تحلیل محاسباتی، از جمله "خوشه‌بندی"، "تحلیل شبکه‌ای" و "مدل‌های زنجیره‌ای" که همگی مبتنی بر استفاده از رویکردهای برنامه‌نویسی مبتنی بر "مدل‌های بزرگ زبان طبیعی" هستند، به دنبال کشف الگوهای همبستگی بین نشانه‌های آغازین (M157) و نشانه‌های همراه را شناسایی کند و نقش این نشانه را در سازمان‌دهی اطلاعات متون آغازعیلامی روشن سازد. نتایج این مطالعه نه‌تنها برای شناخت نظام اداری و نوشتاری جوامع آغازعیلامی اهمیت دارد، بلکه می‌تواند درک ما از فرآیند جهانی پیدایش خط و توسعهٔ نظام‌های مدیریتی نخستین را نیز عمیق‌تر کند.

پیشینه پژوهش

پژوهش‌های انجام‌شده در حوزه متون آغازعیلامی نشان می‌دهند که لوح‌های آغازعیلامی با «سرتیترآغازین/نشانه آغازین» شروع می‌شوند؛ خطی که معمولاً یک یا چند نشانه رده‌بندی‌کننده نهادی/موضوعی را در ابتدای متن می‌آورد و کارکرد فهرست‌واره و سازمان‌دهنده دارد (Dahl, 2013 : 2; Forouzan et al. 2022 : 47). مطالعات کلاسیک نشان داده‌اند که با تکیه بر مقایسه صوری و کارکردیِ نشانه‌ها میان آغازعیلامی و

---

[۲] یک مدل یادگیری ماشین که کلمات یا نشانه‌ها را به بردارهای عددی تبدیل می‌کند، به‌طوری‌که فاصله و زاویه بین بُردارها معنای شباهت یا ارتباط بین آن‌ها را نشان می‌دهد. با بررسی زمینه هر نشانه، یاد می‌گیرد که نشانه‌های مشابه معمولاً در محیط‌های مشابه ظاهر می‌شوند. دو معماری اصلی‌اش (Skip-gram) و (CBOW) هستند. این مدل می‌تواند الگوی هم‌رخدادی نشانه‌ها در لوح‌ها را یاد بگیرد، مثلاً بعد از (M157) کدام نشانه‌ها از نظر معنایی یا ساختاری محتمل‌تر هستند.

[۳] روشی برای تبدیل گره‌های یک شبکه به بُردارهای عددی، مشابه Word2Vec، اما به جای متن از شبکه‌ها یا گراف‌ها استفاده می‌کند. در این مدل گره‌ها را در شبکه (مثلاً هر نشانه به‌عنوان یک گره و ارتباطات توالی یا هم‌رخدادی به‌عنوان یال) با فرآیندهای تصادفی پیمایش می‌کند و از روی مسیرهای پیمایش‌شده، بردار تولید می‌کند. این مدل می‌تواند روابط ساختاری و موقعیت نشانه (M157) در «شبکه نشانه‌ها» را مدل کند تا بفهمد چه نشانه‌هایی از نظر ارتباط شبکه‌ای بعدش محتمل هستند.

[۴] یک مدل آماری که می‌گوید احتمال رخداد یک وضعیت فقط به وضعیت فعلی بستگی دارد، نه کل تاریخچه. این مدل با جدول احتمالات (ماتریس گذار) مشخص می‌کند اگر الان M157 داریم، احتمالاً چه نشانه‌هایی بعدش می‌آید، صرف‌نظر از قبل‌ترها. این مدل می‌تواند سریع و ساده برای پیش‌بینی بعد از (M157) بر اساس فرکانس واقعی توالی‌ها در داده‌های لوح تصمیم بگیرد.

[۵] نوعی شبکه عصبی بازگشتی (RNN) که برای به یاد آوردن وابستگی‌های بلندمدت طراحی شده است. برخلاف مدل مارکوف که فقط وضعیت فعلی را می‌بیند، شبکه عصبی بازگشتی (LSTM) می‌تواند الگوهای چندین نشانه قبل را نیز در نظر بگیرد، زیرا ساختاری دارد که «حافظه طولانی» را حفظ می‌کند. این شیوه می‌تواند وابستگی‌های پیچیده و بلندمدت بین نشانه‌ها را یاد بگیرد و پیش‌بینی دقیقی از نشانه بعد از (M157) بدهد، اگر الگو در فاصله چندین نشانه قبل شروع شده باشد.

ACCEPTED MANUSCRIPT

آغازمیخی (مانند هم‌ارزی‌های «مرد/زن/کودک») و با اتکا به ساختار لوح‌ها و نظام عددی، می‌توان بازسازیِ حوزه‌های معنایی قدمی برداشت. در این سنت، نشانه‌هایی چون «مرد/زن (M388/M72)» و جایگاه‌شان در خطوط آغازین (گاهی به‌مثابه «تیتر» نوعِ محتوایِ لوح) بررسی شده‌اند تا کارکردهای رده‌بندی/طبقه‌نما روشن شود (Dahl, 2018 : 391). این مطالعات همچنین بر تعامل دیرینه مُهر-نوشتار (seals ↔ script) تأکید می‌کنند و منشأ بخشی از نشانه‌ها را در سنت‌های تصویریِ مُهرها جست‌وجو می‌نمایند (Kelley et al. 2024 : 87). اما در حوزه زبان‌شناسی تاریخی، رویکرد اصلی بر تحلیل تطبیقی نشانه‌ها و بازسازی ویژگی‌های صرفی-نحوی و معنایی متمرکز بوده است. پژوهشگران همواره بر پیوند ساختاری میان آغازعیلامی و نظام‌های خطی هم‌عصر مانند آغازمیخی دوره سومریان تأکید داشته‌اند، اما فقدان شواهد زبانی کافی مانع از تطبیق کامل ارزش‌های آوایی شده است (Englund 2004 : 23–24; Born 2023 : 92). از این رو، پرسش‌هایی درباره ماهیت زبان، کارکرد دقیق نشانه آغازین، و روابط میان نشانه‌های نهادی و مفاهیم عددی همچنان محل پرسش است.

۲-۱. تاریخچه کشف و تدوین متون و نشانه‌ها

فعالیت‌های پژوهشی در حوزه آغازعیلامی با کاوش‌های شوش در ابتدای سده بیستم به سرپرستی ونسان شایل (Vincent Scheil) آغاز شد. نخستین انتشار متون در سال ۱۹۰۰ صورت گرفت و طی دهه‌های بعد پژوهشگرانی همچون دمورگان، مک‌کنم، مریگی، انگلوند، دمرو، و دال و دِسِه مجموعه‌های تازه‌ای منتشر کردند (افشاری، ۱۳۹۸). هر انتشار جدید غالباً با یک فهرست نشانه همراه بود؛ فهرست‌هایی که در ابتدا نزدیک به هزار نشانه معرفی می‌کردند. در ادامه و با پیشرفت روش‌های تحلیل گرافوتاکتیکی و رایانه‌ای در قرن بیست‌ویکم، این تعداد به حدود ۱۲۰۰ نشانه متمایز بازبینی شد (Born 2023: 95). در این نظام نوشتاری، نشانه‌ها به‌طور کلی به دو دسته تقسیم می‌شوند: نشانه‌های عددی با پیشوند N و نشانه‌های تصویری با پیشوند M. نشانه‌های عددی دارای ساختار شفاف‌تر بوده و در قالب سیستم‌های ده‌دهی، شصت‌پایه و دیگر گونه‌ها به‌کار رفته‌اند (Dahl, 2018 : 388). نشانه‌های تصویری اما فاقد معادل مستقیم واژگانی‌اند و بیشتر کارکرد ایده‌گرافیک یا احتمالاً هجایی داشته‌اند. تلاش‌های تطبیقی برای همانندسازی این نشانه‌ها با خط میخی بین‌النهرینی (به‌ویژه در حوزه ارزش‌های هجایی) هرگز به‌طور کامل موفق نبوده و بنابراین بخش عمده معناشناسی نشانه‌ها همچنان مبهم باقی مانده است (Englund 2004; Damerow & Englund 1989).

فهرست نشانه‌ها و سنت دسته‌بندی

یکی از دستاوردهای کلیدی سنت اپیگرافی، گردآوری فهرست‌نامه‌های نشانه‌ها بوده است. در این میان، کار تحقیقی "پیرو مریگی[6]" (1974) نقطه عطفی محسوب می‌شود و بعدها یاکوب دال (2002) با ارائه فهرست جامع‌تر، شیوه شماره‌گذاری استاندارد با پیشوندهای M و N را تثبیت کرد. این فهرست‌ها نه تنها امکان شناسایی دقیق نشانه‌هایی چون (M157) به همراه واریانت‌های آن را فراهم ساخته‌اند، بلکه مبنایی برای تحلیل‌های دیجیتال و محاسباتی در پژوهش‌های جدید ایجاد کرده‌اند (Kelley et al. 2022).

ساختار متون: نشانه آغازین، مدخل‌ها و جمع‌بندی‌ها

مطالعات آغازعیلامی (Englund 2004; Hawkins 2015) نشان داده‌اند (جدول یک) که متون آغازعیلامی معمولاً با یک نشانه در تیتر متن آغاز می‌شوند که احتمالاً نمایانگر فرد، خانوار یا نهاد صادرکننده سند بوده است. پس از این بخش، ردیفی از نشانه‌های متنی و سپس بخش‌های عددی (Numeric Notation) قرار می‌گرفت که برای ثبت مقدار کالا، دام یا غله استفاده می‌شد. در انتهای متن نیز غالباً جمع‌بندی و گاهی البته به همراه یک مهر دیده می‌شود. این الگوی نشانه آغازین-بدنه، بعدها پایه بسیاری از تحلیل‌های ساختاری و محاسباتی قرار گرفت (Born 2023; Kelley 2018, 2024).

6 - Piero Meriggi

ACCEPTED MANUSCRIPT

منبع	سال انتشار	روش مورد استفاده	نتایج کلیدی
Scheil (MDP 06, 17, 26, 31)	1900-1949	بازخوانی دست‌نویس بر پایه خط میخی	انتشار اولیه متون و معرفی نشانه‌های آغازعیلامی
Meriggi	1974	فهرست‌نگاری و دسته‌بندی	نخستین فهرست جامع نشانه‌ها از جمله M157
Brice	1962	تحلیل ساختار الفبایی	بررسی کارکرد نشانه‌ها در متون حسابداری
Damerow & Englund	1989	تحلیل عددی و ایدئولوژیک	تاکید بر کارکرد نهادی نشانه‌های آغازین و ساختار کلان متون
Englund	2004	تحلیل ساختاری-زبان‌شناختی	توجه به ارزش هجایی و نقش نشانه‌های آغازین در سازمان‌دهی متن

جدول ۱: جدول زمانی سیرِ تحول مطالعات سنتی در حوزه زبان آغازعیلامی

رویکردهای دیجیتال در خوانش متون باستانی: از توصیف تا تحلیل چندلایه

در دهه اخیر، پژوهش در حوزه متون باستانی از خوانش‌های سنتی و توصیفی به سمت رویکردهای دیجیتال و داده‌محور تحول یافته است (Kelley et al., 2022; 2024). علوم انسانی دیجیتال با بهره‌گیری از ابزارهایی چون مدل‌های آماری و موضوعی مانند تحلیل نهفته (LDA)[7] برای استخراج خوشه‌های معنایی و سنجش هم‌آیندی نشانه‌های آغازین با رده‌های محتوایی (Born 2023; Born et al. 2021; Born et al. 2019; Born et al. Sommerschield, et al. 2023)، بُردارهای معنایی برای تحلیل هم‌نشینی و کشف خانواده‌های نشانه‌ای (2022; Sahala 2021)، و "مدل‌های ترتیبی و عمیق (HMM)" و "ترنسفورمر"ها برای شناسایی خودکار نشانه‌های آغازین و بازسازی نظم نوشتاری (Vaswani et al. 2017; Sommerschield, et al. 2023 : 708)، زمینه تحلیلی تازه‌ای فراهم کرده‌اند. افزون بر این، مدل‌های ترکیبی تصویر-متن امکان بازبینی فهرست نشانه‌ها و کاهش خطاهای ترانویسی را می‌دهند، و چارچوب‌های شبکه‌ای و رده‌بندی مانند (i-Classifier) نیز در طبقه‌بندی نشانه‌ها و روابط نهادی به‌کار گرفته شده‌اند (Selz & Zhang 2024 : 69). این مجموعه چندلایه از روش‌ها، چشم‌اندازی نو برای تحلیل ساختار متون آغازعیلامی و نشانه‌های محوری مانند M157 گشوده است.

مواد و روش ها

نشانه M157 (𒀸)

نشانهٔ (M157) یکی از پرتکرارترین نشانه‌های آغازعیلامی است که هم در تصویرنگاری مُهرها و هم در گل‌نبشته‌ها نقش ساختاری مهمی ایفا می‌کند. نشانه (M157) در متون آغازعیلامی به‌عنوان نماد «خانهٔ/واحد خانگی (Household)» معرفی شده که هم در نقش نماد نهادی روی مُهرها دیده می‌شود و هم در نظام نوشتار به‌عنوان نشانه‌ای با کارکرد «برچسب‌گذار نهادی» مطرح است (Kelley, 2024 : 84). این پیوند مهر-نوشتار از منظر ریشه‌شناسی گرافیکیِ نشانه‌های آغازعیلامی اهمیت دارد و به‌طور ضمنی توضیح می‌دهد چرا (M157) می‌تواند در جایگاه «نشانه سرفصل/نشانه آغازین» عمل کند.

[7] تحلیل نهفته دیریکله (Latent Dirichlet Allocation- LDA) یک مدل آماری برای کشف الگوهای پنهان در مجموعه‌های بزرگ داده‌متنی است. این مدل بر این فرض استوار است که هر سند ترکیبی از چند موضوع و هر موضوع ترکیبی از واژگان است. توزیع هر موضوع در اسناد و توزیع واژگان در هر موضوع، از یک توزیع احتمال خاص به نام «دیریکله» پیروی می‌کند که امکان مدل‌سازی نسبت‌های احتمالی را فراهم می‌سازد. در فرایند LDA، هر سند به‌صورت آماری به درصدهایی از چند موضوع تقسیم می‌شود و هر موضوع نیز مجموعه‌ای از واژگان با احتمال وقوع متفاوت در خود دارد. هدف مدل آن است که، بدون آگاهی پیشین از تعداد و ماهیت موضوعات، این ساختار پنهان را با استفاده از داده‌های موجود شناسایی کند. در تحلیل متون و اسناد آغازعیلامی، می‌توان نشانه‌ها را به‌منزله «واژگان» و هر لوح را به‌مثابه «یک سند» در نظر گرفت. پس LDA می‌تواند بر پایه بسامد و هم‌آیی نشانه‌ها، خوشه‌های معنایی یا کارکردی را آشکار سازد؛ برای مثال، تمایز میان دسته‌های متون اقتصادی، اداری یا دامداری را مشخص کرده و الگوهای کارکردی مشترک یا متمایز را نمایان کند.

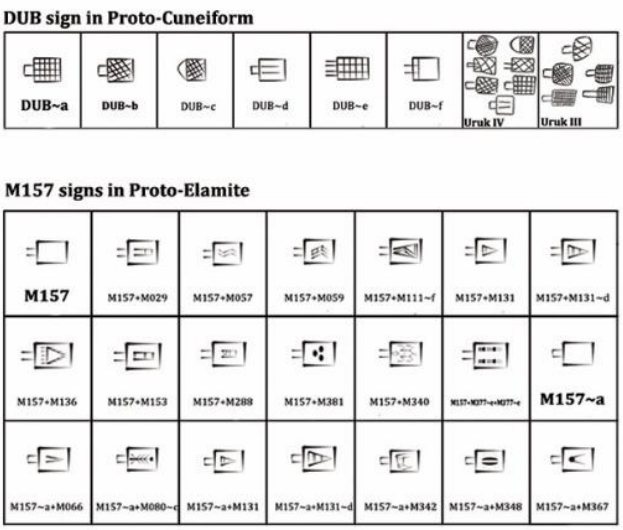
ACCEPTED MANUSCRIPT

در متون آغازعیلامی، (M157) معمولاً در ابتدای گل‌نبشته‌ها و پیش از هرگونه نشانهٔ عددی ظاهر می‌شود و در تحلیل آماری یاکوب دال با ۲۴۷ بار تکرار در صدر فهرست بیشترین نشانه‌های تکرارشده قرار دارد (Dahl, 2002 : 3, Table 2). تفسیرهای سنتی نیز متنوع بوده‌اند؛ مک‌کنم در (MDP 31) آن را معادل انبار غله دانسته است و شایل در (MDP 6) آن را معادل نشانه آغازمیخی DUB ( ) به معنای «لوح/گلنبشته» خوانده و در یادداشت‌های خود در (MDP 17)، نقشی مهم و سازمان‌دهنده برای آن به عنوان نشانه نخست لوح مورد تاکید قرار داده است (Scheil, 1923 : 1). در متون آغازمیخی (تصویر یک)، نشانه DUB به معنای گلنبشته قلمداد شده و شکل آن با یک گلنبشته شباهت کم‌نظیری دارد (Diaco, 2019 : 45). سپس، مریگی بر اساس شباهت‌های گرافیکی با علامت میخی DUB، به آن معنای «حساب» را نسبت داد (Meriggi 1971 : 153). دال نیز بر همین عقیده است و دو نشانه بعدی را نشان از شخص یا خانواری می‌داند که این گلنبشته متعلق به آن است (Dahl, 2002 : 3, p. 13). او معتقد است که این نشانه به همراه نشانه دیگر (M346) که در لیست نشانه‌های دارای بیشترین بسامد حضور دارند، دارای معنایی غیر از محصولات غلات، ظروف یا اشخاص می‌باشد (Dahl, 2002 : 3). دال با اصلاح نشانه‌های لیست مریگی، این نشانه را دارای چندین نوع ساده و پیچیده می‌داند (Dahl, 2002) (تصویر ۱).

آنچه این نقش نشانه‌ای را برجسته می‌کند، حضور آن در مُهرهای آغازعیلامی است. پالادره در بررسی اثرمُهرهای موجود در موزه لوور، نقش (M157) را در ۴۸ اثرمهر یافته‌است (Paladre, 2022). پالادره با بررسی علامت (M157) در نقوش اثرمُهرها که به‌صورت یک مربع با دهانه‌ای باز یا بسته است (Paladre, 2022 : Vol. II, 283; Pittman, 1994 : 148)، متوجه شده که آن نقش، نمادی مهم و مکرر در الواح آغازعیلامی در شوش محسوب می‌شود. این علامت اغلب به عنوان سربرگ الواح مُهروموم‌شده به کار رفته و جالب است که در مواردی که یک مُهر واحد بر روی چندین لوح به کار رفته، سربرگ تمامی آنها بدون استثنا همین علامت یا یکی از گونه‌های آن بوده است (Paladre, 2022 : Vol. I, 314). این نقش روی مُهرهایی با تصاویر متنوع از جمله صحنه‌های اقتصادی، صحنه‌های حیوانی مانند گربه‌سانی در قایق، و به ویژه یک صحنه تعادل قدرت منحصربه‌فرد با حضور پیکر انسانی و تسلیم یک‌طرفه یافت می‌شود (تصویر ۲) (Paladre, 2022 : Vol. III, 38). اگرچه در گذشته این علامت به معنای ساده «لوح» یا «حساب» تفسیر شده بود، اما تکرار آن، محدود بودن مکان آن به شوش، و حضورش در این صحنه‌های خاص، نشان می‌دهد که (M157) نمادی با اهمیت اجتماعی-سیاسی عمیق بوده که احتمالاً مربوط به یک نهاد اداری یا آرشیو خاص در این منطقه است، نمادگرایی که ریشه‌های آن به دوره‌های قدیمی‌تر باز می‌گردد (شکل ۱).

کاترین کلی هم با بررسی و مقایسه نشانه (M157) در مُهر (MDP 16, N334)، که در سنت مُهرسازی آغازعیلامی در محوطه شوش یافت شده، آن را نمونه‌ای برجسته از مهاجرت و تحول تصویر در میان فرهنگ‌های هم‌جوار جنوب عراق و جنوب‌غرب ایران می‌داند (Kelley, 2024 : 86). این نقش مُهر، با وجود تفاوت‌های بنیادین در نظام نوشتاری با مُهرهای اوروکی، نشان‌دهندهٔ اشتراکات تصویری و موضوعی میان این دو حوزه فرهنگی است. یگانگیِ این نشانه آن‌جاست که در سنت مُهرنگاریِ هزاره چهارم-سوم پیش از میلاد، نشانه (M157) در صحنه‌های آیینی/اداری به‌مثابه نماد یک نهاد ظاهر می‌شود و حضور در کنار موتیف‌های حمل‌ونقل/کشتی به‌منزله «علامتِ مؤسسه» به‌کار رفته است (Ibid, 85). در این مُهر (تصویر ۳)، تصویر یک حیوان که به‌نظر می‌رسد گاو نر باشد، درون یک قایق دیده می‌شود. در این نقش، قایق با تزئینات جانبی، تورهایی که در آب آویزان‌اند، و وجود یک جعبه مربع‌شکل در قسمت عقبی قایق، فضایی نمادین و آیینی را تداعی می‌کند. همچنین وجود دو پارو در دو انتهای قایق و ساقه‌های بلند و گره‌خوردهٔ نی، و بدون حضور انسان و البته همراه با نقش یک ماهی، نشان‌دهندهٔ نوعی پیوند میان قدرت تولیدی (گاو نر) و ساختار اجتماعی یا اقتصادی (خانوار) است (Kelley, 2024 : 86).

ACCEPTED MANUSCRIPT



تصویر ۱: تصویر شماره ۱ بیانگر مقایشه شکلی نشانه DUB در متون آغازمیخی و نشانه M157 در متون آغازعیلامی به‌همراه بسامدهای آنها است. بالا: نشانه DUB در متون آغازمیخی با شماره *ZATU086* (Green & Nissen, 1987 : 188)؛ پایین: نشانه M157 در متون آغازعیلامی و انواع آن (Dahl, 2002).

تصویر ۲: نشان‌دهنده استفاده از این نقش در اثرمُهرهای آغازعیلامی در محوطه شوش است: A. نقش مربع هاشوردار به همراه دو جغد از لایه 17B آکروپل شوش (Paladre, 2022 : Vol. II, 138)؛ B. نقش مربع در اثرمُهر آغازعیلامی در شوش (Legrain, 1921 : Pl. XXIII)؛C. نقش مربع در اثرمُهر آغازعیلامی در شوش (Amiet 1972 : pl. 22 et 101, fig. 933)؛ D. نقش مربع در اثرمُهر دوره شوش۲ در شوش (Amiet 1972 : pl. 22 et 101, fig. 646)؛ E. نقش مربع در اثرمُهر آغازعیلامی در شوش (Delaporte 1920 : pl. 43, fig. 10, S 344)؛ F. نقش مربع در اثرمُهر دوره شوش ۲ در شوش (Amiet 1972 : pl. 22 et 101, fig. 681)؛ G. نقش مربع در اثرمُهر دوره شوش ۴ در شوش (Amiet 1972 : pl. 22 et 101, fig. 654)؛ H. نقش مربع در اثرمُهر آغازعیلامی در شوش (Amiet 1972 : pl. 22 et 101, fig. 675)؛ I. نقش مربع در اثرمُهر آغازعیلامی در شوش (Amiet 1972 : pl. 22 et 101, fig. 1013)؛ J. نقش مربع در اثرمُهر آغازعیلامی در شوش (Amiet 1972 : pl. 22 et 101, fig. 674)؛ K. مقایسه موتیف‌های مرتبط با نشانه M157 در مُهرها و گلنبشته‌های آغازعیلامی شوش (Pittman 1994 : Figure 29)؛ L. تصویری از اثرمُهر B. بر روی گلنبشته آغازعیلامی در شوش (Álvarez-Mon 2020 : pl. 29, fig. b)؛ M. تصویری از اثرمُهر I. بر روی گلنبشته آغازعیلامی در شوش (Álvarez-Mon, 2020 : pl. 29, fig. e).

ACCEPTED MANUSCRIPT

شکل ۱: نقشه پراکندگی حضور نشانه M157 تنها در محوطه شوش در مقایسه با حضور دیگر نشانه M136 در جغرافیای فلات ایران (Paladre, 2022 : Vol. II, 281) که بر یگانگی استفاده از این نشانه در شوش تاکید می‌کند.

تصویر ۳: اثر مُهر با نقش مربع به‌همراه گل‌نبشته آن با نشانه آغازین M157 (Paladre, 2022 : 47) که همبستگی استفاده همزمان از نشانه M157 به عنوان یک نهاد/نماد در اثرمُهر و متن در گلنبشته شماره P008022 را نشان می‌دهد.

داده‌ها

معرفی مجموعه متون: تعداد لوح‌ها، نشانه‌ها، توالی‌ها

با ظهور پایگاه‌ها و ویرایش‌های دیجیتالی همانند (Cuneiform Digital Library Initiative = CDLI)، تصاویر با کیفیت بالا و ترانویسی‌های استاندارد شده از گل‌نبشته‌های آغازعیلامی در دسترس پژوهشگران سراسر جهان قرار گرفت. انتشار نسخهٔ تصحیح‌شده نشانه‌ها و ساخت فایل‌های مختلف قابل استفاده پردازش زبان‌های باستانی، تحلیل دیجیتال نشانه‌های آغازعیلامی با دقت بسیار بالاتر ممکن گردید. پیکرهٔ مورداستفاده در این پژوهش شامل ۳٦٤ گلنبشته آغازعیلامی با نشانه (M157) است که از مجموعه لوح‌های اقتصادی و اداری به‌دست‌آمده از محوطه باستانی شوش انتخاب شده‌اند. برای ساخت این مجموعهِ داده‌ها، ابتدا تمامی متون حاوی نشانه (M157) از پایگاه (CDLI) استخراج و به صورت یک فایل متنی با پسوند (.txt) ذخیره گردید. سپس برای اطمینان از محتوای صحیح، تمامی نشانه‌های متون استخراج‌شده، جداگانه از طریق مقایسه نشانه‌های بازنویسی شده براساس فهرست یاکوب دال با تصاویر و طرح‌های موجود در همان پایگاه داده (CDLI) مورد بررسی و مداقّه قرار گرفت و موارد خطاء احتمالی از لیست حذف گردید. در نهایت در فرایند پیش‌پردازش تحلیل‌های محاسباتی مدّنظر در این پژوهش، متون

با استفاده از کدهای برنامه‌نویسی به زبان پایتون مورد پردازش قرار گرفت و در نهایت تعدادِ فراوانی، داده‌های مکانی، و الگوهای هم‌رخدادی نشانه‌ها نیز برای تحلیل‌های بعدی جداگانه محاسبه و ذخیره گردید.[۸]

توصیف چگونگی استخراج داده‌ها از پایگاه‌ها یا اسناد منتشرشده

برای تحلیل‌های محاسباتی، داده‌ها به‌صورت دیجیتالی بازنویسی و استانداردسازی شده و با ابزارهای متن‌کاوی پردازش شده‌اند. در این مرحله، پاک‌سازی داده‌ها (حذف نشانه‌های ناقص یا مبهم)، نرمال‌سازی (هماهنگ‌سازی کدهای نشانه‌ها) و تفکیک نشانه‌های آغازین از متن صورت گرفته است. همچنین فرکانس، موقعیت و الگوهای هم‌رخدادی نشانه‌ها در کنار نشانه آغازین (M157) به‌صورت جداگانه محاسبه و ذخیره شده‌اند. این پیکره نه تنها بستر آماری لازم برای تحلیل‌های بعدی (از جمله خوشه‌بندی، تحلیل شبکه‌ای و مدل‌های پیش‌بینی) را فراهم می‌کند، بلکه به‌عنوان یکی از کامل‌ترین مجموعه‌های دیجیتال‌شدهٔ مربوط به نشانه (M157)، قابلیت مقایسه با متون غیر(M157) را نیز خواهد داشت.

روش‌های تحلیلی

تحلیل آماری: فراوانی، هم‌رخدادی، آزمون‌های χ^2 ، Cramé r's V

کاربرد آمار در تحلیل متون باستانی، به‌ویژه در متون خط میخی و لوح‌های اداری، نقش کلیدی دارد. در پروژه بایگانی دیجیتالی (اِبلا)[۹] تأکید شده است که تنوع قرائت‌ها و املای واژگان، و نیز وجود نشانه‌هایی با ارزش‌های چندگانه، ضرورت بهره‌گیری از ابزارهای آماری را افزایش می‌دهد (Maiocchi & Milano, 2023 : 319). این داده‌ها می‌توانند با آزمون‌های "هم‌رخدادی[۱۰]" و "وابستگی" (χ^2 , Cramé r's V) تحلیل شوند تا ارتباط میان نشانه‌ها و واژه‌ها آشکار شود. در مطالعه متون اکدی نیز از شاخص (PMI) برای استخراج هم‌رخدادی‌های معنی‌دار پیرامون لِمات هدف و کاهش نویز صرفی استفاده شد (Gordin, et al. 2024 : 136). در پژوهش‌های متون آغازعیلامی، از شاخص (Cramé r's V) (با اصلاح سوگیری برگسما) برای سنجش شدت ارتباط میان «نخستین نشانه لوح» و «موضوع غالب لوح» بهره گرفته شد و نتایج نشان داد که بیشترین اطلاعات موضوعی در جایگاه H1 قرار دارد (Born et al. 2019 : 128–130).

تحلیل شبکه‌ای: شناسایی جوامع، مرکزیت نشانه‌ها، تراکم و ماژولاریتی

نشانه‌ها و واژگان متون باستانی را می‌توان به‌صورت شبکه‌ای تصور کرد. در متون اِبلا، نشانه‌ها اغلب در قالب «جعبه‌های معنایی (semantic wholes)» مثل (عدد + اسم یا فعل + نام شخص) ظاهر می‌شوند و همین ساختار بستر تحلیل شبکه‌ای را فراهم می‌کند (Maiocchi & Milano, 2023 : 319). در پژوهش‌های متون اکدی، شبکه‌های هم‌رخدادی بر اساس وزن‌های (PMI) و شبکه‌های مشابهت معنایی بر اساس کسینوس (FastText) ساخته شدند و نشان دادند که خوشه‌ها بیشتر بر حسب ژانر شکل می‌گیرند تا قطبیت احساسی (Yavasan & Gordin, 2025 : 198). در پژوهش‌های متون آغازعیلامی نیز خوشه‌بندی نشانه‌ها مجموعه‌هایی پایدار، مانند گروه‌های مرتبط با دام‌های کوچک را آشکار ساخت (Born & Kelley, 2021 : 2). برای سنجش کیفیت این خوشه‌ها، از شاخص ماژولاریتی بر پایه الگوریتم (Louvain) استفاده شد (Born & Kelley, 2021 : 4).

مدل‌سازی موضوعی (LDA): کشف حوزه‌های معنایی مرتبط با نشانه (M157)

مدل سازی موضوعی (LDA) ابزار مناسبی برای کشف حوزه‌های معنایی در متون باستانی هستند. در پروژه (Potnia)، هدف اصلی تبدیل متون خطی همچون (Linear B) از حالت ترانویسی به کدهای منحصر به فرد بود تا بتوان آن‌ها را وارد مدل‌های زبانی و روش‌هایی مانند آن کرد؛

---

[۸] تمامی مجموعه داده‌های مرتبط با این پژوهش به صورت آنلاین در دسترس می‌باشد. تمامی گلنبشته‌های آغازعیلامی حاوی نشانه M157 از طریق پایگاه داده CDLI (از طریق لینک: https://tinyurl.com/yck9aumt) قابل دسترسی و دریافت است. همچنین، تمامی تحلیل‌های محاسباتی و آماری در این پژوهش، همراه با مجموعه داده‌های پردازش‌شده از طریق پایگاه گیت‌هاب یکی از نویسندگان، به آدرس [https://github.com/SaeedBaghizadeh/M157.git] قابل بررسی و دریافت است. این مسئله به شفافیت کامل و بازتولیدسازی تحقیقات بعدی کمک شایانی می‌کند.

[۹] Ebbla

[۱۰] Co-occurrence

ACCEPTED MANUSCRIPT

چراکه کار با متن اصلی (نه صرفاً ترانویسی) دقت تحلیل‌های معنایی را افزایش می‌دهد (Tour et al., 2025 : 2-3). در خوانش متون آغازعیلامی، مدلی با ۱۰ موضوع ساخته شد و برای تحلیل فاصله موضوع‌ها به‌کار رفت؛ نتایج نشان داد برخی موضوع‌ها به‌شدت با نشانه‌هایی مانند (M056~f و M297~b) پیوند دارند (Born et al. 2019 : 125). موضوع مهم آنست که این مدل حتی در متون ناقص و ناخوانا نیز به‌دلیل تکیه بر هم‌رخدادی کارآمد است (Born et al. 2019 : 130).

بُردارهای معنایی (Word2Vec, Node2Vec): استخراج مشابهت‌های معنایی

پس از یونیکدسازی متون باستانی، امکان تولید جاسازی‌های (embedding) معنایی برای استخراج شباهت‌های میان نشانه‌ها و کلمات فراهم می‌شود (Tour et al., 2025 : 2-3). در خوانش دیجیتالی متون آغازعیلامی نیز پیشنهاد شده است که از بُردار معنایی برای بُردارسازی گره‌های شبکه نشانه‌ها استفاده شود تا مشابهت ساختاری میان نشانه‌ها (مثلاً جایگاه مشترک نشانه M157 با نشانه‌های همراه) استخراج گردد (Born & Kelley, 2021 : 4).

مدل‌های زنجیره‌ای و یادگیری ماشین: استفاده از مدل مارکوف و LSTM برای پیش‌بینی نشانه‌ها و SVM برای طبقه‌بندی متون

مدل‌های زنجیره‌ای و یادگیری ماشین نقش مهمی در بازسازی و پیش‌بینی ساختار متون باستانی دارند. پژوهش‌های اخیر نشان داده‌اند که شبکه‌های عصبی بازگشتی و مدل‌های مارکوف می‌توانند برای بازسازی متون ناقص کارآمد باشند (Tour et al., 2025 : 2)؛ برای نمونه، فتایا و دیگران از شبکه‌های بازگشتی برای بازسازی متون بابلی استفاده کرده‌اند (Fetaya, et al. 2020 : 22743). در بازسازی متون ناقص آغازعیلامی، از مدل یادگیری عمیق و الگوریتم ویتربی برای شناسایی ساختار «نشانه آغازین» استفاده شد که منجر به کشف انواع جدیدی از نشانه‌های آغازین چندنشانه‌ای گردید (Born et al. 2019 : 130). همچنین یک مدل ترنسفورمر نشان داد که «نشانه دوم» قوی‌ترین شاخص برای وجود نشانه آغازین است و فرضیه سنتی تمرکز بر «نشانه نخست» را به چالش کشید (Born et al. 2022 : 113). افزون بر این، دسته‌بندهایی مانند ماشین بُردار پشتیبان می‌توانند بر اساس ویژگی‌های بُرداری یا گرافی برای طبقه‌بندی متون به ژانر/موضوع مناسب به‌کار روند (Saeid et al. 2019; Sommerschield, et al. 2023; Zadworny, & Gordin, 2025).

تحول در خوانش معنایی متون زبان‌های باستانی بازتابی از تغییر روش‌های پژوهش است: از توصیف‌های اپیگرافیک و پابلوگرافیک در اوایل قرن بیستم تا تحلیل‌های آماری و دیجیتالی اخیر. در مطالعات دیجیتال پیش‌رو، نشانه (M157) به‌عنوان یک گِره مرکزی (Hub) در شبکهٔ نشانه‌ها شناسایی شده و معمولاً در مدل‌های مختلف با نشانه‌های ساختاری، اقتصادی و نهادی هم‌خوشه می‌شود (Dahl 2002; Kelley et al. 2024). پژوهش‌های اخیر نشان می‌دهد که این نشانه نقشی ترکیبی و قالب‌بندی دارد و به‌عنوان برچسب نهادی یا نشانه آغازین اسناد اداری عمل می‌کند. بدین‌ترتیب، نشانه‌ای که در ابتدا صرفاً به‌عنوان بازنمایی یک شیء یا کارکرد سادهٔ اپیگرافیک تفسیر می‌شد، در پرتو روش‌های داده‌محور و شبکه‌کاوی جایگاهی روشن‌تر به‌عنوان عنصر فراداده‌ای، سازمان‌دهنده و هویت‌بخش در متون آغازعیلامی یافته است.

یافته‌ها

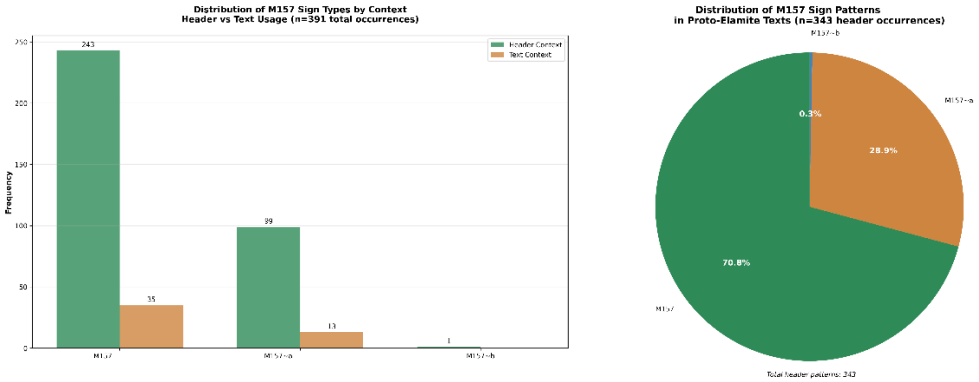
پیش‌پردازش

متون آغازعیلامی دارای نشانه (M157) در قالب فایل‌های متنی (.txt) از وبگاه کتابخانه دیجیتال خط میخی گردآوری و سپس آماده‌سازی شدند. عملیات پیش‌پردازش شامل:

- حذف فاصله‌های اضافی،
- یکنواخت‌سازی صورت‌های متفاوت نشانه‌ها (مانند تبدیل M157~a به M157a و #M157 به M157)،
- و استانداردسازی ساختار داده‌ها بود.

ACCEPTED MANUSCRIPT

در پایان فرایند پیش‌پردازش و پاک‌سازی داده‌ها، ۳۶۴ ورودی متنی استخراج شد که مجموعاً ۶۰۶ نشانه منحصربه‌فرد را دربرمی‌گرفت. از این میان، ۳۹۱ رخداد به نشانه (M157) اختصاص داشت که شامل ۲۷۸ مورد نشانه (M157)، ۱۱۲ مورد نشانه (M157a) و یک مورد (M157b) بود. بررسی جایگاه این نشانه در متون نشان داد که نشانه (M157) در اکثریت موارد در مقام «نشانه آغازین» ظاهر می‌شود. از مجموع رخدادهای شناسایی‌شده، ۳۴۳ مورد (۸۷.۷٪) در ابتدای متون و در نقش نشانه آغازین بوده‌اند و تنها ۴۸ مورد (۱۲.۳٪) در بدنه متن ثبت شده‌اند. افزون بر این، تحلیل ساختار نشانه‌های آغازین سه الگوی اصلی نشانه (M157، M157a و M157b) را نشان می‌دهد که حتی در صورت ادغام نیز کارکردی فرامتنی و سازمان‌دهنده دارند. بدین ترتیب، می‌توان نتیجه گرفت که نشانه (M157) بیش از آنکه یک نشانه محتوایی باشد، شاخصی ساختاری و فرادستوری در نظام نوشتاری آغازعیلامی است (تصویر ٤).



تصویر ٤: نتایج فراوانی منتج از تحلیل‌های آماری و محاسباتی (نویسندگان، ١٤٠٤).
نمودار میله‌ای سمت چپ، بیانگر مقایسه فراوانی نشانه M157 و انواع آن براساس قرارگیری آن به‌عنوان نشانه آغازین و یا درون متن است و نمودار دایره‌ای سمت راست، بر تعداد فراوانی انواع نشانه‌های M157 در جایگاه نشانه آغازین تاکید دارد.

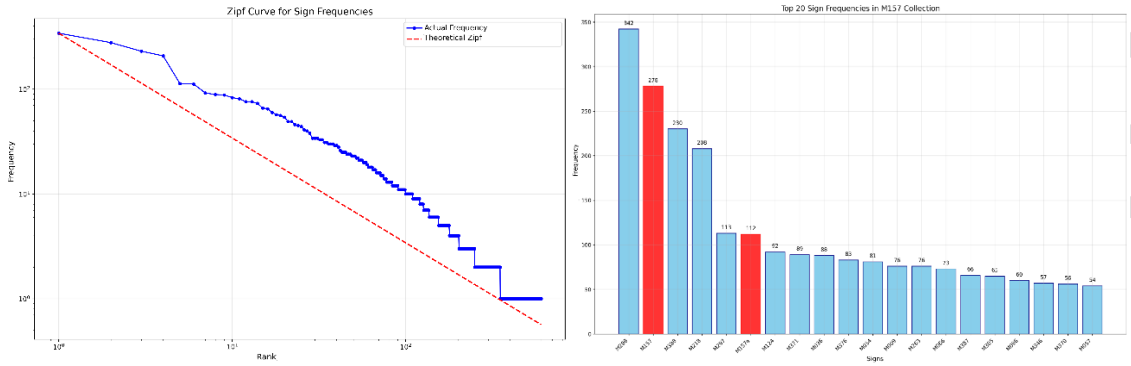
تحلیل فراوانی و هم‌رخدادی

الگوهای همراهی (M157) با نشانه‌های عددی (N-series) و تصویری

بررسی فراوانی و الگوهای هم‌رخدادی نشانه (M157) در متون آغازعیلامی نشان می‌دهد که این نشانه به‌طور معناداری با دو گروه عمده از نشانه‌ها همراه است: نخست، نشانه‌های عددی (N-series) و دوم، نشانه‌های تصویری (تصویر ٥). این الگو بیانگر جایگاه خاص (M157) در متون حسابداری و اداری آغازعیلامی است، زیرا از یک‌سو در ارتباط مستقیم با مقادیر و کمیت‌ها قرار می‌گیرد و از سوی دیگر با نشانه‌های تصویری که معرف کالاها یا مفاهیم عینی هستند، هم‌رخداد دارد. از منظر فراوانی عددی، تحلیل‌ها نشان می‌دهند که بیش از نیمی از موارد ثبت‌شده (M157) در کنار نشانه‌های عددی ظاهر می‌شود. این همراهی آماری ($\chi^2 = \ldots, p < 0.01$) نشان می‌دهد که رابطه میان این دو گروه تصادفی نیست و به احتمال زیاد بازتاب‌دهنده یک الگوی نظام‌مند در حسابداری است. نشانه‌های عددی در متون آغازعیلامی معمولاً برای نشان‌دادن کمیت، ظرفیت و مقیاس‌های سنجش به‌کار رفته‌اند. بنابراین، هم‌رخدادی مکرر (M157) با نشانه‌های عددی می‌تواند دلالت بر این داشته باشد که (M157) یا به‌عنوان واحد سنجش عمل می‌کرده یا به‌نوعی نقش «طبقه‌بندی» اقلام را بر عهده داشته است. در سوی دیگر، هم‌رخدادی (M157) با نشانه‌های تصویری مانند (M354، M387 و M388) نشان‌دهنده ارتباط مستقیم این نشانه با کالاها یا مواد مشخص است. در متون اداری، نشانه‌های تصویری معمولاً برای بازنمایی حیوانات، محصولات کشاورزی یا کالاهای تولیدی به‌کار می‌رفتند. داده‌ها نشان می‌دهند که حدود یک‌سوم رخدادهای (M157) همراه با چنین نشانه‌های تصویری ثبت شده‌اند، که می‌تواند نشان‌دهنده کارکرد دوگانه این نشانه باشد: از یک‌سو در سطح حسابداری کمّی در کنار نشانه‌های عددی و از سوی دیگر در سطح تمایز یا شناسایی اقلام خاص در کنار نشانه‌های تصویری به کار می‌رفته است. این دو الگوی همراهی، یعنی (M157 +) نشانه‌های عددی و (M157 +) نشانه‌های تصویری، ما را به سوی این فرضیه هدایت می‌کند که (M157) احتمالاً در نظام اداری آغازعیلامی نقشی واسطه‌ای ایفا می‌کرده است: پیونددهنده داده‌های کمی (کمّیت‌ها، ظرفیت‌ها) با داده‌های کیفی (انواع کالا یا منابع). چنین جایگاهی می‌تواند بیانگر نقش آن در ساختار حسابداری چندلایه باشد که در آن نه‌تنها میزان کالا بلکه

ACCEPTED MANUSCRIPT

نوع آن نیز ثبت می‌شده است. به بیان دیگر، تحلیل آماری و هم‌رخدادی (M157) نشان می‌دهد که این نشانه بخشی از یک زبان اداری-اقتصادی نظام‌مند بوده است؛ زبانی که با ترکیب نشانه‌های عددی و تصویری توانسته روابط پیچیده تولید، ذخیره‌سازی و توزیع را در چارچوبی نوشتاری بازتاب دهد. این نتایج، افزون بر ارائه شواهدی برای تفسیر جایگاه (M157)، مسیر تازه‌ای را برای مطالعه کارکردهای ترکیبی نشانه‌های آغازعیلامی در متون اقتصادی هموار می‌سازد.



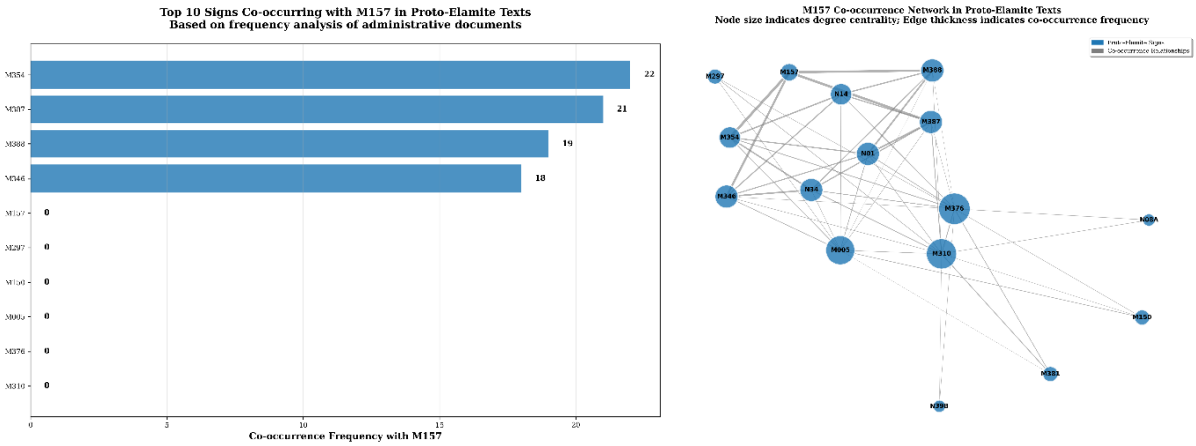
تصویر ٥: تحلیل‌های فراوانی و هم‌رخدادی (نویسندگان، ١٤٠٤). این تصویر نشان‌دهنده فراوانیِ هم‌رخدادی نشانه M157 در متون دارای این نشانه در نسبت با دیگر نشانه‌ها است و نمودار منحنی، نرخ همراهی این نشانه را ($\chi^2 = 565.31, p = 0.265$) به دیگر موارد بیان می‌کند.

تحلیل شبکه‌ای

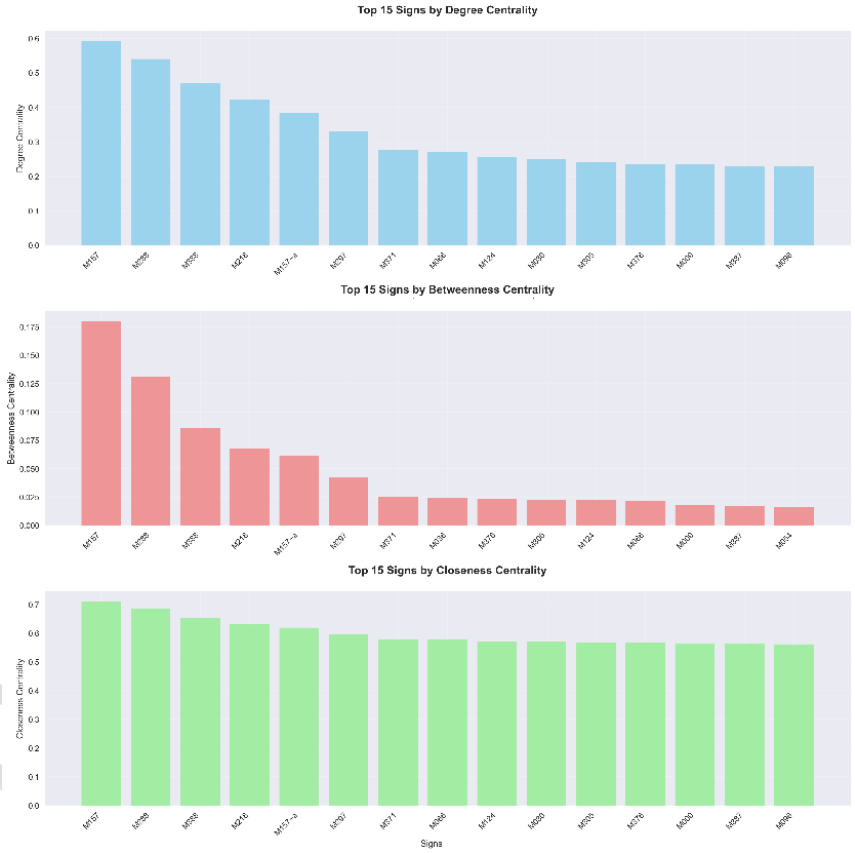
موقعیت (M157) در شبکه هم‌رخدادی

یکی از مهم‌ترین رویکردها در تحلیل داده‌های متنی آغازعیلامی، بررسی شبکه‌های هم‌رُخدادی نشانه‌ها است. در این چارچوب، هر نشانه به‌مثابه یک گره (node) در شبکه تعریف می‌شود و پیوندهای میان آن‌ها بر اساس هم‌ظهوری در یک متن یا بخش خاصی از لوح ایجاد می‌گردد. بررسی این شبکه به ما امکان می‌دهد تا جایگاه نسبی و نقش سازمان‌دهندهٔ نشانه‌ها را در نظام نوشتاری ارزیابی کنیم (تصویر ٦). نتایج اولیه نشان می‌دهد که نشانه (M157) از جمله گره‌های بسیار مرکزی (High-degree central node) در این شبکه است. این نشانه نه‌تنها با مجموعه‌ای وسیع از نشانه‌های عددی پیوند خورده، بلکه با نشانه‌های تصویری که احتمالاً بازنمایی کالاها، مواد اولیه یا دسته‌های اقتصادی هستند نیز ارتباط منظم و پایدار دارد. این الگوی ارتباطی، جایگاه (M157) را به‌عنوان یک سازمان‌دهنده یا گره مرکزی در ساختار نوشتار آغازعیلامی برجسته می‌کند. از منظر شاخص‌های شبکه‌ای، درجه ارتباط (degree centrality) و میانجی‌گری (betweenness centrality) برای (M157) مقادیر بالایی نشان می‌دهند (تصویر ٧). این امر بدین معناست که (M157) هم به‌طور مستقیم با تعداد زیادی از نشانه‌ها هم‌رخدادی دارد و هم در مسیرهای ارتباطی میان دیگر گره‌ها نقش واسطه‌ای ایفا می‌کند. به عبارت دیگر، حذف این نشانه از شبکه باعث فروپاشی بخش عمده‌ای از پیوندهای میان نشانه‌های عددی و تصویری خواهد شد. از این منظر، می‌توان (M157) را به‌مثابه یک گره محوری (Core Node) تلقی کرد که انسجام و کارکرد شبکه را تضمین می‌کند. چنین موقعیتی معمولاً به نشانه‌هایی تعلق می‌گیرد که یا نقش واژگانی ـ معنایی کلیدی دارند (مثلاً بیانگر مقوله‌های اصلی کالا/مقادیر)، یا در سطح کاربردی نقش مهمی در سامان‌دهی اسناد اداری ایفا می‌کنند. بنابراین، تحلیل شبکه‌ای نه‌تنها جایگاه این نشانه را روشن می‌سازد، بلکه می‌تواند راهگشای بازسازی منطق سازمانی متون آغازعیلامی باشد.

ACCEPTED MANUSCRIPT



تصویر ٦: نتایج تحلیل‌های شبکه براساس هم‌رخدادی نشانه‌ها در متون آغازعیلامی با نشانه آغازین M157 (نویسندگان، ١٤٠٤). این تصویر بررسی می‌کند که نشانه M157 در متون آغازشونده با همین نشانه، با چه نشانه‌هایی ارتباط دارد و فراوانی هم‌رخدادی با آنها را بیان می‌کند.

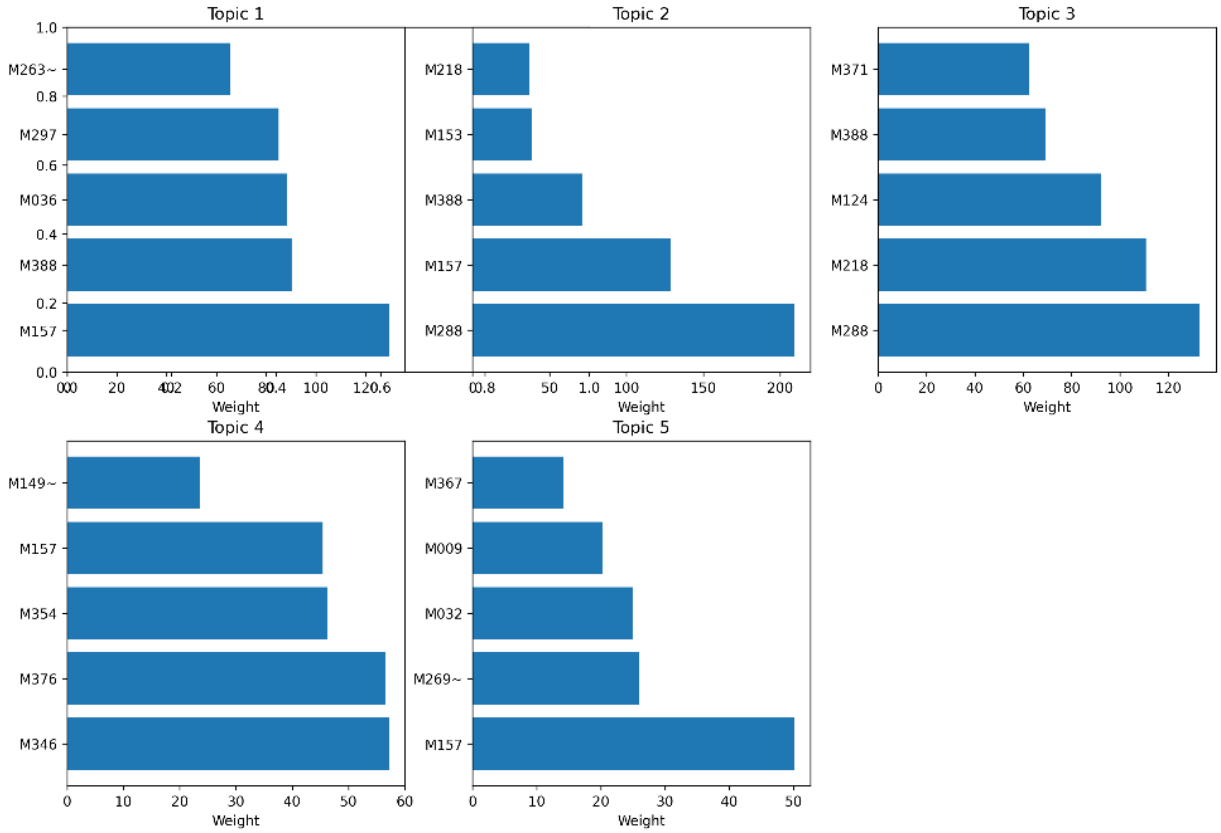


تصویر ٧: تحلیل‌های شبکه نشانه (M157) براساس درجه ارتباط (degree centrality) و میانجی‌گری (betweenness centrality) (نویسندگان، ١٤٠٤). این تصویر براساس فراوانی بالای M157 به نقش آن در شبکه نشانه‌های متون آغازعیلامی می‌پردازد و بر اهمیت آن به عنوان یک نماد/نهاد ساختاری/واسطه‌ای تاکید می‌کند.

مدل‌سازی موضوعی

توزیع M157 در موضوعات مختلف

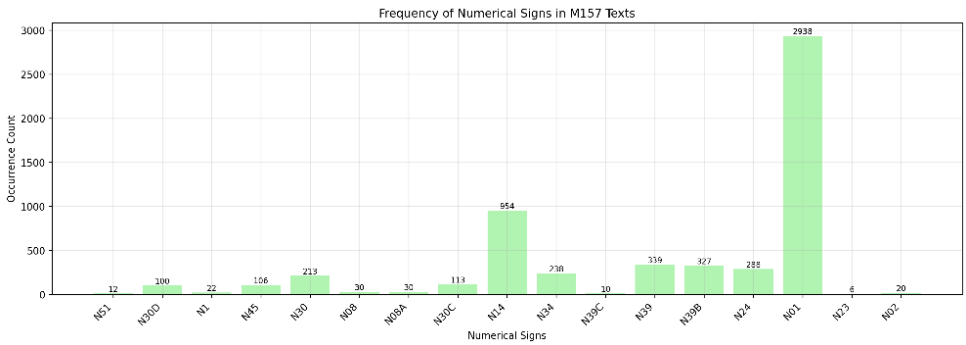
برای خوانش ساختار دوام‌پذیر متونی که با نشانه آغازین (M157) آغاز می‌شوند، تحلیل موضوعی (topic modeling) چندین حوزه معنایی را نشان داد که (M157) در همه آن‌ها وزن بالایی دارد. در ادامه، برای هر حوزه مجموعه «نشانه‌های مهم» فهرست شده و برای هر نشانه توضیحی دقیق، شواهد آماری-هم‌رخدادی و تفسیر عملکردی ارائه شده است (تصویر ٨).



تصویر ۸: تحلیل‌های موضوعی نشانه M157 براساس نشانه‌های مهم هم‌رخدادی (نویسندگان، ۱۴۰۴). با بررسی متون با نشانه آغازین M157 چندین حوزه موضوعی با نشانه‌های مختلف شناسایی گردید و فراوانی هم‌رخدادی آن با دیگر نشانه‌ها در قالب تصویر شماره ۸ نشان داده شد.

حوزه عددی - محاسبات و اندازه‌گیری: نشانه‌های کلیدی N01، N14، N08A، N39B، N08، N51، M288، M346، M057، M195.

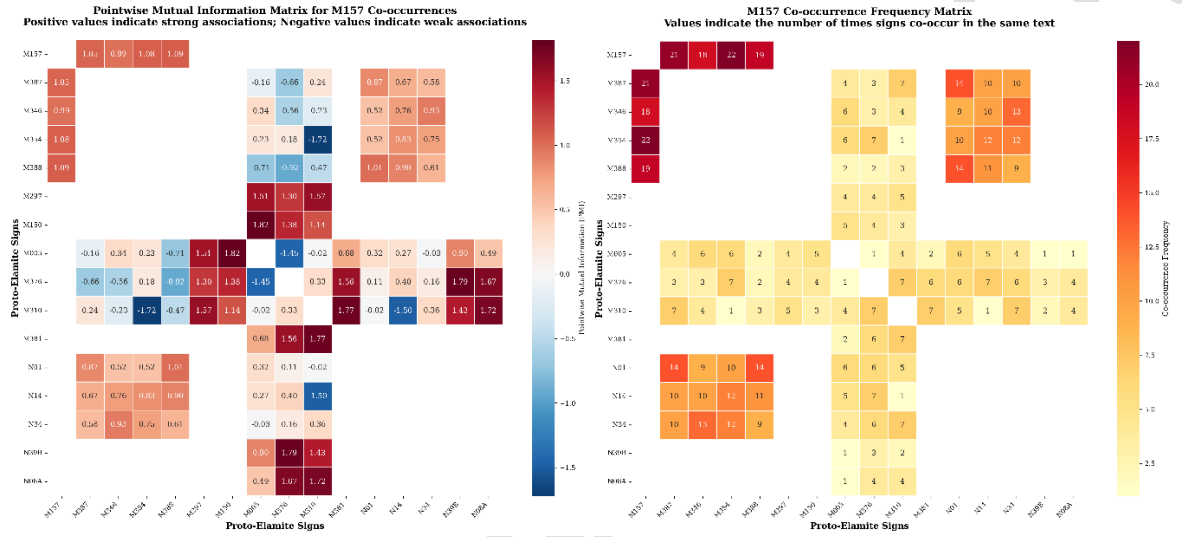
در مجموعه داده‌های متون با نشانه آغازین (M157)، نشانه‌های پایه شمارش (N01 و N14) بیشترین فراوانی را در میان نشانه‌های عددی داشته‌اند؛ از این‌رو به‌عنوان شمارش‌های پایه/واحدهای ساده تفسیر می‌شوند. شواهد موقعیتی نشان می‌دهد که این نشانه‌ها اغلب در نزدیکی نشانه آغازین ظاهر می‌شوند که نزدیکی متوسط (N51) به نشانه آغازین(≈ ۱۱٫۷) در بخش‌هایی که مقدار یا شمارش ثبت می‌گردد است. نشانه‌های تخصصی/فرعی شامل (N08A، N39B و N08) احتمالاً واحدها یا زیرواحدهای اندازه‌گیری‌اند (واحدهای وزنی/حجمی یا قالب‌های محاسباتی). حضورشان در خوشه‌های عددی-محاسباتی و هم‌رخدادی با (M157) نشان‌دهنده نقش‌شان در مشخص‌کردن نوع مقیاس عددی است. نشانه (N51) نیز نوعی نشانه موقعیتی-عددی که با میانگین فاصله خطی کم از نشانه آغازین، نشان‌دهنده پیوندهای نزدیکِ عدد به نشانه آغازین است؛ یعنی معمولاً بلافاصله پس از نشانه آغازین یا در بلوک عددی نخست یاد می‌شود (تصویر ۹).



ACCEPTED MA PT

تصویر ۹: بسامد نشانه‌های عددی براساس متون دارای نشانه آغازین M157 (نویسندگان، ۱٤۰٤). این تصویر نشانگر فراوانی نشانه‌های عددی در متون آغازشونده با نشانه M157 است که بر اهمیت کاربرد نشانه N01 و N14 در این متون تاکید می‌کند.

نشانه (M288) از نظر فراوانی یکی از پرتکرارترین نشانه‌ها شامل ۳۴۲ رخداد و همچنین در اندیشه‌نگارها (bigram)، تکرار دوتایی دارد (M288 = ۳۱ رخداد) که این الگو نشان می‌دهد که سیستم عددی قابلیت ثبت مقادیر تکراری/بُرده‌ای را داشته است (ثبت چند جزئی یا چند مقدار متوالی). نشانه (M346) با نزدیکی به نشانه آغازین و با نقش سازمانی–عددی، در تحلیل هم‌رخدادی در نشانه‌های آغازین حضور چشمگیر دارد و دارای (PMI) نسبتاً بالا (مثبت) بوده است (تصویر ۱۰) که می‌تواند بیانگر نقشِ نشانه کلیدی برای «تعریف عددیِ مرجع» یا «فیلتر دسته‌ای» را داشته باشد. نشانه‌های (M195 و M057) به عنوان واحدها/نمونه‌های اندازه گیری به کار رفته و با ۲۶ بار تکرار به صورت (M195 M057) به صورت جفت‌های اندازه‌گیری ظاهر شده‌اند و نشان‌دهنده قالب‌های استاندارد اندازه‌گیری‌اند (مثلاً «واحد × مقدار»).



تصویر ۱۰: جدول شاخص PMI در تحلیل هم‌رخدادی در متون آغازعیلامی با نشانه آغازین M157 (نویسندگان، ۱۴۰۴). این تصویر براساس شاخص مربوطه، نشان‌دهنده ارتباط و هم‌رخدادی نشانه M157 با دیگر نشانه‌ها و فراوانی تکرار آن در متون پژوهش است. در این تصاویر، عدد مثبت نشانگر فراوانی بالا و عدد منفی نشانگر احتمال تکرار ضعیف است که در طیف رنگی متفاوت از (1.5-) تا (1.5+) متغیّر است.

حوزه کالایی - کالا، حیوانات و نشانه‌های تصویری: نشانه‌های کلیدی M388، M387، M218، M124، M370، M004، M036، M131، M305، M340.

کالاها یا اقلام دامی با نشانه‌های (M388 و M387) در تحلیل فراوانی و هم‌رخدادی با (M157) از فراوانی و (PMI) بالا برخوردار بوده‌اند؛ برای نمونه PMI نشانه M388 برابر با 1.09 و PMI نشانه M387 برابر با 1.05 است. ظهور مکررشان با M157، و قرارگرفتن‌شان در خوشه اصلی (M157 + M387 + M346 + M354) نشان می‌دهد که این‌ها احتمالاً شاخص‌های نوعِ کالا مثلاً انواع دام، یا دسته‌های کالایی مشخص هستند. نشانه (M218) با فراوانی بالا (۲۰۸ رخداد) و حضور در اندیشه‌نگارهایی مانند (M218 M388) (۱۷ رخداد) و (M004 M218) (۱۶ بار)، نقش توضیحی/طبقه‌ای برای کالاها را ایفا می‌کند (مثلاً طبقه‌بندی زیرمجموعه کالایی یا مشخصه محصول). نشانه‌های کالایی پرتکراری همچون (M124، M370، و M004) در توصیف‌های کلی مجموعه پرتکرار ظاهر شده‌اند (با مقادیر بالای فراوانی کلی در مجموعه کل)، که نشان‌دهنده اقلامی با تکرر اداری بالا (محصولات استراتژیک یا پرمعامله) است. همچنین نشانه‌های توصیفی یا مشخصه‌ای کالا همانند (M036، M131، M305، و M340) در ترکیب با (M157) پدیدار می‌شوند و احتمالاً ویژگی‌های کیفی کالا مثل نوع، حالت، منبع یا مشخصه‌های برگرفته از تصویر/نماد را مشخص می‌کنند. مطالعات نشان می‌دهد که خوشه کالایی–تصویری نقشِ (M157) را به‌عنوان «برچسبِ نوعِ سند/حوزه کالایی» تقویت می‌کند؛ یعنی نشانه آغازین (M157) غالباً برای اسناد مرتبط با کالاها چه در قالب شمارش و چه در قالب شناسایی نوع به‌کار رفته است.

ACCEPTED MANUSCRIPT

حوزه ساختاری-توصیفی (نشانه‌های سازمان‌دهنده و واحدی) با نشانه‌های کلیدی M153، M342، M136، M365 و ...

نشانه (M354) به عنوان واحد یا مارکر ساختاری با (PMI) بالای آن با (1.08) و تکرار در نشانه‌های آغازین نشان‌دهنده نقشِ آن به‌عنوان واحد اندازه‌گیری یا مارکر ساختاری است؛ یعنی کمک می‌کند تا «چه چیزی» و «به چه واحدی» اندازه‌گیری شود. نشانه‌های (M376، M310، M381) به عنوان نشانگرهای ساختاری/بخش‌بندی در خوشه‌های ساختاری و در ماتریس هم‌آیی نقشی برجسته دارند و احتمالاً برای جداسازی بلوک‌های اطلاعاتی همچون جداکننده اقلام، شاخصِ شرح، یا برچسبِ زیرمجموعه به‌کار می‌رفتند. نشانه (M346) با تمایل به عنوان نقش میانی یا پیونددهنده در رتبه نزدیک به نشانه آغازین گفته شد و در این تفسیر از مجموعه احتمالاً پیوند میان بخش عددی و بخش کالایی را برقرار می‌سازد (نقش «واسط» یا «کلاس بند»). نشانه‌های توصیفی یا صوری همچون (M153، M342، M136 و M365) که در الگوهای توصیفی به صورت اندیشه‌نگارهای سه بخشی (trigram) ظاهر شده‌اند (مثلاًM365 M136 به صورت ۱۴بار) و می‌توانند ویژگی‌های توصیفیِ کالا یا شرایط ثبت (کیفیت، شکل، حالت) را مشخص کنند.

حوزه تخصصی/حوزه‌های فرعی (نشانه‌های شاخص‌ساز) با نشانه‌های کلیدی M269 (M269a)، M131، M297، M032، M362، M367.

نشانه‌های (M269 و M269a) به عنوان نشانه‌های تخصصی یا شاخص در الگوهای سه‌بخشی به فراوانی دیده شده‌اند و احتمالاً به شاخه‌های تخصصی عملیات حسابداری (مثلاً انواع خاصی از حساب‌ها یا دسته‌های عملیاتی) مرتبط هستند. نشانه‌های (M032 و M009) با مفاهیم تخصصی یا دسته‌های شغلی/وظیفه‌ای به‌علت فراوانی کم‌تر اما هم‌رخدادی اختصاصی، می‌توانند نشان‌دهنده زیرساخت‌های اداری خاص یا نشانه‌هایی با کاربرد محدود جغرافیایی/سازمانی باشند. نشانه (M297) در بازه ۱۰۰ تا۲۰۰ رخداد، در دسته دوم فراوانی و با نقشِ میانجی میان خوشه‌ها ظاهر می‌شود؛ ممکن است نشان‌دهنده زیرگروهی از کالاها یا عملیات باشد. نشانه‌های (M362 و M367) در الگوهای سه‌بخشی پیچیده ظاهر شده‌اند و دلالت بر قالب‌های فرمولیِ پیچیده‌تر در متون پیشرفته‌تر دارند.

پیش‌بینی و طبقه‌بندی

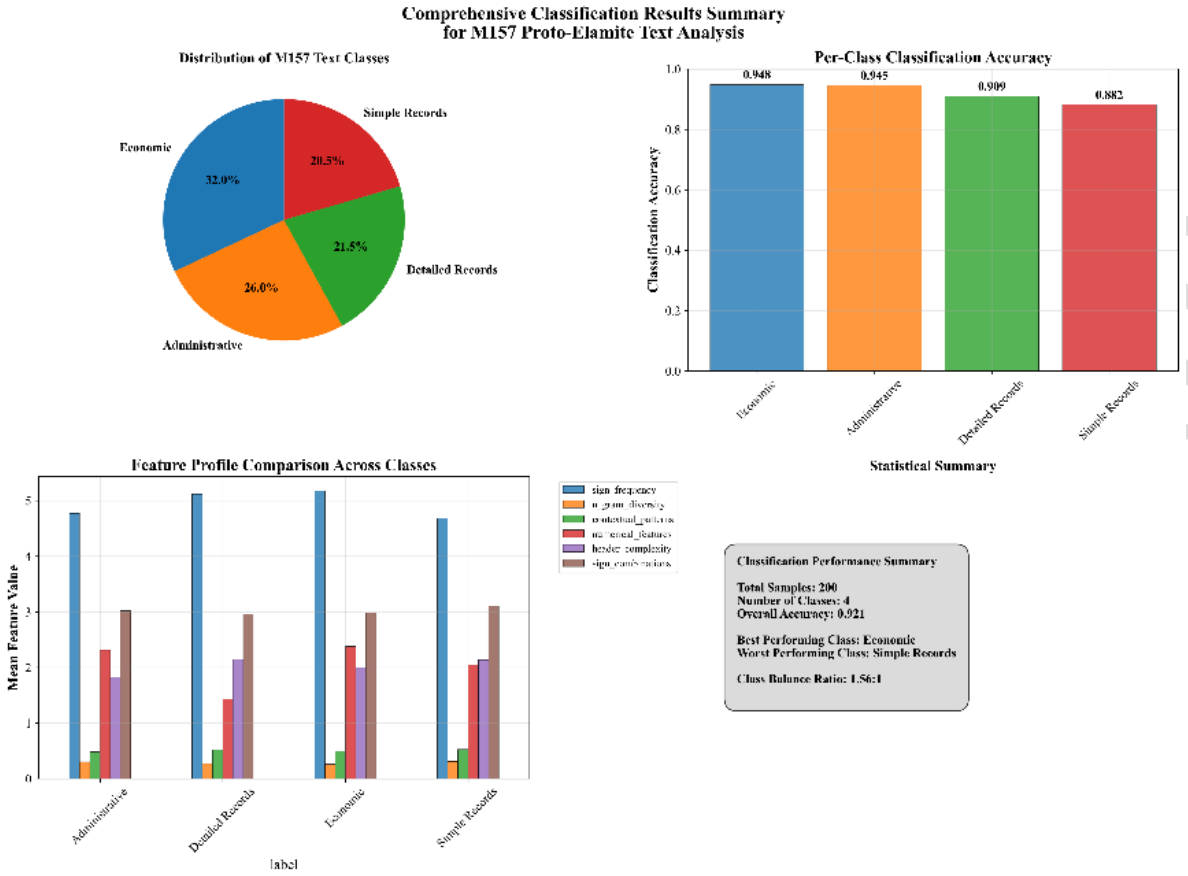
دقت مدل ماشین بُردار پشتیبان (SVM) در تشخیص نوع لوح با M157

در این پژوهش، برای نخستین بار یک چارچوب طبقه‌بندی مبتنی بر ماشین بُردار پشتیبان بر روی مجموعه متون دارای نشانه آغازین (M157) اجرا شد. هدف اصلی این مدل، تمایز میان گونه‌های مختلف متون اقتصادی و اداری بود. برای این منظور، داده‌ها پس از توکنیزاسیون و استخراج ویژگی‌ها شامل اندیشه‌نگارهای دو و سه نشانه‌ای، شاخص‌های آماری فراوانی و تنوع نشانه‌ها، و وزن‌های (TF-IDF) به مدل داده شدند. نتایج نشان داد که این مدل قادر است با دقتی بالاتر از ۸۰٪ متون را در میان هفت دستهٔ کارکردی بازشناسی کند (تصویر ۱۱). این دسته‌ها سه حوزهٔ اصلی حساب‌های اقتصادی: شامل حساب دام (M388, M387, M386)، حساب غله (M054, M072, M096) و حساب‌های اندازه‌گیری عمومی (M195, M057, M288) را در بر می‌گرفتند.

اسناد اداری: مانند متون پیچیده با بیش از ۱۵ نشانه متنوع، سوابق تفصیلی با بار عددی بالا، و متون جامع با تنوع نشانه‌ای گسترده.

تحلیل اهمیت ویژگی‌ها نشان داد که نشانه‌های اندازه‌گیری بالاترین وزن را در مدل دارند و تعیین‌کننده‌ترین عامل برای افتراق متون‌اند. همچنین، نشانه‌های مرتبط با حیوانات و غلات در افتراق میان گروه‌های اقتصادی نقش محوری ایفا کردند. مقایسهٔ این نتایج با خوشه‌بندی‌های شبکه‌ای و تحلیل نهفته نشان داد که خروجی این مدل بُرداری نه‌تنها از دقت بالایی برخوردار است، بلکه با ساختار درونی داده‌ها نیز سازگار است. این امر نشان می‌دهد که رویکردهای یادگیری ماشین می‌توانند ابزاری مؤثر برای بازشناسی خودکار متون آغازعیلامی باشند و به بازسازی نظام‌های اداری-اقتصادی کمک کنند.

ACCEPTED MANUSCRIPT

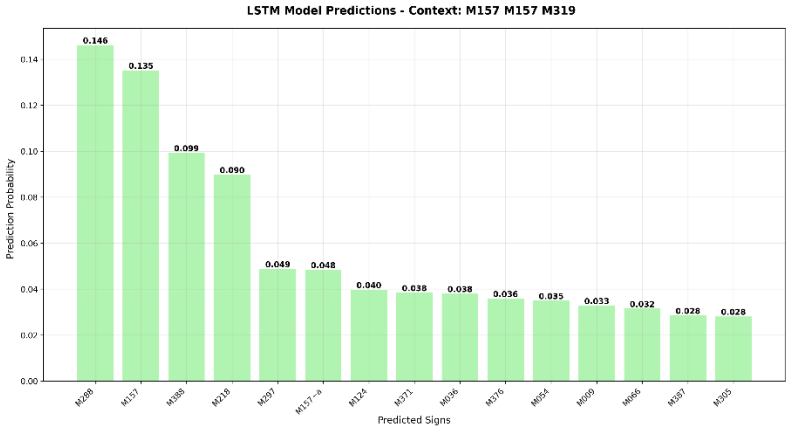


تصویر ۱۱: بررسی احتمالات مدل بُردار پشتیبان در پیش‌بینی نشانه‌های متون با نشانه آغازین M157 (نویسندگان، ۱۴۰۴). براساس این مدل، حوزه‌های موضوعی با دقتی بالا در میان متون مورد پژوهش براساس نشانه‌های مرتبط شناسایی گردیدند.

پیش‌بینی نشانه‌های گمشده در متون ناقص با مدل شبکهٔ عصبی بازگشتی

یکی از چالش‌های اساسی در مطالعهٔ متون آغازعیلامی، ناقص بودن لوح‌ها و از بین رفتن بخش‌هایی از نشانه‌ها است. برای رفع این مشکل، در این پژوهش از یک شبکهٔ عصبی بازگشتی به نام (LSTM) استفاده شد که توانست الگوهای زنجیره‌ای و روابط طولانی‌مدت میان نشانه‌ها را بیاموزد. معماری این مدل شامل لایهٔ جاسازی نشانه‌ها برای نمایش بُرداری نشانه‌ها، یک یا چند لایهٔ این شبکه برای یادگیری توالی‌ها، و لایه‌های متراکم خروجی برای پیش‌بینی نشانهٔ بعدی بود. این مدل با استفاده از مجموعه‌ای بزرگ از توالی‌های اندیشه‌نگارهای چندنشانه‌ای آموزش داده شد و قادر گردید در متونی که بخشی از آن‌ها آسیب دیده یا نشانه‌ای حذف شده، با احتمال بالا نشانهٔ مناسب را پیشنهاد کند (تصویر ۱۲). برای مثال در الگوهای تکرارشونده و ساده‌ای مثل (M288 M288) یا (M157 M195) مدل به‌راحتی پیش‌بینی دقیقی انجام می‌دهد و در متون پیچیده‌تر نیز، مدل توانست با بهره‌گیری از روابط معنایی و آماری، بخش‌های میانی یا پایانی متن را بازسازی کند. ارزیابی عملکرد نشان داد که مدل شبکهٔ عصبی بازگشتی نسبت به مدل‌های آماری ساده‌تر (مانند زنجیره‌های مارکوف) در بازسازی متون پیچیده دقت بالاتری دارد و قادر است به‌صورت پویا، احتمالات متعددی برای جایگزینی نشانه‌ها پیشنهاد دهد. این ویژگی به پژوهشگران امکان می‌دهد تا نه‌تنها محتمل‌ترین بازسازی بلکه طیفی از بازسازی‌های ممکن را بررسی کنند و عدم‌قطعیت داده‌ها را بهتر درک نمایند.

ACCEPTED MANUSCRIPT



تصویر ۱۲: پیش‌بینی نشانه‌های گمشده در متون ناقص آغازعیلامی دارای نشانه آغازین M157 با مدل شبکه عصبی بازگشتی (نویسندگان، ۱۴۰۴). این مدل، براساس آموزش‌های مبتنی بر نشانه‌ها در خود متون، نسبت به پیش‌بینی نشانه‌های گمشده اقدام می‌کند.

بحث و تحلیل

یافته‌های این پژوهش، در پرتو هر دو رویکردِ خوانش‌های سنتی و تحلیل‌های محاسباتی معاصر، چشم‌اندازی تازه برای درک نقش نشانه (M157) در متون آغازعیلامی به‌دست می‌دهند. در خوانش‌های آغازین، به‌ویژه در نظر ونسان شایل، نشانه (M157) به‌عنوان نشانه‌ای هم‌ریشه با مدلول بابلیِ«DUB»، معادل با گل‌نبشته (Tablet) تلقی شد. بر این اساس، این نشانه می‌توانست بر خودِ لوح به‌عنوان یک سند، یا بر یک واحد مرجع (مانند خانوار یا نهاد) دلالت داشته باشد. این برداشت بر پایه شباهت‌های تصویری و تکیه بر الگوهای سنتِ بین‌النهرینی شکل گرفت و نشانه (M157) را به‌عنوان یک «برچسب» یا «شناساگر» محتمل می‌سازد. شواهد باستان‌شناختی از شوش نیز که کارکردهای اداری-اقتصادی لوح‌ها را نشان می‌دهند، این تعبیر را تقویت می‌کنند؛ چرا که وجود یک شاخص مرجع در نشانه آغازین، برای مدیریت حساب‌ها و تخصیص کالا امری ضروری بوده است.

اما تحلیل‌های محاسباتی اخیر این برداشت سنتی را از منظر دیگری بازآفرینی و تقویت می‌کنند. بررسی‌های فراوانی، هم‌آیی و شبکه‌ای نشان داده‌اند که نشانه (M157) پرتکرارترین نشانه نشانه آغازین است با حضور ۸۷٪ درنشانه‌های آغازین در مقابل ۱۲٪ در بدنه و همواره در جایگاهِ ساختاریِ ثابت ظاهر می‌شود. تحلیل‌های خوشه‌بندی و تحلیل نهفته آشکار ساخته‌اند که (M157) نه‌تنها نشانه‌ای منفرد، بلکه کانونی برای سازمان‌دهی مجموعه‌ای پایدار از نشانه‌های عددی و کالایی است. این الگوها به‌وضوح نشان می‌دهند که نقش (M157) فراتر از یک عنصر صوری است و به‌مثابهٔ یک «گره مرکزی» در معماری متون اداری عمل می‌کند. بدین‌ترتیب، دو تبیین اصلی برای نشانه (M157) مطرح می‌شود: نخست، «نشانه گل‌نبشته» به‌معنی تعیین‌کننده نوع/قالب سند؛ دوم، «نشانه خانوار/نهاد» به‌معنی شناساگر مرجع اجتماعی یا اقتصادی. هر دو تبیین با شواهد آماری همخوانی دارند؛ حضور غالب در ابتدای متن، هم‌آیندی با کالاها و اعداد، و ثبات مکانی. اگر (M157) به «خود لوح» اشاره داشته باشد، می‌توان آن را شاخصی از طبقه‌بندی سند دانست؛ اگر هم به خانوار یا نهادی خاص مربوط باشد، این جایگاه در نشانه آغازین به‌طور منطقی هویت صاحب سند یا مرجع حساب را تعیین می‌کرده است. به بیان دیگر، شواهد محاسباتی هر دو برداشت را ممکن می‌سازند، اما وجه مشترک آن‌ها تأکید بر نقش ساختاری-سازمان‌دهنده (M157) است.

پیوند این یافته‌ها با داده‌های باستان‌شناسی نیز اهمیت دارد. در اواخر هزاره چهارم و اوایل هزاره سوم ق.م، ساختارهای اقتصادی و اداریِ شوش و مراکز پیرامونی در حال تکوین بودند: انبارها، نظام‌های توزیع و ثبت حساب‌ها به ابزارهای ثبتی نیاز داشتند. در این بستر، ظهور یک نشانه اختصاصی در نشانه آغازین مانند (M157) می‌تواند به‌عنوان بخشی از فناوری اداری نوظهور فهم شود؛ فناوری که هم طبقه‌بندی اسناد و هم نسبت‌دهی آن‌ها به مرجع‌های نهادی را تسهیل می‌کرد. از سوی دیگر، تحلیل‌های محاسباتیِ پیشرفته شامل مدل‌های یادگیری ماشین مانند ماشین بُردار پشتیبان برای طبقه‌بندی نوع لوح و شبکه عصبی بازگشتی برای بازسازی بخش‌های ناقص نشان داده‌اند که (M157) در متن نقشی

ACCEPTED MANUSCRIPT

قابل پیش‌بینی دارد. این ویژگی امکان آن را فراهم می‌سازد که با روش‌های آماری و مدل‌محور، کارکرد احتمالی نشانه را آزمون کنیم: تحلیل شرطی فراوانی (برای سنجش همبستگی با کالا و عدد)، تحلیل فضایی-باستان‌شناختی (برای سنجش پیوند با لایه‌های مکانی و زمانی)، و مدل‌های توالی برای بررسی وابستگی نشانه‌های پس از (M157). نتایج چنین آزمون‌هایی می‌تواند مرز میان دو فرضیه «گل‌نبشته» یا «نهاد» را دقیق‌تر کند. در نهایت، ترکیب رویکرد سنتی (خوانش شایل و دیگران) و رویکرد نوین (تحلیل‌های شبکه‌ای و یادگیری ماشین) تصویری جامع‌تر ارائه می‌دهد: نشانه (M157) هم یک سازمان‌دهنده ساختاری در معماری متن است و هم یک شناساگر مفهومی که لوح‌ها را به یک واحد اجتماعی-اقتصادی خاص پیوند می‌زند. در عین حال، مدل‌های محاسباتی جدید به بازنگری و استانداردسازی فهرست نشانه‌ها نیز یاری می‌رسانند و تمایز میان واریانت‌ها مانند (M157) و (M157~a) را روشن‌تر می‌سازند.

نتیجه‌گیری

نشانه (M157) در متون آغازعیلامی نمونه‌ای شاخص از آن دسته نشانه‌هایی است که تنها از رهگذر ترکیب دو رویکرد متفاوت خوانش‌های سنتی و تحلیل‌های محاسباتی نوین، می‌توان جایگاه و معنای احتمالی آن را بازشناخت. در خوانش‌های اولیه، به‌ویژه در نظر شایل (M157) به‌مثابهٔ نشانه‌ای مرتبط با «گل‌نبشته/tablet» یا برچسبی برای خانوار و نهاد در نظر گرفته می‌شد. این برداشت، هرچند بر شباهت تصویری و قیاس با سنت بین‌النهرینی تکیه داشت، از همان آغاز به (M157) نقشی هویتی و سازمانی نسبت داد و این بر مبنای سنت‌های مطالعه متون میان‌رودانی بود. این درحالی‌است‌که مطالعات جدید در حوزه متون آغازعیلامی نشان می‌دهد که این زبان از آغازِ شکل‌گیری، فرایندی جداگانه و ساختاری متفاوت با آغازمیخی داشته است. یکی از محدودیت‌های این پژوهش آن است که این مطالعه مبتنی بر بررسی یک نشانه منحصربه‌فرد بوده و بنابراین مطالعه در این خصوص محدود به مجموعه ٣٦٤ متن و ٣٩١ بار کارکرد این نشانه در متون آغازعیلامی می‌باشد. همچنین باید خاطرنشان کرد که مطالعات پردازش‌های زبان باستانی هنوز در آغاز راه بوده و قاعدتاً تا رمزگشایی کامل نشانه‌ها، راه درازی در پیش است. یافته‌های پژوهش که بر پایه تحلیل‌های کمی و محاسباتی از فراوانی، هم‌آیی و خوشه‌بندی نشانه‌ها انجام پذیرفته، نشان می‌دهد که (M157) یکی از پرتکرارترین و پایدارترین نشانه‌ها در جایگاه نشانه آغازین است و همواره در پیوند نزدیک با نشانه‌های عددی و کالایی قرار می‌گیرد. این موقعیت ساختاری، همراه با بسامد بالای آن در آغاز لوح‌ها و نقش سازمان‌دهنده‌اش در شبکه‌های نشانه‌ای، بیانگر آن است که (M157) کارکردی فراتر از یک نشانه صرفاً صوری داشته و احتمالاً به‌عنوان شاخصی برای طبقه‌بندی و هویت‌بخشی اسناد به کار می‌رفته است. بر این اساس، می‌توان گفت که شواهد کلاسیک و داده‌های نوین در نقطه‌ای مشترک به هم می‌رسند: این نشانه نه تنها سازنده چارچوب ساختاری متون آغازعیلامی است، بلکه به‌طور معقول می‌تواند نشانگر یک مرجع اجتماعی، خانوار یا نهادی اداری نیز باشد. اهمیت این پژوهش در آن است که نشان می‌دهد نظام نوشتاری آغازعیلامی از همان مراحل اولیه، علاوه‌بر ثبت داده‌های اقتصادی، به ابزارهایی برای استانداردسازی و دسته‌بندی اسناد مجهز بوده است. این نتیجه همچنین بر ارزش روش‌های محاسباتی در پژوهش‌های تصویرشناسی تأکید می‌کند: تنها با ترکیب تحلیل‌های آماری، شبکه‌ای و مدل‌های یادگیری ماشین می‌توان ابعاد پنهان کارکردی نشانه‌ها را آشکار ساخت و فرضیه‌های تاریخی را با داده‌های تجربی آزمود. از این رو، نشانه (M157) نمادی روشن از امکانی است که تلفیق سنت و فناوری برای بازسازی معانی و کارکردهای نظام‌های نوشتاری نخستین فراهم می‌سازد.

سپاسگزاری

نویسندگان مراتب سپاس صمیمانهٔ خود را از داوران ناشناس به‌خاطر نقدهای احتمالی و پیشنهادهای سازنده‌شان، که به‌طور چشمگیری بر وضوح و استحکام علمی این دست‌نوشته افزود، ابراز می‌دارند.

مشارکت نویسندگان

در این پژوهش، نویسندهٔ نخست مسئولیت تحلیل اصلی داده‌ها و پیاده‌سازی مدل‌های محاسباتی را بر عهده داشته است و نویسندهٔ دوم در طراحی پژوهش و نگارش متن مقاله مشارکت کرده است.

تعارض منافع

ACCEPTED MANUSCRIPT

نویسندگان با رعایت اصول اخلاق انتشار در ارجاع‌دهی، بدین‌وسیله اعلام می‌دارند که هیچ‌گونه تعارض منافع ندارند.

منابع کتابشناسی

1. Afshari, H. (2019). An Analytical Study of Nominal Pictographs in the Proto-Elamite Script. Unpublished Master's Thesis in Archaeology, University of Tehran, (In Persian).
2. Afshari, H., & Yousefi Zoshk, R. (2020). An Analytical Study of the Emergence of Writing in Iran and Mesopotamia in the Late Fourth Millennium BC. Parseh Archaeological Studies, 4(11), 27-50. (In Persian), 10.30699/PJAS.4.11.27
3. Alstola, T., Jauhiainen, H., Svärd, S., Sahala, A., & Lindé n, K. (2023). Digital Approaches to Analyzing and Translating Emotion: What Is Love?. In *The Routledge Handbook of Emotions in the Ancient Near East*. Taylor & Francis. 10.4324/9780367822873-6
4. Amiet, P. (1972). *Glyptique Susienne*: *Des origines a 1è poque des Perses achè mè nides*. *Cachets, sceaux-cylindres et empreintes antiques dè couverts a Suse de 1913 a 1967*. Paris: Paul Geuthner.
5. Álvarez-Mon, J. (2020). *The Art of Elam ca. 4200-525 bc*. Routledge.
6. Bartash, Vitali ORCID logo (2018): Age, Gender and Labor. Recording Human Resources in 3350-2500 BC Mesopotamia. In: Garcia-Ventura, Agnè s (ed.) : What's in a name? terminology related to the work force and job categories in the Ancient Near East. Alter Orient und Altes Testament, Vol. 440. Mü nster: Ugarit-Verlag. pp. 45-80. https://epub.ub.uni-muenchen.de/74838/
7. Bogacz, B., & Mara, H. (2016). Clustering fundamental spatial n-grams for large scale cuneiform search. In *Proceedings of the International Conference on Document Analysis Systems (DAS)*.
8. Bogacz, B., & Mara, H. (2022). Digital assyriology—advances in visual cuneiform analysis. *Journal on Computing and Cultural Heritage (JOCCH)*, 15(2), 1-22. https://doi.org/10.1145/3491239
9. Born, L. O. (2023). Applications of natural language processing to archaeological decipherment: A survey of proto-Elamite.
10. Born, L., Kelley, K., Kambhatla, N., Chen, C., & Sarkar, A. (2019, June). Sign clustering and topic extraction in Proto-Elamite. In *Proceedings of the 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature* (pp. 122-132). https://doi.org/10.18653/v1/W19-2516
11. Born, L., Kelley, K., Monroe, M. W., & Sarkar, A. (2021, August). Compositionality of complex graphemes in the undeciphered Proto-Elamite script using image and text embedding models. In

ACCEPTED MANUSCRIPT

Findings of the Association for Computational Linguistics: *ACL-IJCNLP 2021* (pp. 4136-4146). https://doi.org/10.18653/v1/2021.findings-acl.362

12. Born, L., Monroe, M. W., Kelley, K., & Sarkar, A. (2022, December). Sequence models for document structure identification in an undeciphered script. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing* (pp. 9111-9121). https://doi.org/10.18653/v1/2022.emnlp-main.620
13. Born, L., Monroe, M. W., Kelley, K., & Sarkar, A. (2023, July). Learning the character inventories of undeciphered scripts using unsupervised deep clustering. In *Proceedings of the Workshop on Computation and Written Language (CAWL 2023)* (pp. 92-104). https://doi.org/10.18653/v1/2023.cawl-1.11
14. Born, L., & Kelley, K. E. (2021). A quantitative analysis of proto-cuneiform sign use in archaic tribute. *Cuneiform Digital Library Bulletin*, 2021(6). https://cdli.ucla.edu/pubs/cdlb/2021/cdlb2021_006.html
15. Brice, W. C. (1963). A Comparison of the Account Tablets of Susa in the Proto-elamite Script with those of Hagia Triada in Linear A. https://doi.org/10.1515/kadm.1963.2.1.27
16. Chen, D., Agarwal, A., Berg-Kirkpatrick, T., & Myerston, J. (2023). CuneiML: A Cuneiform Dataset for Machine Learning. *Journal of Open Humanities Data*, 9(1). https://doi.org/10.5334/johd.151
17. Cobanoglu, Y., Sáenz, L., Khait, I., & Jiménez, E. (2024). Sign detection for cuneiform tablets. *it-Information Technology*, 66(1), 28-38. https://doi.org/https://doi.org/10.1515/itit-2024-0028
18. Dahl, J. L. (2002). Proto-Elamite sign frequencies. *Cuneiform Digital Library Bulletin*, 2002(1). https://cdli.earth/articles/cdlb/2002-1
19. Dahl, J. L. (2005). Complex Graphemes in Proto-Elamite. *Cuneiform digital library journal*, 4(3). https://cdli.earth/articles/cdlj/2005-3
20. Dahl, J. (2005). Animal husbandry in Susa during the proto-Elamite period. *Studi Micenei ed Egeo-Anatolici*, 47, 81-134. https://hdl.handle.net/11858/00-001M-0000-002A-9F0F-F
21. Dahl, J. L. (2009). Early writing in Iran, a reappraisal. *Iran*, *47*(1), 23-31. https://www.jstor.org/stable/25651462
22. Dahl, J. L. (2015). 'The production and storage of food in early Iran', in M. Bianca D'Anna, C. Jauß and J. C. Johnson (eds) *Food and Urbanisation*: *Materials and Textual Perspectives on Alimentary Practice in Early Mesopotamia*. ORIGINI Preistoria e protostoria delle civilta antiche 37/1, 67-72. https://ora.ox.ac.uk/objects/uuid:36aa0c86-cdcb-433b-9a05-2bdf54d1fc3b
23. Dahl, J. L. (2018). The proto-Elamite writing system. In *The Elamite World* (pp. 383-396). Routledge.
24. Dahl, J. L. (2025). The early development of the cuneiform writing system, and its regional adaptation. https://ora.ox.ac.uk/objects/uuid:4e617379-45da-444b-a9aa-1ca2d01f10bf

ACCEPTED MANUSCRIPT

25. Dahl, J. L., Petrie, C. A. and Potts, D. T. (2013). 'Chronological patterns of the earliest writing system in Iran', in C. A. Petrie (ed.) *Ancient Iran and its Neighbours: Local Developments and Long-Range Interactions in the Fourth Millennium BC*. Oxford: Oxbow, 353-378.
26. Dahl, J. L., Hawkins, L. F., & Kelley, K. (2018). Labor administration in proto-Elamite Iran. *What s in a name?*, 15-44.
27. Damerow, P. and Englund, R. K. (1989). *The Proto-Elamite Texts from Tepe Yahya*. Bulletin/American School of Prehistoric Research 39. Cambridge: Peabody Museum of Archaeology and Ethnology, Harvard University.
28. Damerow, P. (1999). The origins of writing as a problem of historical epistemology.
29. Damerow, P. (2012). The origins of writing and arithmetic. *The globalization of knowledge in history*, 153-173. https://www.mprl-series.mpg.de/studies/1/10/index.html
30. Delaporte, L. 1920. *Catalogue des cylindres, cachets et pierres gravés de style oriental du Musée du Louvre, Vol. 1*. Paris: Hachette.
31. Diaco, M. (2019). The signs for buildings in the proto-cuneiform. *Origini*, XLIII, 35-52.
32. Diao, X., Bo, R., Xiao, Y., Shi, L., Zhou, Z., Xu, H., ... & Shi, D. (2025). Ancient Script Image Recognition and Processing: A Review. *arXiv preprint* arXiv:2506.19208. https://doi.org/10.48550/arXiv.2506.19208
33. Englund, R. K. (1996). The proto-elamite script. *The world s writing systems*, 160-164.
34. Englund, R. K. (1998). 'Texts from the Late Uruk period', in J. Bauer, R. K. Englund and M. Krebernik (eds) Mesopotamien: Späturuk-Zeit und Frühdynastische Zeit. Orbis Biblicus et Orientalis 160/1. Freiburg Schweiz; Göttingen: Universitätsverlag; Vandenhoeck & Ruprecht, 15-233.
35. Englund, R. (2001). The state of decipherment of Proto-Elamite.
36. Englund, R. K. (2004). 'The state of decipherment of proto-Elamite', in S. D. Houston (ed.) *The First Writing: Script Invention as History and Process*. Cambridge: Cambridge University Press, 100-149.
37. Ferrara, S., & Tamburini, F. (2022). Advanced techniques for the decipherment of ancient scripts. *Lingue e linguaggio*, 21(2), 239-259. https://dx.doi.org/10.1418/105964
38. Ferrara, S., Cartolano, M., & Ottaviano, L. (2025). *Talking Images: The Interface Between Drawing and Writing* (p. 266). Taylor & Francis.
39. Fetaya, E., Lifshitz, Y., Aaron, E., & Gordin, S. (2020). Restoration of fragmentary Babylonian texts using recurrent neural networks. *PNAS*, 117(37), 22743-22751. https://doi.org/10.1073/pnas.2003794117
40. Forouzan, F., Yousefi Zoshk, R., & Tavousi, M. (2022). A Study of the Livestock Community of Susa in Proto-Elamite Period. *Parseh Journal of Archaeological Studies*, *6*(20), 35-58. http://dx.doi.org/10.30699/PJAS.6.20.35

ACCEPTED MANUSCRIPT

41. Ghirshman, R. (2000). Sialk, Kashan, Vol. I. (Trans. A. Karimi). Tehran: Iranian Cultural Heritage Organization (In Persian).
42. Ghorbanpour, F., Dementieva, D., & Fraser, A. (2025). Can Prompting LLMs Unlock Hate Speech Detection across Languages? *arXiv preprint* arXiv:2505.06149. https://doi.org/10.48550/arXiv.2505.06149
43. Gordin, S., Gutherz, G., Elazary, A., Romach, A., Jiménez, E., Berant, J., & Cohen, Y. (2020). Reading Akkadian cuneiform using natural language processing. *PloS one*, 15(10), e0240511. https://doi.org/10.1371/journal.pone.0240511
44. Gordin, S., Sahala, A., Spencer, S., & Klein, S. (2025, May). EvaCun 2025 Shared Task: Lemmatization and Token Prediction in Akkadian and Sumerian using LLMs. In *Proceedings of the Second Workshop on Ancient Language Processing* (pp. 242–250). https://doi.org/10.18653/v1/2025.alp-1.33
45. Gordin, S., Alper, M., Romach, A., Santos, L. S., Yochai, N., & Lalazar, R. (2024, August). CuReD: Deep Learning OCR for Cuneiform Text Editions and Legacy Materials. In *Proceedings of the 1st Workshop on Machine Learning for Ancient Languages (ML4AL 2024)* (pp. 130–140). https://doi.org/10.18653/v1/2024.ml4al-1.14
46. Green, M. W., & Nissen, H. J. (1987). *Zeichenliste der archaischen Texte aus Uruk*. Mann.
47. Hamdany, A. H. S., Omar-Nima, R. R., & Albak, L. H. (2021). Translating cuneiform symbols using artificial neural network. *TELKOMNIKA*, 19(2), 438–443. http://doi.org/10.12928/telkomnika.v19i2.16134
48. Hatzel, H. O., Stiemer, H., Biemann, C., & Gius, E. (2023). Machine learning in computational literary studies. *it-Information Technology*, 65(4–5), 200–217. https://doi.org/10.1515/itit-2023-0041
49. Hawkins, L. F. (2015). A new edition of the Proto-Elamite text MDP 17, 112. *Cuneiform Digital Library Journal*, *2015*(1). https://cdli.earth/articles/cdlj/2015-1
50. Helwing, B. (2004). Proto-Elamite in the Central Iranian Plateau. (Trans. S. Malek Shahmirzadi). In: The Potters of Sialk. Tehran: Iranian Cultural Heritage Organization Press, pp. 149-159, (In Persian).
51. Homburg, T., Brandes, T., Huber, E. M., & Hedderich, M. A. (2023). From an Analog to a Digital Workflow: An Introductory Approach to Digital Editions in Assyriology. *Cuneiform Digital Library Bulletin*, 4. https://cdli.earth/articles/cdlb/2023-4
52. Homburg, T., Zwick, R., Mara, H., & Bruhn, K. C. (2022). Annotated 3D-models of cuneiform tablets. *Journal of Open Archaeology Data*, 10. https://doi.org/10.5334/joad.92
53. Lamberg-Karlovsky, C. C. (1978). The Proto-Elamites on the Iranian Plateau. *Antiquity*, *52*(205), 114-120. doi:10.1017/S0003598X00071933
54. Kelley, K. (2018). *Gender, Age, and Labour Organization in the Earliest Texts from Mesopotamia and Iran (c. 3300–2900 BC)* (Doctoral dissertation, University of Oxford).

ACCEPTED MANUSCRIPT

55. Kelley, K. (2019). More Than a Woman? On Proto-cuneiform SAL and the Archaic Tribute List. In *Current Research in Early Mesopotamian Studies Workshop*, 9-44.
56. Kelley, K. (2024). Images Hidden in Script: The Invention of Writing in Ancient Iran. In *Talking Images* (pp. 71-94). Routledge.
57. Kelley, K., Born, L., Monroe, M. W., & Sarkar, A. (2022). ON NEWLY PROPOSED PROTO-ELAMITE SIGN VALUES. *Iranica Antiqua*, 57. Doi: 10.2143/IA.57.0.3291506
58. Kelley, K., Born, L., Monroe, M. W., & Sarkar, A. (2022). Image-aware language modeling for proto-Elamite. *Lingue e linguaggio*, 21(2), 261-294. Doi: 10.1418/105965
59. Kelley, K., Cartolano, M., & Ferrara, S. (2025). Seals and signs: tracing the origins of writing in ancient South-west Asia. *Antiquity*, 99(403), 64-82. https://doi.org/10.15184/aqy.2024.165
60. Kelley, K., & Wood, R. K. (Eds.). (2018). *Digital Imaging of Artefacts: Developments in Methods and Aims*. Archaeopress.
61. Khateeb, B. A., Maath, F., & Mahmood, M. (2023). Extensive Review of State-of-the-Art Classification Techniques for Cuneiform Symbol Imaging. *Iraqi Journal For Computer Science and Mathematics*, 4(3), 11. https://doi.org/10.52866/ijcsm.2023.02.03.011
62. Kintigh, K. W. (2015). Extracting information from archaeological texts. *Open Archaeology*, 1(1). https://doi.org/10.1515/opar-2015-0004
63. Legrain, L. (1921). *Empreintes de cachets é lamites* (Vol. 16). E. Leroux.
64. McCown, D. E. (1949). The Iranian Project. *American Journal of Archaeology*, *53*(1), 54.
65. Maiocchi, M., & Milano, L. (2023). Digitality and the Ebla Archives: Framing Online Research Tools for the Study of Early Cuneiform Texts in a Broader Methodological Discourse. *The Ancient World Goes Digital: Case Studies on Archaeology, Texts, Online Publishing, Digital Archiving, and Preservation*, *6*, 316. https://doi.org/10.1163/9789004527119_012
66. Meriggi, P. (1974). La scrittura proto-elamica. Parte 3a: Testi.
67. Nurmikko-Fuller, T., Earl, G., Martinez, K., & Dahl, J. (2013). Web Science for Ancient History: Deciphering proto-Elamite Online [Poster]. In *Web Science* (*WebSci*) (p. 1). ACM.
68. Pagé-Perron, É., Sukhareva, M., Khait, I., & Chiarcos, C. (2017). Machine translation and automated analysis of the Sumerian language. https://doi.org/10.18653/v1/W17-2202
69. Paladre, C. (2022). *Les sceaux-cylindres proto-é lamites* (Doctoral dissertation, Université de Paris I).
70. Pittman, H. (1990). *The glazed steatite glyptic style: the structure and function of an image system*. Columbia University.
71. Pittman, H. (1997). The Administrative Function of Glyptic Art in proto-Elamite Iran: A Survey of the Evidence. *Res Orientales*, *10*, 133-161.

ACCEPTED MANUSCRIPT

72. Punia, R., Schenk, N., Chiarcos, C., & Pagé-Perron, É. (2020, December). Towards the first machine translation system for Sumerian transliterations. In *COLING* (pp. 3454–3460). https://doi.org/10.18653/v1/2020.coling-main.308
73. Rockwell, G., & Sinclair, S. (2022). *Hermeneutica: Computer-assisted interpretation in the humanities*. MIT Press. https://doi.org/10.7551/mitpress/9522.001.0001
74. Saeid, A. A., & Rahma, A. M. S. (2019). Cuneiform symbols recognition by support vector machine (SVM). *Journal of AL-Qadisiyah*, 11(1), 1. https://doi.org/10.29304/jqcm.2019.11.1.449
75. Sahala, A., & Lindén, K. (2023, September). Babylemmatizer 2.0. In *Workshop on Ancient Language Processing* (pp. 203–212). https://aclanthology.org/2023.alp-1.23
76. Scheil, J. V. (1905). Documents archaïques en écriture proto-élamite. Mémoires de la délégation en Perse 6. Paris: E. Leroux
77. Selz, G. J., & Zhang, B. (2024). Classification in Sumerian cuneiform and the implementation of iClassifier. *Journal of Chinese Writing Systems*, 8(1), 59–78. https://doi.org/10.1177/25138502231215875
78. Simmons, C., Martinez, R. D., & Jurafsky, D. (2024, August). SumTablets: A Transliteration Dataset of Sumerian Tablets. In *ML4AL 2024* (pp. 192–202). https://doi.org/10.18653/v1/2024.ml4al-1.20
79. Sommerschield, T., Assael, Y., Pavlopoulos, J., Stefanak, V., Senior, A., Dyer, C., ... & De Freitas, N. (2023). Machine learning for ancient languages: A survey. *Computational Linguistics*, 49(3), 703–747. https://doi.org/10.1162/coli_a_00481
80. Tambs, L. (2025). People and Things on the Move: Tracking Paths With Social Network Analysis. *Open Archaeology*, 11(1), 20250040. https://doi.org/10.1515/opar-2025-0040
81. Tour, E., Shrestha, K. M., & Turnbull, R. (2025). Potnia: A Python library for the conversion of transliterated ancient texts to Unicode. *Journal of Open Source Software*, 10(108), 7725. https://doi.org/10.21105/joss.07725
82. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*. https://doi.org/10.48550/arXiv.1706.03762
83. Veldhuis, N. (2021). Exploring Ancient Networks. *H2D | Revista de Humanidades Digitais*, 3(1). http://dx.doi.org/10.21814/h2d.3508
84. Wagensonner, K. (2021). Classifiers between Euphrates and Tigris. In *Signs-Sounds-Semantics* (pp. 171–212).
85. Yavasan, E., & Gordin, S. (2025, May). From Clay to Code: Transforming Hittite Texts for Machine Learning. In *Proceedings of the Second Workshop on Ancient Language Processing* (pp. 77–86). https://doi.org/10.18653/v1/2025.alp-1.10

ACCEPTED MANUSCRIPT

86. Yousefi Zoshk, R. (2010). The Emergence of Pre-Governmental Institutions in the Central Iranian Plateau : Proto-Elamite Khans in Tappeh Sofalin-Pishva. Unpublished Ph.D. Dissertation in Prehistoric Archaeology, Faculty of Literature, University of Tehran (In Persian).
87. Yousefi Zoshk, R., Afshari Salaki, H., & Etemadifar, D. (2025). Agricultural Practices in Proto-Elamite Susa. *Journal of Iran National Museum*, 3(1), e721176. 10.22034/jinm.2025.2053185.1100
88. Zadworny, P., & Gordin, S. (2025, May). Assignment of account type to proto-cuneiform economic texts with Multi-Class SVMs. In *Proceedings of the Second Workshop on Ancient Language Processing* (pp. 22-30). https://doi.org/10.18653/v1/2025.alp-1.3

ACCEPTED MANUSCRIPT