

## ارائه چارچوبی برای داده گان ملی با تمرکز بر توسعه حاکمیت داده

نادر نقشینه\* فاطمه فهیم نیا\*\* حمیدرضا احمدیان\*\*\*<sup>۱</sup>

\* دانشیار، دکتری علم اطلاعات و دانش شناسی، دانشیار دانشکده مدیریت، دانشگاه تهران

\*\* دانشیار، دکتری علم اطلاعات و دانش شناسی، دانشیار دانشکده مدیریت، دانشگاه تهران

\*\*\* دانشجوی دکتری علم اطلاعات و دانش شناسی، دانشگاه تهران

تاریخ پذیرش: ۱۴۰۰/۰۲/۱۸

تاریخ دریافت: ۱۳۹۹/۱۰/۲۴

نوع مقاله: پژوهشی

### چکیده:

هدف این پژوهش طراحی جامع برای دادگان ملی می‌باشد و صرفاً به مفهوم جمع‌آوری و یا ذخیره‌سازی اطلاعات نمی‌باشد؛ بلکه سناریوهای راهبردی، توسعه، دسترس‌پذیری، ارتقا زیرساخت و نظایر آن را به عنوان اهداف فرعی در نظر دارد. مهم‌ترین مؤلفه‌های برنامه موفقیت حاکمیت داده براساس الگوریتم ژنتیک مشخص شد و شاخص‌های کلیدی بر اساس مدل دلفی احصاء شدند و شناسایی مؤلفه‌های حاکمیت داده در ساماندهی دادگان ملی با استفاده از تحلیل عاملی اکتشافی و روش متدولوژی سیستم‌های نرم، استخراج مؤلفه‌های دادگان ملی در راستای توسعه حاکمیت داده‌ها صورت گرفت. در نهایت از روش واریماکس برای ماتریس همبستگی و تأثیر مؤلفه‌ها بر همدیگر استفاده شد و با استفاده از الگوریتم ژنتیک، بهینه‌سازی لازم صورت گرفت و با تحلیل عاملی مرتبه اول و دوم، مدل نهایی شد. با استفاده از روش کتابخانه‌ای و تحلیل متون لیستی از مؤلفه‌ها و شاخص‌ها استخراج شد. سپس با اعمال نظر ۲۲ خبره و متخصص، ۱۰ مؤلفه اساسی نهایی شدند. برای تحلیل داده‌ها از نرم افزار smartPLS استفاده شده است و جامعه آماری پرسشنامه مدیران فناوری اطلاعات، اساتید دانشگاه، صاحبان کسب‌وکارها بوده‌اند. روش ترکیبی اکتشافی، مدل پیشنهادی و شاخص‌های استخراج شده مبتنی بر مدل مفهومی، مبنای ارائه چارچوب دادگان ملی برای توسعه حاکمیت داده می‌باشد. سپس شناسایی شاخص‌های احصاء شده در فراترکیب در پنج گروه شاخص‌های ورودی، خروجی، نتیجه، فرآیند و تأثیر دلفی خلاصه شدند.

**واژگان کلیدی:** حاکمیت داده، خزانه داده، دادگان ملی، راهبری داده، نظام داده و اطلاعات، مدیریت داده.

## ۱. مقدمه

دادگان ملی، یعنی مجموعه از داده‌های منتج شده از تعامل سیستمی در کشور که اندازه آنها فراتر از حدی است که با نرم افزارها یا روشهای معمول بتوان آنها را در یک زمان قابل قبول؛ دریافت، ذخیره، مدیریت و پردازش کرد [۱ و ۲]. مهمترین مولفه دادگان ملی حجم می باشد و دلیل این امر، تولید مداوم داده توسط ابزارها و سیستمهای دیجیتال است. نکته بعدی در ساختار است که دادگان ملی بر خلاف نگاه سنتی؛ غیرمتمرکز و توزیع شده است و ساختار منسجمی ندارد (بدون ساختار/ نیمه ساختار) [۳] از سوی دیگر مدل داده‌ها در داده‌های سنتی، مدلی ثابت است در حالی که داده ملی، مدل مشخص و ثابتی ندارند. ارتباطات داخلی بین داده‌های سنتی پیچیده است در حالی که در دادگان ملی ارتباط داخلی پیچیده‌ای بین داده‌ها وجود ندارد. فرآیند جمع آوری دادگان ملی، در ابتدا با هدف ثابتی نیست. مقدار زیادی داده جمع آوری می‌شود و این که ارزش یا پتانسیل استفاده از داده چقدر است بعد از جمع آوری مشخص می‌گردد. به همین دلیل گفته می‌شود که داده‌ها برای اهدافی جمع آوری شده‌اند که از قبل بطور دقیق مشخص نشده‌اند [۴ و ۵] دادگان ملی به این دلیل که صرفاً حجم زیادی از اطلاعات را در اختیار دارند اهمیت چندانی ندارند؛ بلکه این نحوه ذخیره‌سازی، بازیافت، استفاده و نحوه بهره‌برداری از این اطلاعات است که اهمیت دارد. پژوهش‌های متعدد، کمتر به موضوع سناریوهای یکپارچه مدیریت داده‌های ملی پرداخته شده است و نگاه در طرح مدل به سمت داده و اطلاعات یک سازمان متصور شده است. نکته دیگر توجه به این موضوع است که در اکثر مطالعات معماری ذخیره سازی داده‌های با ابعاد متوسط مبنا قرار گرفته شده است.

در حالی که امروزه گفتمان «داده به عنوان دارایی» به خوبی پذیرفته شده است، سازمان‌ها، از جمله بخش دولتی، اغلب در مدیریت و ارزش گذاری داده‌ها به مانند سایر دارایی‌های مرتبط با موفقیت آنها ناکام هستند. این امر امکان استفاده از فرصتهای ایجاد شده توسط «داده‌های گسترده‌ای از اطلاعات که قبلاً غیرممکن بود» را تضعیف

می‌کند. سازمان‌های دولتی و خصوصی اغلب با چالش‌هایی نظیر میراثی که از مدل‌های تجاری آنالوگ به ارث رسیده‌اند، زیرساخت‌های داده‌های منسوخ شده تا شکاف مهارت، موانع نظارتی، عدم رهبری و پاسخگویی و و فرهنگ سازمانی مستعد ابتکار و نوآوری دیجیتالی و تغییر روبرو هستند [۶]. چالش‌های جدیدی نیز ناشی از سوءاستفاده از داده‌های شهروندان و موارد سوءاستفاده، بطور عمده توسط سازمان‌های بخش خصوصی است. این امر با ناتوانی دولت‌ها در انجام اقدامات پیشگیرانه، همگام با تحولات فناوری و درک پیامدهای سیاسی داده‌ها از نظر اعتماد و حقوق اساسی یکی شده است [۷]. بنابراین نقش و جایگاه حاکمیت داده نه تنها در شفافیت دادگان، بلکه در توسعه و توزیع بهینه آن می‌تواند موثر باشد. حاکمیت داده (حکمرانی داده) [۲].

مهمترین جنبه‌های مدل‌سازی از دادگان، راهبرد حاکمیت صحیح آن با اهداف متعدد از جمله توسعه کسب و کار همراه است. طی همه سالها با مفاهیمی همچون بانک اطلاعاتی، دریاچه داده، بازارگاه داده مواجه بوده ایم که در سناریوی یکپارچگی دادگان و مدیریت یکنوا در سطح ملی پاسخگو نمی‌باشد زیرا در کنار تبعات سیاسی و مالکیت دادگان، مشکلات دوجندان در مدیریت و اصالت پذیری دادگان را مواجه هستیم. بنابراین طرح خزانه داده ملی می‌تواند راهکاری مطمئن برای برون رفت از مکانیزمهای اشتباه در حوزه دادگان ملی باشد. خزانه داده‌ها<sup>۳</sup> یعنی دسترسی به داده‌های بسیارزیادی که طی سالیان متمادی در ارتباط با موضوعات متعدد گردآوری شده است تا در فرایند پرس‌وجو، تلفیق یا ترکیب اطلاعات به تصمیم‌گیری ویا در شرایط تصمیم‌سازی به آنها کمک نماید. مخزن داده همچنین به عنوان یک کتابخانه داده یا بایگانی داده‌ها شناخته می‌شود [۸]. خزانه داده یک اصطلاح کلی برای مراجعه به مجموعه داده‌های جدا شده جهت استخراج برای گزارش و تجزیه و تحلیل داده‌ها است. خزانه داده داده یک زیرساخت بزرگ دیتابیس است که مجموعه داده‌ها را برای تجزیه و تحلیل داده، به اشتراک گذاری و گزارش جمع آوری، مدیریت و ذخیره می‌کند و یک اصطلاح

## ۲. مرور پیشینه پژوهش

### الف - مطالعات حوزه حاکمیت داده در ایران

با توجه به اهمیت توجه به دادگان ملی و تسری موضوعات حاکمیت داده، انتظار می‌رود این موضوع در پژوهش‌های حوزه مدیریت کشور مورد توجه قرار گرفته باشد؛ اما مروری بر متون مدیریت دادگان در ایران نشان می‌دهد که علیرغم اهمیت روزافزونی که مدل و معماری سازماندهی و استفاده از دادگان ملی در کشور دارد ولی کمتر به این موضوع پرداخته شده است. بخش زیادی از پژوهش‌های سالهای اخیر به موضوع حاکمیت داده در سازمانها و بخش‌های کاری از قبیل بانک و بیمه پرداخته و پیشنهادات حاصل شده از تحقیقات قابلیت انشار به کل کشور به عنوان یک اکوسیستم کامل را نداشته است از جمله این پژوهش‌ها می‌توان به مطالعه حاکمیت داده در حوزه سلامت [۱۲] و حاکمیت داده محور با رویکرد دولت الکترونیکی [۱۳]، مدلسازی داده‌های بزرگ [۱۴]، بررسی جوانب انبارداده [۱۵]، بهبود کارایی در انبار داده [۱۶]، حاکمیت داده در صنعت بیمه [۱۷]، حاکمیت داده در صنعت بانکی [۱۸]، چارچوب‌های حاکمیت داده [۱۹] اشاره نمود. تعدادی از پژوهش‌ها نیازمندی و مولفه‌های کلیدی موردنیاز از جمله کیفیت داده و استاندارد سازی و کدینگ را هدف قرار دادند از جمله می‌توان به مطالعاتی نظیر مدیریت کیفیت داده، سازماندهی و پاسخگویی [۲۰]، الزامات ارزیابی بلوغ مدیریت داده [۲۱]، خوشه بندی داده‌ها با هدف کاهش زمان جستجو [۲۲]، بررسی امنیت در انبار داده‌ها [۲۳]، بهبود عملکردی امنیت یکپارچگی پایگاه داده‌های تجاری [۲۴]، تهدیدات و تمهیدات پایگاه داده [۲۵] می‌توان اشاره داشت. از زمره مطالعات اندکی که در ارتباط با موضوع می‌توان اشاره داشت، مطالعه ارائه یک روش جدید برای ذخیره سازی و دسترسی به داده‌ها در پایگاه داده‌های بزرگ توزیع شده مبتنی بر استخراج، پالایش و بارگذاری اطلاعات [۲۶] و طراحی چارچوب حاکمیت داده در مراکز تبادل داده و

تا حدودی کلی است که برای مراجعه به مقصد تعیین شده برای ذخیره داده‌ها استفاده می‌شود.

"خزانه داده" شامل ایجاد و نگهداری یک مجموعه داده‌ای و فرآیند اکتساب اطلاعات سودمند از داده‌های ذخیره شده است؛ به عبارت ساده‌تر، روشی است برای فرآیند جمع‌آوری داده‌ها (از انواع مختلف که برای سازمان داخلی یا خارجی هستند)، ثبت آن در یک قالب خاص، و ایجاد امکان بازرسی داده‌ها برای کاربران نهایی سازمان به روش‌های مختلف به منظور ارزش‌سنجی داده‌ها صورت می‌گیرد [۹]. مخازن داده با بهره‌گیری از معماری، متدولوژی، و فناوری ویژه یک سازمان را قادر می‌سازد که داده‌های خود را به نحوی بهینه بکار گرفته تا بار کاری بر روی سیستم‌های پایگاه داده اصلی به کمترین مقدار برسد. یک خزانه داده، داده‌های خود را از یک پایگاه اصلی و عملیاتی گرفته و با انجام پردازش‌هایی بر روی آن‌ها به گونه‌ای آن‌ها را ذخیره می‌کند که این داده‌ها به فرمت و اشکال بسیار موثرتری قابل دریافت باشند [۶]. خزانه داده ملی<sup>۴</sup> همانند یک بانک اطلاعات است که به دنبال حفظ و ترویج داده‌های مربوط به منابع یک کشور است. یک خزانه ملی داده معمولاً توسط نهادهای مستقر، با رویکرد کنترل و پشتیبانی از تبادل، ضبط، انتقال و توزیع اطلاعات تأسیس می‌شود، و هدف نهایی برای ارائه ابزار و اطلاعات به کشور برای اطمینان از رشد، توانایی دولت، کنترل، استقلال و حاکمیت اطلاعات است [۱۰]. دو دلیل اساسی هر کشور برای ایجاد مخزن داده‌های ملی، حفظ داده‌های تولید شده در داخل کشور توسط صنعت و ترویج سرمایه‌گذاری در کشور با استفاده از داده‌ها برای کاهش خطرات اکتشاف، تولید و حمل و نقل است. برخی از مخازن داده‌های ملی کاملاً توسط دولت یک کشور تأمین می‌شوند و برخی از آنها سیستم‌های ترکیبی هستند که بخشی از آن توسط صنعت و دولت تأمین می‌شوند. مخزن داده‌های ملی معمولاً هزینه‌های درخواست داده و بارگیری داده را می‌پردازند [۱۱]. یک مخزن داده باید دارای ۴ ویژگی باشد که عبارت است از: موضوع گرا، یکپارچه، غیرفرار و متغیر با زمان است

۴. National Data Repository (NDR)

۳. مدل مفهومی مبتنی بر شاخص های مورد انتظار و میزان

همبستگی شاخص ها برای داده گان ملی چیست؟

۴. بازیگران اصلی و ارتباط بین بازیگران چگونه است؟

۵. معماری نگهداری و مدیریت داده گان ملی چگونه است؟

۶. پارامترهای حفظ داده گان ملی و تحقق حداکثری امنیت

داده گان چیست؟

۷. مدل بهینه برای دادگان ملی باهدف تحقق حاکمیت در بخش

فناوری اطلاعات چگونه است؟

#### ۴. روش پژوهش

رویکرد پژوهش حاضر کیفی است از روش فراترکیب جهت ارزیابی منابع علمی استفاده شده است فراترکیب به عنوان یکی از حوزه های فرامطالعه<sup>۵</sup>، رویکردی نظام مند برای ترکیب پژوهش های کیفی مختلف در راستای کشف زمینه های فرعی و اصلی است که موجب ارتقای دانش جدید شده و دید جامعی از حوزه مورد بررسی به وجود می آورد. این روش مستلزم آن است که پژوهشگر بازنگری دقیق و عمیقی پیرامون موضوع پژوهش انجام دهد و یافته های پژوهش های کیفی مرتبط را ترکیب و از این طریق نمایش جامعی از پدیده های مورد بررسی را نشان دهد. به عبارت دیگر استفاده از فراترکیب خاصیت هم افزایی دارد و نتیجه را به دست می دهد که بزرگتر از مجموعه بخش هایش است. به طور معمول بررسی مقاله ها و مستندات علمی با توجه به عمق مطالعه در این رویکرد کمتر از ۷۰ مورد است [۳۰ و ۳۱].

این روش به طور ویژه در علوم که مطالعات آن بیشتر مبتنی بر تحلیل کیفی است و مبانی نظری وسیعی را شامل نمی شوند کاربرد دارد از آنجا که مطالعات مرتبط با مفهوم خزانه داده در حال حاضر بیشتر کیفی است و تا زمان انجام پژوهش حاضرمدل جامعی که ابعاد مختلف خزانه داده ملی را در برگیرد یافت نشده است؛ روش فراترکیب به عنوان روش مناسبی برای شناسایی و به دست آوردن ترکیب جامعی از مولفه ها و ابعاد اصلی تحقیق بکار گرفته شده است. از این رو برای بهره گیری از روش فراترکیب در این پژوهش، از

اطلاعات ملی [۲۷] می باشد. بنابراین پژوهشگر نتوانست پژوهشی با هدف یافتن مدل مرجعی برای حاکمیت دادگان بزرگ و یا مدلسازی راهبردی دادگان ملی بیابد.

#### ب - مطالعات حوزه خزانه داده

نوکلای<sup>۵</sup> و همکاران (۲۰۱۹) در پژوهشی با عنوان «حاکمیت داده ها در سیستم عامل های دیجیتال» یک چارچوب مقدماتی مبتنی بر ادبیات برای مدیریت داده های پلتفرم با پنج حوزه کیفیت داده های اصلی، مالکیت و دسترسی؛ مباحثت کیفیت داده های پلت فرم و ارزش استفاده از داده ها پیشنهاد دادند [۲۸]. الحسن<sup>۶</sup> و همکاران (۲۰۱۹) پژوهشی با عنوان «عوامل موفقیت بحرانی برای مدیریت داده ها انجام دادند. هدف از این تحقیق، ارائه فاکتورهای مهم موفقیت برای حاکمیت داده ها است. این مقاله به یک مطالعه موردی واحد می پردازد که داده ها از طریق مصاحبه های نیمه ساختار یافته به دنبال رویکرد فاکتورهای مهم موفقیت جمع آوری شده و با استفاده از تکنیک های کدگذاری باز، محوری و انتخابی مورد تجزیه و تحلیل قرار می گیرد. این یافته ها به عنوان هفت فاکتور مهم موفقیت ارائه شده است که به ترتیب از نظر اهمیت رتبه بندی می شوند. علاوه بر این، نیاز به درک بهتر روابط (بهم پیوسته) بین فاکتورهای مهم موفقیت را برجسته شده است [۲۹].

#### ۳. هدف و پرسش های پژوهش

هدف پژوهش حاضر تعیین شاخص های کلیدی راهبری دادگان ملی با هدف ایجاد چارچوبی برای دادگان ملی با هدف توسعه حاکمیت داده است. در این راستا پژوهش حاضر درصدد است که به سوالات زیر پاسخ دهد.

۱. مولفه های کلیدی و تاثیرگذار حاکمیت داده برای داده گان ملی چیست؟
۲. ناحیه دانشی حاکمیت داده برای داده گان ملی دارای چه ویژگی هایی می باشد؟

از حذف موارد تکراری و کمتر وابسته، ۴۷ مقاله برای بررسی بیشتر باقی ماند و پس از بررسی نتایج و روش مدلسازی و چکیده مقالات، ۳۱ مقاله به عنوان مقالات مرتبط با طراحی مولفه های کلیدی مدنظر قرار گرفت. در ادامه نیز با استفاده از روش طوفان فکر با متخصصان و تطابق با پارامترهای استخراج شده از صاحب‌ها، مولفه‌ها نهایی گردید.

### گام دوم: ترسیم ناحیه دانشی چارچوب حاکمیت داده

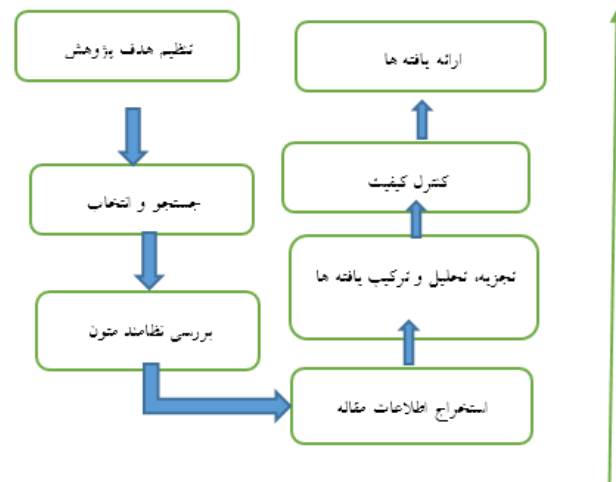
نیازمند توجه به مولفه های راهبردی حاکمیت داده در حوزه راهبری دادگان می باشیم و برای ترسیم ناحیه دانشی حاکمیت داده که در واقع پاسخ به سوال دوم پژوهش می باشد. جستجوی مقالات در پایگاه های اطلاعاتی WOS, Scopus, Science Direct, ProQuest با کلید واژه های Models, schema به همراه عبارات Data trust, Data Governance, Data Security, data Managements، با اعمال محدودیت های زبان انگلیسی انجام گرفت. در این پژوهش، کلیه مطالعات تحت پوشش پایگاه ها که تا آخر خرداد سال ۱۳۹۹ منتشر شده بودند، در نظر گرفته شد. در مجموع ۸۲ مقاله مرتبط بررسی گردید و با استخراج مولفه های دانشی حاکمیت، شاخص های مورد نیاز احصا شد. ناحیه دانشی بر اساس مطالعات پژوهشگران پیشین بر سه محور استانداردسازی، مدیریت و انتشار، شفافیت پیش بینی گردید. مراحل جمع آوری اطلاعات با توجه به استفاده از الگوی پژوهش به شرح ذیل انجام شد:

مرحله اول: سوالات تحقیق مشخص شد و با توجه به سؤالیهای "مولفه های کلیدی و تاثیرگذار حاکمیت داده برای دادگان ملی چیست؟" و "ناحیه دانشی حاکمیت داده برای دادگان ملی دارای چه ویژگی هایی می باشد؟" به یک چارچوب اولیه از مولفه ها و ناحیه کاری مورد نظر رسیدیم.

مرحله دوم: مقالات مناسب برای احصاء شاخص ها شناسایی شد؛

مرحله سوم: اطلاعات مدنظر برای شناسایی زیر شاخص ها انجام پذیرفت؛

روش هفت مرحله ای سندلوسکی و بارسو<sup>۸</sup> (۲۰۰۷) استفاده شده است که در شکل (۱) نشان داده شده است.



شکل ۱. مراحل هفت گانه فراترکیب [۳۲]

روش فراترکیب به عنوان روشی کامل شناخته می شود، زیرا از فنون، روش ها و ابزارهای مختلفی جهت شناسایی پدیده موردنظر استفاده می کند و پایش های مختلفی در بین تعداد زیادی از منابع مورد استفاده قرار می گیرد. اعتبار این روش نیز در خود آن تأیید می شود. جامعه آماری پژوهش شامل کلیه آثار علمی مربوط به مبانی و ابعاد خزانه داده شامل ۱۸۷ مقاله مرتبط بین سالهای ۲۰۱۰ تا ۲۰۲۰ میلادی برای منابع انگلیسی و ۱۳۸۸ تا ۱۳۹۹ خورشیدی برای منابع فارسی است که از این تعداد و پس از طی مراحل غربالگری تعداد ۴۵ اثر برای استخراج مدنظر قرار گرفت.

### گام اول: تعیین مؤلفه های کلیدی مدل

برای پاسخ به سوال اول پژوهش، جستجوی مقالات در پایگاه های اطلاعاتی WOS, Scopus, Science Direct, ProQuest با کلید واژه های Priority, Structure, Models, Data Governance, Data Aggregation, Data Security, Distributed Networks, Security Managements با اعمال محدودیت های زبان انگلیسی انجام گرفت. در این مطالعه، کلیه مطالعات تحت پوشش پایگاهها که تا آخر خرداد سال ۱۳۹۹ منتشر شده بودند، در نظر گرفته شد. پس

۸. Sandelowski & Barroso

## ۵-۲- ناحیه دانشی حاکمیت داده برای داده گان ملی

### دارای چه ویژگی هایی می باشد؟

از مجموع بررسی ۸۲ مقالات حوزه حاکمیت دانش با تمرکز بر مهمترین ظرفیت های آن بر دادگان ملی، جدول زیر (جدول ۲): پارامترهای احصاء شده برای ناحیه دانشی حاکمیت داده) استخراج گردید. مهمترین دلیل احصاء ناحیه دانشی توجه به مهمترین ویژگیهای مورد انتظار در راهبری و سازماندهی دادگان ملی می باشد.

جدول ۲. ویژگی های احصاء شده برای دادگان

ردیف	موضوع	منبع
۱	استفاده از محیط مناسب فرهنگی، سیاسی، قانونی، نظارتی، نهادی، سازمانی و فنی	[۳۳]؛ [۱]
۲	نظام سیاست گذاری منسجم در حوزه داده	[۳۴]؛ [۱]
۳	مدل های نگهداری و دسترسی به دادگان	[۳۵]؛ [۳۶]؛ [۳۴]؛ [۳۷]؛ [۳۸]
۴	محرماتگی، امنیت، دسترس پذیری در نحوه ذخیره سازی	[۳۴]؛ [۳۶]؛ [۳۹]؛ [۲۸]
۵	تضمین مدیریت منابع داده ای	[۴۰]؛ [۴۱]
۶	یکپارچگی، قابلیت استفاده	[۳۴]
۷	نظارت بر داده گان، شیوه های بهینه در سازماندهی	[۳۷]؛ [۴۲]
۸	اطمینان از کیفیت و اطلاعات	[۴۳]؛ [۴۴]؛ [۲۸]

مرحله چهارم: مدل مفهومی اولیه متناسب با مولفه ها، شاخص ها و زیر شاخص ها تبیین شد؛

مرحله پنجم: مولفه ها با بررسی و تحلیل خبرگانی نهایی شد؛

مرحله ششم: به منظور تطبیق و نهایی سازی زیر شاخص ها مصاحبه با ۲۰ خبره در حوزه مدیریت داده صورت پذیرفت؛

مرحله هفتم: شاخص ها و زیر شاخص ها احصاء شده در دو گروه "مدیریت و راهبری" و "فنی، حقوقی" بصورت پرسشنامه منتشر گردید؛

مرحله هشتم: داده های گردآوری شده از پرسشنامه تحلیل و واکاوی آماری شد؛

مرحله نهم: مدل ارزیابی و نهایی گردید؛

## ۵. یافته ها

### ۵-۱- مولفه های کلیدی و تاثیرگذار حاکمیت داده

#### برای داده گان ملی چیست؟

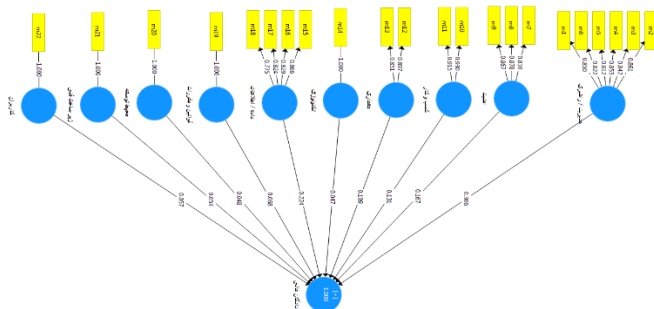
پس از بررسی مطالعاتی که به حاکمیت داده پرداخته اند و مصاحبه با صاحب نظران، ده مولفه اصلی طبقه بندی شده اند (جدول ۱). حروف بیانگر کد تخصیص داده شده به هر مولفه می باشند. همچنین طی مصاحبه ها هدفگذاری تبیین برخی شاخص های کلیدی برای هر مولفه مدنظر قرار گرفت و منطبق با مطالعات بالادستی، لیستی از این شاخص ها احصاء گردید.

جدول ۱. پارامترهای کلیدی

مدیریت راهبری	امینت	کسب و کار	معماری	تکنولوژی	داده/اطلاعات	قوانین	محیط توسعه	زیر ساخت	کاربران
M	S	B	A	T	D	R	E	I	U

۳	کسب و کار	B	مدیریت دانش / کسب و کار
۴	معماری	A	معماری داده، اطلاعات / متدلوژی توسعه
۵	تکنولوژی	T	پلت فرم های نرم افزاری
۶	داده / اطلاعات	DI	داده استاندارد / داده باز/ مدیریت داده / فراداده
۷	قوانین و مقررات	R	قوانین و مقررات
۸	محیط توسعه	ED	مهارت و دانش
۹	زیرساخت فنی	I	زیر ساخت فنی، ارتباطی
۱۰	کاربران	U	کاربران بالقوه

برای تحلیل جدول فوق، پرسشنامه لازم تهیه و به طرق مختلف منتشر گردید. تحلیل ضریب مسیر به عنوان یکی از تحلیل های این جدول مدنظر قرار گرفت. تحلیل ضریب مسیر، بیان کننده وجود رابطه علی خطی و شدت و جهت این رابطه بین متغیرهاست. در حقیقت همان ضریب رگرسیون در حالت استاندارد است که در مدل های ساده تر رگرسیون ساده و چندگانه مشاهده می شود. عددی بین ۱- تا +۱ است که اگر برابر با صفر شوند، نشان دهنده ی نبود رابطه ی علی خطی بین دو متغیر پنهان است. با توجه به شاخص های احصا شده در این قسمت، نسبت به بررسی و تحلیل عاملی مرتبه دوم مولفه های مورد مطالعه، به شرح زیر اقدام می شود:



۹	مستند سازی و تهیه نسخه پشتیبان از داده	[۴۵]
۱۰	ضرورت وجود مدل و معماری مطلوب	[۴۰]؛ [۴۶]؛ [۴۱]

ویژگی ها با بسامد تکرار بالاتر و تمرکز بر راهبری دادگان با توجه به احصا و کد گذاری نظر خبرگانی به صورت شکل زیر جمع بندی گردید. بنابراین ناحیه دانشی برای دادگان ملی به صورت شکل (۲) ترسیم گردید.



شکل ۲. ناحیه دانشی احصا شده

### ۵-۳- مدل مفهومی منطبق بر شاخص های مورد انتظار

#### و میزان همبستگی شاخص ها مبتنی بر

#### حاکمیت داده برای داده گان ملی چیست؟

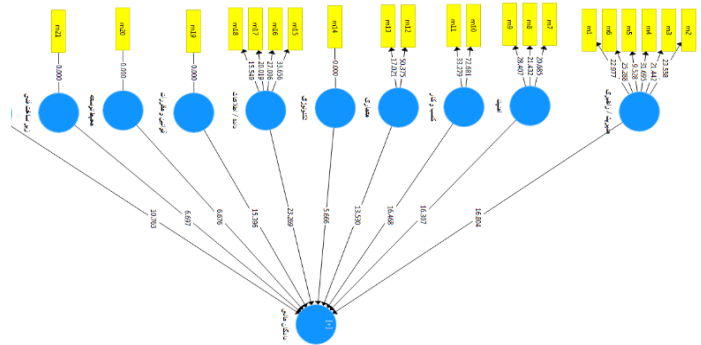
با استخراج مصاحبه های انجام شده و تطبیق شاخص های احصا شده با مقالات نمونه ( بررسی ۸۲ مقاله مرتبط تا تاریخ خرداد ۱۳۹۹ مدنظر قرار گرفته است) جدول زیر (جدول ۳) به عنوان جدول شاخص های راهبری دادگان ملی ترسیم گردید.

#### جدول ۳. جمع بندی شاخص ها راهبری دادگان ملی

ردیف	مولفه	نماد	شاخص
۱	مدیریت / راهبری	M	منابع / پشتیبانی بهینه خدمات / حریم شخصی کاربران / توسعه سرزمینی / ساختار / برنامه راهبردی
۲	امنیت	S	دسترس پذیری / احراز هویت / مالکیت معنوی

شکل ۳. تحلیل ضریب بتا

با توجه به نتایج حاصل شده در این قسمت مشاهده شد که بیشترین علیت را مدیریت/راهبری با ضریب ۰,۳ دارد. همچنین در ادامه نیز کاربران با ضریب ۰,۰۵، زیر ساخت فنی با ضریب ۰,۰۵، محیط توسعه با ضریب ۰,۰۴، قوانین و مقررات با ضریب ۰,۰۶، داده و اطلاعات با ضریب ۰,۲، تکنولوژی با ضریب ۰,۰۴، معماری با ضریب ۰,۱ و کسب و کار با ضریب ۰,۱ و امنیت با ضریب ۰,۱ از علیت های متفاوتی برخوردار هستند.



شکل ۴. برازش مدل در مقطع T-value

با توجه به نتایج حاصل شده در این قسمت مشاهده شد که بیشترین علیت را مدیریت/راهبری با ضریب ۱۶,۸۴ دارد. همچنین در ادامه نیز کاربران با ضریب ۱۰,۷۶، زیر ساخت فنی با ضریب ۶,۶۹، محیط توسعه با ضریب ۶,۶۷، قوانین و مقررات با ضریب ۱۵,۳۹، داده و اطلاعات با ضریب ۲۳,۲۶، تکنولوژی با ضریب ۵,۶۶، معماری با ضریب ۱۳,۵۳ و کسب و کار با ضریب ۱۶,۴۴ امنیت با ضریب ۱۶,۳۰ از علیت های متفاوتی برخوردار هستند.

همچنین با تحلیل برای پایایی و روایی مولفه ها مشاهده شد که تمامی مولفه های مورد مطالعه در سطح معناداری کمتر از ۵ درصد، ضریب تی ویو بیشتر از ۱,۹۶ و همچنین ضرایب بتای قابل قبولی برخوردار هستند. شاخص آلفای کرونباخ برای تمامی مولفه های مورد مطالعه بیشتر از ۰,۷ می باشد، شاخص CR نیز در مقدار عددی قابل قبولی است.

روایی واگرایی قابل قبول یک مدل، حاکی از آن است که یک سازه در مدل تعامل بیشتری با شاخص های خود دارد تا با سازه های دیگر. روایی واگرایی وقتی در سطح قابل قبول است که AVE برای هر سازه بیشتر از واریانس اشتراکی بین آن سازه و سازه های

جدول ۴. بررسی پایایی مولفه های مدل شده

	Cronbach's Al...	rho_A	Composite Rel...	Average Varian...
امنیت	0.811	0.812	0.888	0.726
تکنولوژی	1.000	1.000	1.000	1.000
داده / اطلاعات	0.842	0.845	0.894	0.679
داندگان عالی	0.954	0.958	0.959	0.517
زیر ساخت فنی	1.000	1.000	1.000	1.000
قوانین و مقررات	1.000	1.000	1.000	1.000
محیط توسعه	1.000	1.000	1.000	1.000
مدیریت / راهبری	0.893	0.902	0.920	0.659
معماری	0.845	0.708	0.867	0.765
کاربران	1.000	1.000	1.000	1.000
کسب و کار	0.825	0.831	0.919	0.851

دیگر(یعنی مربع مقدار ضرایب همبستگی بین سازه ها) در مدل باشد.

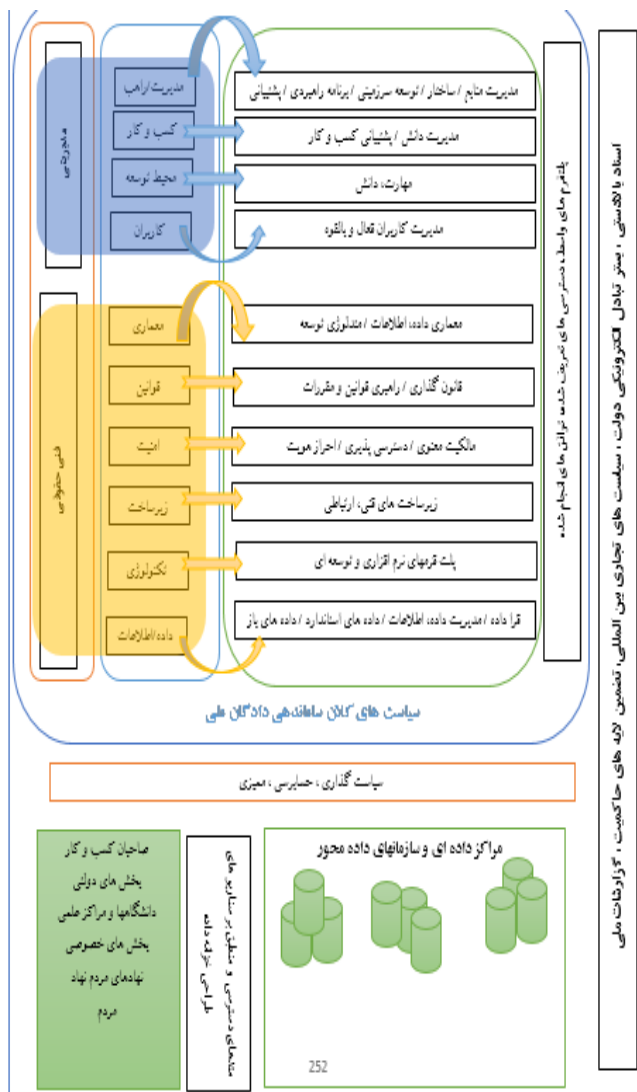
جدول ۵. بررسی روایی واگرایی توسط ماتریس فورنر لاکر

	امنیت	تکنولوژی	داده / اطلاعات	داندگان عالی	زیر ساخت فنی	قوانین و مقررات	محیط توسعه	مدیریت / راهبری	معماری	کاربران	کسب و کار
امنیت	0.852										
تکنولوژی	0.495	1.000									
داده / اطلاعات	0.774	0.452	0.824								
داندگان عالی	0.872	0.538	0.937	0.719							
زیر ساخت فنی	0.426	0.234	0.575	0.602	1.000						
قوانین و مقررات	0.630	0.404	0.758	0.771	0.478	1.000					
محیط توسعه	0.445	0.237	0.532	0.541	0.431	0.448	1.000				
مدیریت / راهبری	0.702	0.390	0.755	0.881	0.473	0.613	0.388	0.811			
معماری	0.708	0.512	0.747	0.816	0.488	0.570	0.347	0.631	0.874		
کاربران	0.433	0.333	0.636	0.643	0.515	0.489	0.446	0.481	0.532	1.000	
کسب و کار	0.802	0.477	0.820	0.875	0.442	0.651	0.413	0.685	0.710	0.527	0.922



با توجه به نهایی شدن مولفه‌ها و شاخص‌های مورد انتظار می‌توان مدل مفهومی برای دادگان ملی را ترسیم نمود. کسب و کار و یا هر مکانیزم خدمت، نیازمند داده است و اساساً داده و اطلاعات عامل هویت ابزار هستند. بنابراین در مورد دادگان ملی علاوه بر پارادایم‌های مدیریت دادگان و تضمین امنیت و کیفیت، نیازمند ترسیم یک اکوسیستم جامع در تحقق حاکمیت با رویکرد سیاستگذاری و نظارت هستیم. این ادبیات در قالب ساختار یکپارچه و توزیع یافته بلوغ می‌یابد. خزانه داده به مفهوم ایجاد یک ساختار یکسان و یکباره برای دادگان یک کشور نیست و نباید تلقی شود که به دنبال یک انبار داده بزرگ در ابعاد ملی باشیم. بنابراین با توجه به شاخص‌های احصا شده که پیشتر به آن پرداخته شد نیازمند طراحی یک مدل مفهومی اولیه برای دادگان ملی هستیم. این مدل مفهومی ارتباط بین سازماندهی دادگان ملی و شاخص‌های احصا شده را نشان می‌دهد و می‌تواند به ایه‌های تصمیم‌گیری کمک به طراحی مدل اجرایی و تجاری نماید.

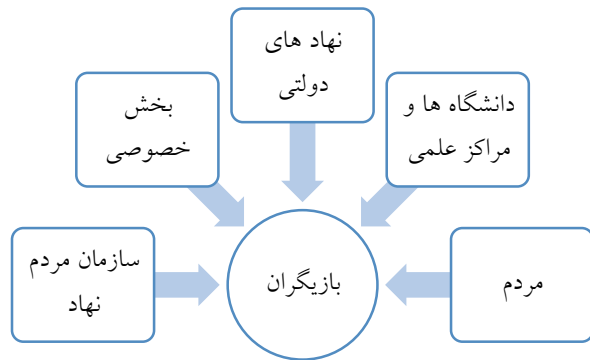
مباحث حاکمیت داده دارای نقاط مشترک زیادی با حوزه فناوری اطلاعات می‌باشد از جمله در زمینه‌های داده‌های بزرگ، هوش تجاری، سیستم عامل منابع سازمانی، مدیریت چرخه محصول، مدیریت اطلاعات تولید محصول، مهارت مدیریت داده، مهارت حاکمیت داده. بنابراین می‌توان انتظار داشت بسیاری از شاخص‌ها و رویکرد طراحی مدل مطلوب دارای نقاط اشتراک واضحی با طرح‌های فناوری اطلاعات داشته باشد. زیرا غایت اجرایی رسیدن به مدلی است که قابل ممیزی و نظارت فنی نیز باشد. در شکل (۵) مدل مفهومی پژوهش ترسیم شده است.



شکل ۵. مدل مفهومی نگهداری دادگان ملی (خزانه داده)

## ۵-۴- بازیگران اصلی و ارتباط بین بازیگران چگونه

است؟



شکل ۶. ماهیت بازیگران مدل

## ۵-۵- معماری نگهداری و مدیریت داده گان ملی

چگونه است؟

پیچیدگی فرآیند توسعه مخزن (خزانه) داده برای موفقیت نیاز به پیروی از یک روش شناختی دارد. یک رویکرد گسترده پذیرفته شده برای توسعه، روش ترکیبی است که در آن الزامات و منابع داده باید در مدل جدید مخزن (خزانه) داده جای بگیرند (مته و تروجلی، ۲۰۱۲). از سوی دیگر برای راهبری داده در یک سازمان نوین داده محور، نیاز به معماری یکپارچه برای ذخیره داده در مرحله اول و سپس پردازش آن هست که مباحث مدیریتی و راهبری داده هم در آن لحاظ شده باشد [۴۷]. در طراحی خزانه داده، ترکیبی از هر دو معماری متمرکز و نیمه متمرکز را خواهیم داشت و مهمترین دلیل آن گستردگی و عدم تطابق فرآیندی در حوزه های مختلف می باشد. بنابراین موضوع اول مولفه های مدنظر در معماری و مسئله بعدی شاخص های معماری می باشد. در شکل (۷) مولفه های معماری مطلوب در طراحی خزانه داده به تصویر کشیده شده است. توسعه مخزن (خزانه) داده یک کار مداوم است که در آن دائماً الزامات جدید اضافه می شود بنابراین نیازمند یک معماری مدل محور هستیم. برای اینکه بتوان شاهد حکمرانی داده در تمامی واحدهای اداری بود یعنی تمامی داده های سازمانی را به صورت یکپارچه و مدیریت شده و در راستای افزایش بهره وری و تصمیم سازی های کلان به کار برد، نیازمند یک معماری داده

توجه به بازیگران در طراحی مدل، جزئی از اقدامات اصلی در موفقیت مدل می باشد زیرا یکی از مهمترین سوالات در زمینه طراحی، پاسخ به سوال "چه کسی" می باشد. بنابراین این سوال به عنوان یکی از سوالات اصلی پژوهش مطرح می باشد. اقدامات کلان در حوزه دادگان، در هسته اصلی خود به تیم های اجرایی نیاز دارند تا میزان آمادگی آنها را برای اجرای پروژه ارزیابی کنند. این ارزیابی شامل تعیین چگونگی دستیابی به داده و اطلاعات است. در مسیر تعامل ذی نفعان، کیفیت داده ها باید ارزیابی شوند و همچنین بررسی اینکه آیا داده ها می توانند نیازهای تجاری همه سیستم ها را پشتیبانی کنند. رهبری سازمان نیاز به ایجاد ضرورت های تجاری دارد. هر ذی نفع نیاز به مستندسازی و اولویت بندی کارهای کیفیت داده ای دارد که باید تکمیل شوند و چالش های طولانی مدت مانند فقدان استانداردهای مستند، درک ابرداده و کیفیت پایین داده ها باید مورد توجه قرار گیرد (لوشین، ۲۰۱۰). توجه به ذی نفعان که می توانند داخلی و یا خارجی باشند از اهمیت بالایی برخوردار است. ذی نفع داخلی شامل افراد یا گروه های هستند که متولی نگهداری و تضمین کیفیت و استاندارد داده ای می باشند. در مطالعه ۵ گروه ذینفع شناسایی گردید که هم تولید کننده و هم استفاده کننده از دادگان می باشند. همانطور که در شکل زیر (شکل ۶) مشاهده می شود فارغ از حجم تراکنش ها می توان بازیگران اصلی را رصد نمود.

[۵۲]، [۵۶]، [۵۷]	نمای خزانه داده قبل از پیاده سازی، تشخیص مقادیر پیش محاسباتی، طراحی جدول های ابعاد و واقعی، دریافت و سازماندهی	۷
[۵۳]	توجه به چارچوب های امن سازی	۸
[۵۳]، [۵۵]	توجه به معماری همبسته اطلاعات	۹
[۵۸]، [۵۶]	درک و تحلیل درست از نیازمندی ها	۱۰

جدید منطبق با نیازهای عصر کلان داده هستیم. ساختاری که ضمن حفظ استقلال نرم افزارها و داده‌ها در بخش‌های مختلف، به تحلیلگران و مدیران اجازه بررسی و استفاده از تمامی اطلاعات سازمان را به صورت مدیریت شده بدهد و انواع گزارش‌گیری‌ها و تحلیلها را برای آنها میسر سازد [۴۷].

با بررسی ۵۶ مقاله و تفکیک ۲۳ مقاله مجموعه‌ای از مهمترین پارامترهای موثر بر معماری خزانه داده استخراج گردید و پارامترها استخراج شده در مصاحبه‌های تخصصی با صاحب‌نظران به بررسی و چالش قرار داده شد و مجموعه عوامل در جدول پارامترهای موثر بر معماری (جدول ۶) و شکل ۷ (نمودار احصا شده عوامل موثر بر معماری) نمایش داده شده است. اهمیت این سوال در پژوهش به جهت ضرورت پرداختن به چارچوب معماری در کنار شاخص‌های احصا شده دو چندان می‌باشد.

جدول ۶. پارامترهای موثر بر چارچوب معماری

ردیف	موضوع	منبع
۱	امکان دید یکپارچه ، ادغام طرح واره های جزئی	[۴۸]، [۴۹]، [۵۰]
۲	توجه به نیازمندی های کاربران نهایی	[۵۰]، [۵۱]
۳	منبع حمایت و پشتیبانی	[۵۳]، [۵۲]، [۵۰]
۴	سطح مورد نیاز برای مدیریت اطلاعات	[۵۴]، [۵۰]
۵	مسائل فنی	[۵۶]، [۵۵]، [۵۰]
۶	مخزن داده در طول حیات خود، مجموعه‌ای از عملیات را تکرار می‌کند	[۵۲]



شکل ۷: نمودار احصا شده عوامل موثر برای معماری

در ادامه از بعد اجرایی، لازم است تا زیر شاخص‌های لازم برای شاخص معماری به عنوان یکی از مولفه‌های کلیدی طراحی دادگان

شده و پس از ارزیابی و بهینه یابی سه شاخص دسترس پذیری، مالکیت معنوی، احراز هویت طبق جدول (۷) نهایی گردید:

جدول ۷. نمای مولفه و شاخص های اصلی در امنیت

امنیت			مولفه
احراز هویت	مالکیت معنوی	دسترس پذیری	شاخص ها

مبتنی بر هریک از شاخص ها، زیرشاخص های قابل پیش بینی با استناد به مقالات و مصاحبه خبرگانی احصا گردید که در جدول (۸) نشان داده شده است:

جدول ۸. شاخص ها و زیر شاخص های امنیت

مالکیت معنوی	احراز هویت	دسترس پذیری
علائم تجاری	امضاء دیجیتال	دسترس مجاز
حقوق تجاری	استاندارد پذیرش	آدرس دهی
حقوق ثبت داده	الگوریتم نفوذ	کتابخانه داده ای
حق نشر		کاتالوگ داده
حق استفاده		سیستم لاگ
حق داده های		سرویس های امن
محرمانه		احراز هویت
		حسابرسی
		SLA تبادل امن

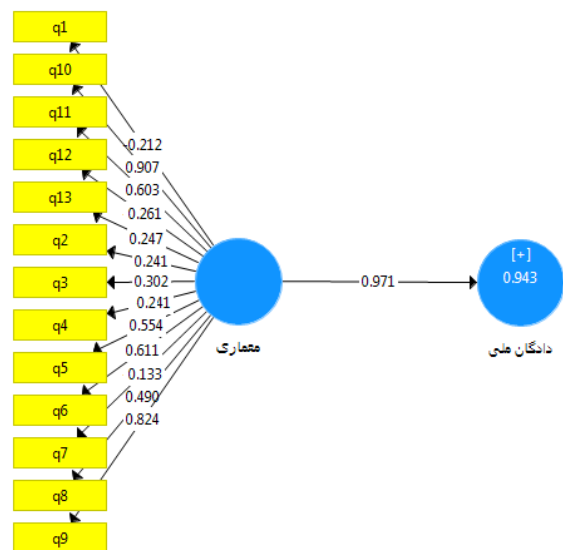
با توجه به داده های جمع آوری شده از پرسشنامه، همبستگی شاخص ها بررسی گردید که در نمودارهای زیر قابل رویت می باشد:



شکل ۹. بررسی رابطه مولفه ها در متغیر احراز هویت

نتایج حاکی از وجود رابطه معنادار علی در مولفه های خرد در تبیین مولفه ی هدف که احراز هویت می باشد، است. در بررسی ها

ملی در نظر گرفته شود. در پروسه مصاحبه با خبرگان و مستند بر کلمات کلیدی استخراج شده، ۱۳ زیرشاخص استاندارد، کدینگ سرویس ها، تعداد SLA، قوانین اجرایی، سرویس دیکشنری، سرویس کاتالوگ، سیستم پیگیری فرآیندها، سیستم پردازش فرآیندها، نرخ تبادل، استاندارد شبکه، استاندارد دسترسی، ساختار دادگان، ظرفیت زیرساخت فنی، احصا شد. و با گردآوری اطلاعات از پرسشنامه و تحلیل لازم وابستگی و همگرایی زیرشاخص های احصا شده تعیین گردید.



شکل ۸. بررسی رابطه مولفه ها در متغیر معماری

با مشاهده شد که زیرشاخص دهم که نرخ تبادل داده بده است در رابطه شدید تری به نسبت سایر مولفه ها با مقدار عددی ۰٫۹۰۷ می باشد.

## ۵-۶- پارامترهای حفظ داده گان ملی و تحقق

### حداکثری امنیت داده گان چیست؟

امنیت یکی از مولفه های اصلی مدل مدنظری باشد و دلیل طرح مجزای آن، اهمیت پایه ای آن در مدل و تحقق حداکثری آن می باشد. شاخص های دسترس پذیری، امنیت داده/اطلاعات، احراز هویت، مالکیت معنوی، حریم شخصی کاربران، پلتفرم، زیرساخت و تبادل برای آن احصا

## ۵-۷- مدل بهینه برای دادگان ملی باهدف تحقق

### حاکمیت در بخش فناوری اطلاعات چگونه

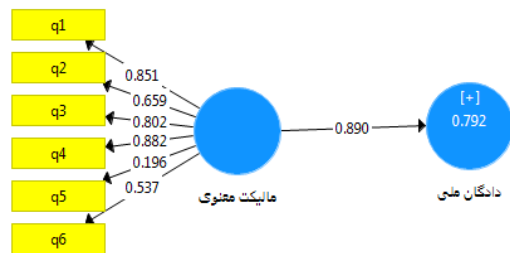
است؟

شاخص‌های برآزش و نکویی محاسبه شده حاکی از مطلوبیت مدل طراحی شده می‌باشد. در مدل نهایی ۱۰ مولفه، ۲۱ شاخص و ۱۵۴ زیرشاخص احصا گردید که در جدول ۹ قابل مشاهده می‌باشد.

جدول ۹. نمای کلی از مولفه‌ها، شاخص‌ها

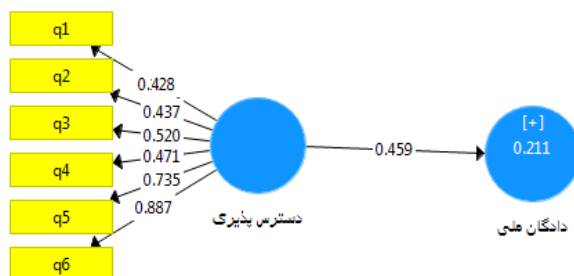
ردیف	مولفه	نماد	شاخص	تعداد زیرشاخص
۱	مدیریت / راهبری	M	منابع / پشتیبانی / بهینه خدمات / حریم شخصی کاربران / توسعه سرزمینی / ساختار / برنامه راهبردی	۳۵
۲	امنیت	S	دسترس پذیری / احراز هویت / مالکیت معنوی	۱۸
۳	کسب و کار	B	مدیریت دانش / کسب و کار	۱۳
۴	معماری	A	معماری / داده، اطلاعات / متدلوژی توسعه	۱۰
۵	تکنولوژی	T	پلت فرم‌های نرم افزاری	۱۱
۶	داده / اطلاعات	DI	داده استاندارد / داده باز / مدیریت داده / فراداده	۳۲
۷	قوانین و مقررات	R	قوانین و مقررات	۶
۸	محیط توسعه	ED	مهارت و دانش	۵

م‌شخص گردید که زیر شاخص الگوریتم نفوذ در رابطه شدیدتری به نسبت سایر مولفه‌ها با مقدار عددی ۰,۹۵۰ می‌باشد.



شکل ۱۰. بررسی رابطه مولفه‌ها در متغیر مالکیت معنوی

با توجه به نتایج حاصل شده در این قسمت مشاهده می‌شود که مولفه‌های مورد مطالعه در ضرایب عددی بیشتر از ۰,۱ می‌باشند که این امر نشان دهنده وجود رابطه معنادار علی در مولفه‌های خرد در تبیین مولفه‌ی هدف که مالکیت معنوی می‌باشد، است و زیر شاخص حق نشر در رابطه شدیدتری نسبت به سایر مولفه‌ها قرار دارد.



شکل ۱۱. بررسی رابطه مولفه‌های سازنده دسترس پذیری

با توجه به نتایج حاصل شده در این قسمت مشاهده می‌شود که مولفه‌های مورد مطالعه در ضرایب عددی بیشتر از ۰,۳ می‌باشند که این امر نشان دهنده وجود رابطه معنادار علی در زیر شاخص‌ها در تبیین مولفه‌ی هدف که دسترس پذیری می‌باشد، وجود دارد. همچنین تحلیل‌ها نشان می‌دهد زیر شاخص سرویس‌های امن، در رابطه شدیدتری به نسبت سایر مولفه‌ها با مقدار عددی ۰,۸۸۷ می‌باشد.

۹	زیرساخت فنی	۱	زیر ساخت فنی، ارتباطی	۱۷
۱۰	کاربران	۱۱	کاربران بالقوه	۷

در مجموع شاخص های احصاء شده، ۱۵۴ زیر شاخص که دارای بیشترین تاثیر در استقرار و پایداری خزانه داده را دارا می باشند، انتخاب شدند. هدف غائی برای دادگان ملی استقرار خزانه داده ای می باشد که بتوان در کنار پارامترهای مدیریتی لازم، حاکمیت داده را نیز تثبیت نمود. پیچیدگی فرآیند توسعه مخزن (خزانه) داده برای موفقیت نیاز به پیروی از یک روش شناختی دارد. یک رویکرد گسترده پذیرفته شده برای توسعه، روش ترکیبی است که در آن الزامات و منابع داده باید در مدل جدید مخزن (خزانه) داده جای بگیرند [۵۹]. از سوی دیگر برای راهبری داده در یک سازمان نوین داده محور، نیاز به معماری یکپارچه برای ذخیره داده در مرحله اول و سپس پردازش آن هست که مباحث مدیریتی و راهبری داده هم در آن لحاظ شده باشد [۴۷]. بنابراین معماری انتخاب شده برای محصول نهایی استقرار خزانه داده با یک معماری سه لایه منطبق بر اهداف، شاخص ها و زیر شاخص های می باشد. و این معماری سه لایه می تواند ذات داده گان را بیان نماید. در جدول ۱۰ لیستی از مولفه ها، شاخص ها و زیر شاخص های نهایی شده قابل مشاهده می باشد.

جدول ۱۰. شاخص ها و زیر شاخص های نهایی شده برای خزانه دادگان

مولفه	شاخص	زیر شاخص
داده / اطلاعات	داده های باز	شفافیت
		دسترس پذیری
		API
داده های استاندارد	داده های استاندارد	مرجع پذیری
		قابل انتشار
		عدم دستکاری
		توکن بندی
		پاک بودن داده

کدینگ داده ای مستند	مدیریت داده	
کاتالوگ داده ای		
کیفیت داده		
حجم داده		
دسترس پذیری		
توسعه و انتشار		
مالکیت داده		
حذف		
آرشیو		
استاندارد پذیرش		
Open API		
حجم داده		
داده های تمیز		
قابل انتشار		
دسترس پذیری		
کاتالوگ داده	فرا داده	
دیکشنری داده		
قواعد و مقررات		
تبادل		
کیفیت داده	مدیریت منابع	مدیریت/راهبری
SLA		
دیکشنری متا داده		
دسترس پذیری		
استاندارد		
توزیع پذیری		
بهینه سازی		
سازماندهی		
تجهیزات		
دسترسی		
آرشیو		
توزیع مطلوب		
میزان پذیرش		

راهبری خزانه داده			توسعه بازار	برنامه راهبردی	
حریم خصوصی	قانون گذاری		اولویت های راهبردی		
داده گان عمومی			نظارت		
قانون استفاده			برنامه مالی		
مالکیت داده			تعهد مدیریتی		
اصالت سنجی			اعتماد سازی		
امنیت			اثربخشی	پشتیبانی	
تسری پذیری			مدیریت هزینه	متمرکز و بهینه	
اتکا پذیری			پاسخگویی مطلوب		
تنظیم مقررات			مرکز تماس		
اپراتور			سرویس در لحظه		
مالکیت معنوی			اولویت بندی سرویس		
دسترسى مجاز	دسترس پذیری	امنیت	استاندارد خدمات		
آدرس دهی			دردسترس بودن		
کتابخانه داده ای			حجم صف درخواستها		
کاتالوگ داده			پشتیبانی برخط		
سیستم لاگ			پرسش و پاسخ		
سرویس های امن			پوشش جغرافیایی		
احراز هویت			دسترس پذیری به دادگان		
حسابرسی			نیروی انسانی	ساختار	
SLA تبادل امن			مدل اقتصادی		
امضاء دیجیتال	احراز هویت		توزیع پذیری	توسعه	
استاندارد پذیرش			دسترس پذیری	سرزمینی	
الگوریتم نفوذ			شبکه ملی اطلاعات		
علائم تجاری	مالکیت معنوی		جامعیت قوانین	قوانین و مقررات	قوانین و مقررات
حقوق تجاری			مجوزهای دسترسی		
حقوق ثبت داده			توافق تعامل داده ای		
حق نشر			حمایت های حاکمیت		
حق استفاده			همکاری بین بخشی		
حق داده های محرمانه					
کتابخانه دانش	مدیریت دانش	کسب و کار			
اطلاعات آزاد					

ارائه چارچوبی برای داده گان ملی با تمرکز بر توسعه حاکمیت داده

open API		
پایگاه قواعد و مقررات		
تبادل پذیری		
امضاء دیجیتال		
حسابرسی		
فایروال		
سیستم مانیتورینگ		
قواعد دسترسی		
شبکه ارتباطی	زیرساخت فنی،	زیرساخت فنی
پهنای باند	ارتباطی	
توان محاسباتی		
فرکانس ارتباطی		
دسترسی		
پرسه (roaming)		
اتصال اختصاصی		
پرسرعت		
پهنای باند مبتنی بر تقاضا		
سرویس یک به یک		
سرویس یک به چند		
مدیریت IP		
شبکه اختصاصی		
مجازی		
سرویس مبتنی بر زیرساخت شبکه		
سرویس ابری		
توزیع سرورها		
مراکز داده		
دسترسی فیبری		
سیستم عامل		
مشاوره	مهارت و دانش	محیط توسعه
آموزش		

ذخیره دانش		
هوش تجاری		
سفارشی سازی اطلاعات	کسب و کار	
دسترس پذیری		
مدل داده		
مدل کسب و کار		
کسب و کار داده محور		
هوش تجاری		
میز خدمت		
زمان ارائه		
پاسخگویی		
مستند سازی		
هویت دیجیتال	کاربران بالقوه	کاربران
امضاء دیجیتال		
کاربران همزمان		
دسترس پذیری		
میزان آگاهی		
مسئولیت پذیری		
امنیت کاربران		
کدینگ سرویس ها	معماری	معماری
سرویس دیکشنری	داده/اطلاعات	
سرویس کاتالوگ		
نرخ تبادل		
ساختار دادگان	متدولوژی توسعه	
ساختار پروژه		
استاندارد توسعه		
ترافیک داده		
زمان دسترسی		
شبکه امن		
رابط کاربری	پلت فرم های نرم افزاری	تکنولوژی
هوش تجاری		
داده های آزاد		



مستند سازی		
فرهنگ سازی		
محتوای آموزشی		

## ۶. بحث و نتیجه گیری

برای نکویی و برازش مدل، ۵ شاخص را مدنظر قرار گرفت. شاخص RMR یکی از شاخص‌هایی است که بد بودن مدل تجربی را نشان می‌دهد هر دو شاخص RMR و SRMR ریشه مربعات حاصل از تفاوت بین ماتریس کوواریانس نمونه باقی مانده و مدل کوواریانس فرض شده می‌باشد. شاخص SRMR بین ۰ تا ۱ تغییر می‌کند که برای مدل‌هایی با برازش خوب این مقدار زیر ۰,۰۵ است که در تحقیق مقدار ۰,۲۰۹ حاصل شد. میزان شاخص‌خی- دو نیز ۷۱۲,۱۵ محاسبه گردید که مطلوب می‌باشد و شاخص NFI از شاخص‌های برازش نسبی، شامل گروهی از شاخص‌ها می‌شوند که مقادیر کای اسکوئر را با مدل پایه ای مقایسه می‌نمایند در واقع این شاخص‌ها بدنبال مقایسه بین یک مدل خاص با سایر مدل‌های ممکن می‌پردازد. با توجه به نتایج حاصل شده در این قسمت مقدار عددی آن برابر با ۰,۹۱۱ و شاخص R<sup>2</sup> نیز ۰,۹۲ بدست آمد که بیان از پذیرش مدل می‌باشد.

از معیار یا شاخص Gof برای بررسی اعتبار یا کیفیت مدل در تحلیل pls استفاده شد. این شاخص عددی بین صفر تا یک است که هر چه مقدار آن به یک نزدیکتر باشد نشان از اعتبار و کیفیت بالاتر مدل است. معمولاً مقادیر به دست آمده از این فرمول که بالاتر از ۳۵ درصد یا ۳۵۰ هزارم باشد نشان از اعتبار قابل قبولی در مدل pls است در تحلیل نهایی مقدار این شاخص ۰,۵ بوده که بانگر مطلوبیت مدل می‌باشد. بنابراین با توجه به نتایج حاصل شده در این قسمت مشاهده می‌شود که شاخص‌های برازش در مقادیر عددی مطلوبی بوده و در نهایت میتوان بیان نمود که مدل برازش یافته از نکویی برخوردار است.

اولین اهمیت توجه به دادگان ملی و ایجاد خزانه داده، تمرکز بر توسعه کسب و کار در کشور می‌باشد و یکی از اهداف این تحقیق،

توجه به چارچوب‌های این مهم می‌باشد. در توسعه فضای کسب و کار بایستی به ادبیات و مکانیزم‌های راهبردی، همچون حاکمیت داده<sup>۹</sup> و حاکمیت اطلاعات متناسب با سطح کاربردی و مشتریان توجه داشت. در مدل مورد تحقیق ایجاد جنبه‌های کسب و کار و توسعه متوازن مدنظر بوده و توجه شده است که حاکمیت داده، مجموعه‌ای از شیوه‌ها و فرآیندهای است که به تضمین مدیریت منابع داده‌ای در یک سازمان کمک می‌کند. گرچه اهمیت توجه به موضوع تسری حاکمیت داده با توجه به این مفهوم که حاکمیت داده اغلب شامل مفاهیمی همچون نظارت بر داده‌ها، کیفیت داده‌ها و نظایر آن با رویکرد کمک به شرکت در کنترل بهتر منابع اطلاعاتی، از جمله روش‌ها، فن‌آوری‌ها و رفتارها در کنار مدیریت مناسب داده‌ها می‌باشد، بوده است اما سایر جنبه‌های مورد نظر در قالب مفاهیم اولیه از در کنار هم قرار گرفتن داده‌های ملی با مفهومی همچون پارامترهای اولیه در مدل در نظر گرفته شده است. از سوی دیگر شامل موضوعاتی همچون امنیت و حفظ حریم خصوصی، یکپارچگی، قابلیت استفاده، ادغام، انطباق، در دسترس بودن، نقش‌ها و مسئولیت‌ها، و مدیریت کلی جریان‌های داده‌های داخلی و خارجی در یک سازمان می‌باشد [۶۰]. باید توجه داشت که حاکمیت داده، یک سیستم تصمیم‌درست و مسئولیت‌پذیر برای فرآیندهای مرتبط با اطلاعات است که بر اساس مدل‌های توافق شده‌ای و با رویکرد دسترسی کنترل شده، شرایط استفاده بهینه و مشخص و متدهای اجرایی قابل قبول، اجرا می‌شود (۶۱).

مدل طراحی شده می‌تواند مبنای برنامه ریزیهای اجرای در کشور در حوزه دغدغه افزایش دادگان قرار گیرد و می‌توان از آن به عنوان الگویی پایه در مطالعات بعدی نیز استفاده نمود.

## مراجع

[۱] Pauleen, D. and Wang, W. (۲۰۱۷), "Does big data mean big knowledge? KM perspectives on big data and analytics", Journal of Knowledge Management, Vol. ۲۱ No. ۱, pp. ۱-۶.

۹. Data Governance

[۱۱] Treloar, Andrew (۲۰۰۹). Design and Implementation of the Australian National Data Service. The International Journal of Digital Curation ۱(۴)-۱۳۷-۱۲۵

[۱۲] سامنی، سارا؛ رامین نصیری و مهران محسن زاده (۱۳۹۸). مقایسه مدلها و چارچوبهای حاکمیت داده با هدف دستیابی به پیش نیازها، موانع و مؤلفه‌های یک برنامه حاکمیت داده برای داده‌های حوزه سلامت، ششمین کنفرانس بین المللی فناوری اطلاعات، کامپیوتر و مخابرات، گرجستان، دانشگاه پیام نور مرکز بین الملل گرجستان.

[۱۳] جان نثاری، مصطفی؛ امیرزاده گوغری، احسان؛ عبدالله زاده، علیرضا (۱۳۹۸). حکمرانی داده محور. کار و جامعه. ۳۲۰. ص ۱۷-۳۲.

[۱۴] صادقی هفشجانی، اسماعیل و درویشی، اصغر، ۱۳۹۸، مدل سازی داده های بزرگ با رویکرد کاهش افزونگی در پایگاه داده های پردازش موازی، ششمین کنفرانس بین المللی یافته های نوین علوم و تکنولوژی با محوریت علم در خدمت توسعه، تهران، <https://civilica.com/doc/۹۲۰۱۱۷>

[۱۵] پدیداران مقدم، فرهنگ، سنگ سفیدی، سیما (۱۳۹۸). مروری بر جوانب مختلف انبار داده‌ها، سومین همایش ملی دانش و فناوری مهندسی برق، کامپیوتر و مکانیک ایران، تهران، موسسه برگزار کننده همایش‌های توسعه محور دانش و فناوری سام ایرانیان.

[۱۶] پدیداران مقدم، فرهنگ و معصومه رجبی (۱۳۹۷). افزایش امنیت در انبار داده، دومین کنفرانس ملی فناوری‌های نوین در مهندسی برق و کامپیوتر، اصفهان، موسسه آموزش عالی جهاد دانشگاهی استان اصفهان.

[۱۷] میرزایی پری، یحیی، خلیلی، مهدی، رضایی، اطهر (۱۳۹۷). حکمرانی داده؛ ضرورتی برای مواجهه با تحولات پیش روی صنعت بیمه، بیست و پنجمین همایش ملی بیمه و توسعه، تهران، پژوهشکده بیمه.

[۱۸] لگزائی، علیرضا؛ نسرین کاظمی و مسعود عسگری مهر (۱۳۹۷). چارچوبی جهت پیاده سازی حاکمیت داده در صنعت بانکداری ایران، دومین کنفرانس بین المللی مدیریت سیستم های فازی، ایوان کی، دانشگاه ایوان کی- موسسه بین المللی مهد پژوهش ره پویان حقیقت.

[۱۹] مرعشی پور، امید (۱۳۹۳) پژوهشی با عنوان «ارائه چارچوب حاکمیت داده‌های بزرگ توانمندساز مدیریت جریان دانش (مورد مطالعه: بانک مرکزی جمهوری اسلامی ایران). پایان نامه کارشناسی ارشد. مدیریت

[۲] Tan, K. (۲۰۱۸), "Managerial perspectives of big data analytics capability towards product innovation", Strategic Direction, Vol. ۳۴ No. ۸, pp. ۳۳-۳۵.

[۳] Cervone, H. (۲۰۱۶), "Organizational considerations initiating a big data and analytics implementation", Digital Library Perspectives, Vol. ۳۲ No. ۳, pp. ۱۳۷-۱۴۱.

[۴] Ahmed, W. and Ameen, K. (۲۰۱۷), "Defining big data and measuring its associated trends in the field of information and library management", Library Hi Tech News, Vol. ۳۴ No. ۹, pp. ۲۱-۲۴.

[۵] Amankwah- Amoah, J. (۲۰۱۵), "A unified framework for incorporating decision making into explanations of business failure", Industrial Management & Data Systems, Vol. ۱۱۵ No. ۷, pp. ۱۳۴۱-۱۳۵۷. <https://doi.org/۱۰.۱۱۰۸/IMDS-۰۳-۲۰۱۵-۰۰۸۵>

[۶] Alhassan, I., Sammon, D. and Daly, M. (۲۰۱۹), "Data governance activities: a comparison between scientific and practice-oriented literature", Journal of Enterprise Information Management, ۳۱ (۲): ۳۰۰-۳۱۶.

[۷] فتح اله زاده، فرحناز و مصطفی امینی (۱۳۹۶). مروری بر مدل‌های بلوغ حکمرانی داده‌ها در بنگاه‌های اقتصادی، پنجمین همایش مدیران فناوری اطلاعات، تهران، پژوهشگاه علوم و فناوری اطلاعات ایران و نهاد ریاست جمهوری.

[۸] طالبی، سجاد (۱۳۹۸). پردازش پایگاه داده‌ها و مقایسه بین روش‌های داده کاوی، سومین کنفرانس ملی ایده‌های نوین در فنی و مهندسی، رشت، سازمان بسیج مهندسی صنعتی استان گیلان.

[۹] Alam Khan Fakhri, Awais Ahmad, Muhammad Imran, Mafawez Alharbi, Mujeeb-ur-rehman, Bilal Jan (۲۰۱۷). Efficient data access and performance improvement model for virtual data warehouse, Sustainable Cities and Society, Volume ۳۵, Pages ۲۳۲-۲۴۰.

[۱۰] Serrano-Vicente, R., Melero, R. and Abadal, E. (۲۰۱۸), "Evaluation of Spanish institutional repositories based on criteria related to technology, procedures, content, marketing and personnel", Data Technologies and Applications, Vol. ۵۲ No. ۳, pp. ۳۸۴-۴۰۴.

[۲۹] Alhassan I., Sammon, D., Daly., M (۲۰۱۹). Critical Success Factors for Data Governance: A Theory Building Approach, Information Systems Management, ۳۶:۲, ۹۸-۱۱۰

[۳۰] سهرابی، بابک، اعظمی، امیر، یزدانی، حمیدرضا (۱۳۹۰). آسیب شناسی پژوهش‌های انجام شده در زمینه مدیریت اسلامی با رویکرد فرا ترکیب. چشم انداز مدیریت دولتی. (۶) ۲، ۹-۲۴

[۳۱] نیرومند، پوراندخت، رنجبر، محبوبه، بامداد صوفی، جهانیار (۱۳۹۰). شایستگی‌های حرفه‌ای مدیران و توسعه پایدار منطقه ویژه اقتصادی انرژی پارس (ارائه یک چارچوب پیشنهادی). فصلنامه مدیریت و منابع انسانی در صنعت نفت، (۸) ۲، ۷۳-۱۰۱.

[۳۲] Sandelowski, M., Barros, J. (۲۰۰۷). Handbook for Synthesizing Qualitative Research. Springer publishing company Inc.

[۳۳] صدیقی، امیرحسین (۱۳۹۶). حکمرانی مسئولانه داده از منظر اجتماعی، سومین کنفرانس بین‌المللی مدیریت و مهندسی صنایع، تهران، دانشگاه مقدس اردبیلی.

[۳۴] Gartner. (۲۰۰۹). Gartner EXP Worldwide Survey of More than ۱,۵۰۰ CIOs Shows IT Spending to Be Flat in ۲۰۰۹. STAMFORD, Conn

[۳۵] Golfarelli M. and Rizzi S. (۲۰۰۹) Data warehouse design, McGraw Hill ۲۰۰۹.

[۳۶] Dori, D., Feldman, R., Sturm, A. (۲۰۰۸) From conceptual models to schemata: An object-process-based data warehouse construction method, Information Systems, ۷۳۳, Issue ۶, Elsevier.

[۳۷] Chang, Chih-Hung Fuu-Cheng Jiang, Chao-Tung Yang, Sheng-Cang Chou (۲۰۲۰). On construction of a big data warehouse accessing platform for campus power usages, Journal of Parallel and Distributed Computing, Volume ۱۳۳, Pages ۴۰-۵۰.

[۳۸] Parn, E., Edwards, D. (۲۰۱۹), "Cyber threats confronting the digital built environment: Common data environment vulnerabilities and block chain deterrence", Engineering, Construction and Architectural Management, Vol. ۲۶ No. ۲, pp. ۲۴۵-۲۶۶.

[۳۹] Wang, Y. and Zhang, J. (۲۰۱۷), "Exploring topics related to data mining on Wikipedia", The Electronic Library, Vol. ۳۵ No. ۴, pp. ۶۶۷-۶۸۸.

- مدیریت فناوری اطلاعات - مدیریت دانش. دانشگاه آزاد اسلامی واحد تهران مرکزی، دانشکده مدیریت.

[۲۰] فتح اله زاده، فرحناز (۱۳۹۶). مدیریت کیفیت داده؛ سازماندهی پاسخگویی و مدل حکمرانی داده آن، هفتمین کنفرانس ملی مدیریت، اقتصاد و حسابداری، تبریز، سازمان مدیریت صنعتی تبریز.

[۲۱] فتح اله زاده، فرحناز (۱۳۹۶). الزامات ارزیابی مدل بلوغ داده: آیا مدل‌های حکمرانی داده این الزامات را برآورده می‌کنند، دهمین همایش ملی ادکا، تهران، اتحادیه انجمن‌های علمی دانشجویی علم اطلاعات و دانش شناسی ایران.

[۲۲] سدید، جواد؛ صاحبی وایقان، سعیده؛ رضائیان، هانی (۱۳۹۶). بررسی استفاده از خوشه بندی جهت کاهش زمان پرس و جوهای تجمیع داخل پایگاه داده مکانی مطالعه موردی: رسترهای بارش. اطلاعات جغرافیایی. ۱۰۳، ص ۵-۱۶.

[۲۳] پدیداران مقدم، فرهنگ؛ حاتمی، سید شهلا (۱۳۹۶). بررسی مدل امنیتی انبار داده مبتنی بر متادیتا. تهران: کنفرانس بین‌المللی فناوری اطلاعات، مهندسی کامپیوتر و مخابرات. حامی؛ انجمن هم‌اندیشان مبتکر رادمان.

[۲۴] نوایی راد، سمیه؛ میرعابدینی، سید جواد (۱۳۹۵) پژوهشی با عنوان «بهبود عملکردی امنیت یکپارچگی پایگاه داده‌های تجاری به کمک الگوریتم. شبک. ۲(۱۴). ص ۱-۱۲.

[۲۵] رفیعی کرکوندی، مجتبی؛ ایزدی، سید کامیار؛ خوش صفت، ابوالفضل (۱۳۹۴). تهدیدات و تمهیدات پایگاه داده به عنوان سرویس در رایانش ابری از دیدگاه پدافند غیر عامل. پدافند غیر عامل. ۲۲، ص ۶۳-۷۴.

[۲۶] لطفی احمدآبادی، محمد، ثقفی، محبوبه (۱۳۹۸). ارائه یک روش جدید برای ذخیره سازی و دسترسی به داده‌ها در پایگاه داده‌های بزرگ توزیع شده مبتنی بر ipfs و etl، ششمین کنفرانس بین‌المللی یافته‌های نوین علوم و تکنولوژی با محوریت علم در خدمت توسعه، تهران، دبیرخانه دائمی کنفرانس.

[۲۷] مرتضوی، محمدرضا (۱۳۹۴). طراحی چارچوب حاکمیت داده در مراکز تبادل داده و اطلاعات ملی (مورد مطالعه: وزارت علوم، تحقیقات، و فناوری). پایان نامه دکتری. : مدیریت - مدیریت فناوری اطلاعات: مدیریت - مدیریت فناوری اطلاعات.

[۲۸] Nokkala, T., Hannu, S., Jouko T., (۲۰۱۹). Data Governance in Digital Platforms. Conference: AMCIS ۲۰۱۹, At Cancún, México

- [۵۰] Astriani, W., Trisminingsih, R. (۲۰۱۶). Extraction, Transformation, and Loading (ETL) Module for Hotspot Spatial Data Warehouse Using Geokettle, *Procedia Environmental Sciences*, Volume ۳۳, Pages ۶۲۶-۶۳۴.
- [۵۱] Bouaziz Senda, Ahlem Nabli, Faiez Gargouri (۲۰۱۹). Design a Data Warehouse Schema from Document-Oriented database, *Procedia Computer Science*, ۱۵۹: ۲۲۱-۲۳۰.
- [۵۲] Mourtzis Dimitris, Vasilios Samothrakis, Vasilios Zogopoulos, Ekaterini Vlachos (۲۰۱۹). Warehouse Design and Operation using Augmented Reality technology: A Papermaking Industry Case Study, *Procedia CIRP*, ۷۹: ۵۷۴-۵۷۹
- [۵۳] Van der Heide G., N.D. Van Foreest, K.J. Roodbergen (۲۰۱۸). Optimizing stock levels for rental systems with a support warehouse and partial backordering, *European Journal of Operational Research*, ۲۶۵ (۱): ۱۰۷-۱۱۸
- [۵۴] Zekri Manel, Sahbi Zahaf, Faiez Gargouri (۲۰۱۹). Specification of the data warehouse for the decision-making dimension of the Bid Process Information System, *Procedia Computer Science*, ۱۵۹: ۱۱۹۰-۱۱۹۷
- [۵۵] Smith J. S., Ansari, M. N (۲۰۱۹). Robust analysis and optimization of a novel efficient quality assurance model in data warehousing. *Computer Science*, DOI:۱۰,۱۰۱۶/J.COMPELECENG.۲۰۱۹,۰۲,۰۰۳ Corp us ID: ۸۶۶۷۲۰۸۹
- [۵۶] Horta, M., Coelho, F., Relvas, S. (۲۰۱۶). Layout Design Modelling for a Real World Just-in-Time Warehouse. *Computers & Industrial Engineering* ۱۰۱: ۱-۹. DOI:۱۰,۱۰۱۶/j.cie.۲۰۱۶,۰۸,۰۱۳
- [۵۷] Vela Belén, Carlos Blanco, Eduardo Fernández-Medina, Esperanza Marcos (۲۰۱۲). A practical application of our MDD approach for modeling secure XML data warehouses, *Decision Support Systems*, ۵۲ (۴): ۸۹۹-۹۲۵.
- [۵۸] Gosain Anjana, Heena (۲۰۱۵). Literature Review of Data Model Quality Metrics of Data Warehouse, *Procedia Computer Science*, ۴۸: ۲۳۶-۲۴۴.
- [۵۹] Maté Alejandro, Juan Trujillo (۲۰۱۴). Tracing conceptual models' evolution in data warehouses by [۴۰] Tenopir, C., Dalton, E. D., Allard, S., Frame, M., Pjesivac, I., Birch, B., (۲۰۱۵). Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *Plos One*, ۱۰(۸). doi: ۱۰,۱۳۷۱/journal.pone.۰۱۳۴۸۲۶
- [۴۱] Hardy Ian, M., Obaidul, H., Vicente, R., (۲۰۱۸). Data for learning? Confirming and contesting performative practices of data governance. *Global Studies of Childhood*. Vol ۸, Issue ۴.
- [۴۲] Buffenoir, E.; Bourdon, I. (۲۰۱۳). Managing extended organizations and data governance. *Adv. Intell. Syst. Comput.* ۲۰۵, ۱۳۵-۱۴۵.
- [۴۳] Loshin, D. (۲۰۱۰). *Operationalizing Data Governance through Data Policy Management*; Knowledge Integrity, Inc.: Washington, DC, USA, ۱۳۲-۱۴۴
- [۴۴] Kim, S. and Lee, W. (۲۰۱۴), "Global data repository status and analysis: based on Korea, China and Japan", *Library Hi Tech*, Vol. ۳۲ No. ۴, pp. ۷۰۶-۷۲۲.
- [۴۵] Berro, A., Megdiche, I., Teste, O. (۲۰۱۵). Graph-based ETL processes for warehousing statistical open data. In *ICEIS*, (Vol. ۲۰۱۵ pp. ۲۷۱-۲۷۸).
- [۴۶] Merkus Jan, Remko Helms, and Rob Kusters (۲۰۱۹). *Data Governance and Information Governance: Set of Definitions in Relation to Data and Information as Part of DIKW*. International Conference on Enterprise Information Systems.
- [۴۷] [۴۷] بنائی، سیدمجتبی و محسن صابری (۱۳۹۶). دریاچه داده، بستری ضروری برای حکمرانی داده در سازمانها، پنجمین همایش مدیران فناوری اطلاعات، تهران، پژوهشگاه علوم و فناوری اطلاعات ایران و نهاد ریاست جمهوری.
- [۴۸] Kern, Rafał. Adrianna Kozierekiewicz, Marcin Pietranik, (۲۰۲۰). The data richness estimation framework for federated data warehouse integration, *Information Sciences*, Volume ۵۱۳, ۲۰۲۰, Pages ۳۹۷-۴۱۱, ISSN ۰۰۲۰-۰۲۵۵
- [۴۹] Amuthabala, P., Santhosh, R., (۲۰۱۹). Robust analysis and optimization of a novel efficient quality assurance model in data warehousing. *Computer Science*, DOI:۱۰,۱۰۱۶/J.COMPELECENG.۲۰۱۹,۰۲,۰۰۳ Corp us ID: ۸۶۶۷۲۰۸۹

using the model driven architecture, Computer Standards & Interfaces, ۳۶(۵): ۸۳۱-۸۴۳.

[۶۰] Michel knight, dataversity, ۲۰۱۹

[۶۱] DG (DG: Data Governance Institute; It is a non-government organization that work on data governance framework solutions), ۲۰۱۹.