

مطالعه مبانی، مدل‌ها و مسائل گزینش و مدیریت داده‌های پژوهشی در محیط‌های علمی و دانشگاهی

اکبر مجیدی: دانشجوی دکتری علم اطلاعات و دانش‌شناسی، دانشکده مدیریت، دانشگاه تهران، تهران، ایران (نویسنده مسئول).
akbarmajidi@ut.ac.ir

دکتر نادر نقشیینه: استادیار علم اطلاعات و دانش‌شناسی، دانشگاه تهران، تهران، ایران.

دکتر محمدرضا اسمعیلی گیوی: استادیار مدیریت سیستم‌ها، دانشگاه تهران، تهران، ایران.

دکتر محمودرضا هاشمی: استادیار فناوری اطلاعات، دانشگاه تهران، تهران، ایران.

چکیده

دریافت: ۱۳۹۵/۱۰/۱۸
پذیرش: ۱۳۹۶/۸/۲۰

زمینه و هدف: گزینش و مدیریت داده‌های پژوهشی حوزه نوظهور آموزشی و پژوهشی در علم اطلاعات و دانش‌شناسی است. این حوزه دارای مسائل و ابعاد گسترده‌ای است. هدف این مقاله مطالعه، شناسایی و بحث و بررسی مبانی و مفاهیم، مدل‌ها و چهارچوب‌ها و ابعاد و چالش‌های گزینش و مدیریت داده‌های پژوهشی در محیط‌های علمی و دانشگاهی است.

روش پژوهش: مقاله حاضر یک مقاله مروری است و از روش کتابخانه‌ای برای گردآوری مدارک و متون علمی و پژوهشی این حوزه استفاده شد. در این پژوهش پایگاه‌های اطلاعات علمی خارجی و داخلی و همچنین منابع وب با کلیدواژه‌های «data curation»، «digital curation»، «research data management»، «research data curation» و معادل آنها در فارسی جست و جو شدند. پس از حذف منابع تکراری و منابع نامرتبط، در مجموع ۱۳۲ منبع تفکیک و محتوی آنها ارزیابی و تحلیل نهایی شد.

یافته‌ها: تحلیل متون نشان داد که گزینش و مدیریت داده‌های پژوهشی حوزه‌ای نوظهور با مسائل پیچیده و ابعاد مختلفی است که طیف وسیعی از مسائل آموزشی، سازمانی، فرهنگی، فناوری، قانونی، امنیتی را شامل می‌شود. این حوزه از دهه ۲۰۰۰ میلادی مورد توجه دولت‌ها، سازمان‌های تأمین مالی پژوهش، دانشگاه‌ها قرار گرفته و اجتماع پژوهشی قوی از پژوهشگران مختلف به خصوص پژوهشگران علم اطلاعات و دانش‌شناسی، آرشیو و سیستم‌های اطلاعاتی حول آن شکل گرفته است. امروزه گزینش و مدیریت داده‌های پژوهشی به عنوان یکی از اجزای اصلی سیاست‌گذاری کلان علم و فناوری مطرح شده و نیازمند حمایت دولت و نهادهای قانون‌گذار و سیاست‌گذار است. مدل‌ها و چهارچوب‌های مختلفی در سطوح مختلف ملی، موسسه یا اجتماع خاص و رشته برای درک

ابعاد گزینش و مدیریت داده‌های پژوهشی و اجرای آن ترسیم شده‌اند. همچنین مطالعه متون نشان داد که کتابخانه‌ها و کتابداران از شایستگی و صلاحیت لازم برای برعهده گرفتن نقش‌ها و مسئولیت‌های گزینش و مدیریت داده‌های پژوهشی در دانشگاه و موسسات علمی برخوردارند. با این حال، گزینش و مدیریت داده‌های پژوهشی در اجرا و عمل با چالش‌هایی از قبیل حریم خصوصی، مالکیت داده، حقوق مالکیت فکری، عدم به اشتراک‌گذاری داده توسط محققان، نبود زیرساخت‌های مناسب مدیریت داده، عدم آگاهی و شناخت و دانش درباره فرایند گزینش و مدیریت داده و غیره روبرو است.

نتیجه‌گیری: علیرغم پژوهش‌های گسترده در خارج، این حوزه در ایران چندان مورد توجه واقع نشده است. بنابراین این مقاله به طور نسبتاً جامع به موضوع و ابعاد گزینش و مدیریت داده‌های پژوهشی می‌پردازد و می‌تواند برای پژوهشگران، سیاست‌گذاران علم و فناوری، کتابداران به منظور تحقیق و اجرای خدمات گزینش و مدیریت داده‌های پژوهشی مفید باشد.

کلیدواژه‌ها: گزینش داده، مدیریت داده‌های پژوهشی، داده‌های پژوهشی، مدل‌های چرخه حیات، مدل‌های چرخه حیات گزینش داده، گزینش دیجیتال

مقدمه

مانند مرکز داده بلورشناسی کمبریج^۲ که در سال ۱۹۶۵ در بخش شیمی دانشگاه کمبریج^۳ تأسیس شد. در کنار این فعالیت‌ها، موزه‌های تاریخ طبیعی هم در طی سالیان متمادی داده‌های علمی را در قالب قطعات فیزیکی گزینش می‌کردند. این تلاش‌ها و پیشرفت‌ها، نقطه آغازین توجه به حفاظت و مدیریت دارایی‌های دیجیتال و به خصوص داده‌ها در

از اوایل دهه ۱۹۶۰، با تأسیس آرشیوهای داده علوم اجتماعی، تلاش‌های اولیه در زمینه حفاظت، اشتراک‌گذاری و قابل دسترس ساختن دارایی‌های دیجیتال آغاز شد. علاوه بر فعالیت‌های آرشیوهای داده در علوم اجتماعی، در علوم شیمی و زیست‌شناسی نیز آرشیوها و پایگاه‌های داده گزینش شده^۱،

^۲Cambridge Crystallographic Data Centre (CCDC)

^۳Department of Chemistry at the University of Cambridge

^۱Curated databases

کیفیت داده. رویه‌های گزینش دیجیتال به بهبود کیفیت داده کمک می‌کند و قابلیت اعتماد داده را بهبود می‌بخشد و اطمینان می‌دهد که داده‌ها به عنوان یک رکورد رسمی اعتبار دارند.^۳ تشویق به استفاده مجدد و اشتراک‌گذاری داده. رویه‌های گزینش، اشتراک‌گذاری و استفاده از داده را با بکار بردن استانداردهای مشترک تشویق و کمک می‌کند و با فراهم کردن اطلاعاتی درباره بافت و منشأ داده موجب می‌شود که در طول زمان به صورت کامل مورد بهره‌برداری قرار گیرد.^۴ حفاظت از داده. رویه‌های گزینش دیجیتال از کهنگی و اتلاف داده محافظت می‌کنند. گزینش، ارزش بلندمدت داده موجود را با در دسترس قرار دادن آن برای انجام تحقیق جدید ارتقا می‌دهد. استفاده مجدد از داده و ترکیب داده پژوهشی از منابع و حوزه‌های چندگانه به محققان اجازه خواهد داد که از سنج‌های گذشته برای تولید دانش جدید به شیوه موثر و بهنگام استفاده کنند (زیمرمن^۷، ۲۰۰۸). کوهن و همکاران^۸ (۲۰۱۴) گزینش و مدیریت داده‌های پژوهشی را به عنوان ابزاری برای بهبود کیفیت داده و تسهیل استفاده مجدد از آن می‌دانند و اعتقاد دارند که گزینش و مدیریت داده پژوهشی خوب می‌تواند ضمن کاهش تلاش مجدد و اجتناب از اتلاف داده به ایجاد همکاری موثر پژوهشی بین موسسات، گروه‌ها و افراد کمک کند.

با توجه به شناخت ارزش داده‌های پژوهشی و اهمیت گزینش و مدیریت آنها، موسسات مختلفی در سطح بین‌المللی و ملی از قبیل، مرکز گزینش دیجیتال^۹ در بریتانیا، مجله بین‌المللی گزینش دیجیتال^{۱۰} و ائتلاف حفاظت دیجیتال^{۱۱} راه‌اندازی و مشغول فعالیت هستند. علاوه بر این، به دلیل اینکه امروزه داده‌های پژوهشی به یک عنصر اصلی در سیاست‌گذاری علم و فناوری تبدیل شده است، لذا بسیاری از موسسات و نهادهای سرمایه‌گذار و سیاست‌گذار در تحقیق نیز بر اهمیت مدیریت داده‌های پژوهشی به خصوص برای پروژه‌هایی که با بودجه عمومی انجام می‌شوند، تاکید می‌کنند و رهنمودها، خط‌مشی‌ها و سیاست‌گذاری‌های مختلفی تدوین و منتشر کرده‌اند، از جمله، سازمان توسعه و همکاری اقتصادی^{۱۲}، بنیاد ملی علوم، کمیته مشترک سیستم‌های اطلاعاتی^{۱۳}، شوراهای

دانشگاه‌ها و سازمان‌های دولتی بود. لکن، آنچه اکنون در حوزه حفاظت، گزینش و مدیریت داده‌های پژوهشی^۱ در جریان است، به لحاظ نظریه و عمل از آن چیزی که در اوایل دهه ۱۹۶۰ انجام می‌گرفت، متفاوت است. امروزه، ضرورت توجه به گزینش و مدیریت داده‌های پژوهشی و سرمایه‌گذاری در آن فقط با انگیزه نگهداری و آرشیو داده‌های پژوهشی نیست، بلکه به دلیل ارزش بالقوه و حیاتی داده‌های پژوهشی است که باید در طول چرخه حیات خود، از نقطه خلق تا دفع داده، به صورت دقیق و مناسب حفظ و نگهداری شوند و از طریق ایجاد ارزش افزوده امکان کشف، بازیابی و استفاده مجدد از آنها در آینده با هدف تولید دانش جدید و دستیابی به راه‌حل‌ها و نتایج ارزشمند در مواجهه با چالش‌ها و مسائل پیچیده جوامع انسانی و محیطی فراهم شود (چاودری^۲، ۲۰۱۰). با آغاز دهه ۲۰۰۰ و با ظهور تحقیقات داده‌بر^۳، یا «علم الکترونیکی»^۴ و با افزایش حجم داده در «علم بزرگ»^۵ (یورگمن^۶، ۲۰۰۹)، بحث‌های مربوط به ارزش داده‌های پژوهشی و لزوم گزینش، حفاظت و مدیریت آنها با نگاهی عمیق‌تر و گسترده‌تر در رشته‌ها، مراکز علمی و پژوهشی و سازمان‌های تأمین‌کننده اعتبار پژوهش افزایش یافت. امروزه گزینش و مدیریت داده پژوهشی به خاطر دلایل متعددی به یک ضرورت در دنیای علم تبدیل شده است. گزینش و مدیریت داده پژوهشی می‌تواند از چهار جهت برای جامعه پژوهش سود رساند: ۱. بهبود دسترسی. رویه‌های گزینش داده دیجیتال امکان دسترسی مستمر به داده را فراهم می‌کند و سرعت دسترسی به داده‌های قابل اعتماد و طیفی از داده‌هایی را که می‌توانند در دسترس باشند، افزایش می‌دهد ۲. بهبود

^۱ گزینش داده معادل «data curation» و مدیریت داده‌های پژوهشی معادل «research data management» بکار رفته است. لازم به تذکر است که معادل دقیق و مورد توافق رسمی میان متخصصان و محققان در داخل کشور در مورد اصطلاح curation وجود ندارد. البته معادل‌های دیگری از قبیل نظارت، مدیریت چرخه حیات داده از سوی برخی از متخصصان علم اطلاعات و دانش‌شناسی پیشنهاد شده است. نویسنده هیچ ادعایی ندارد که معادل گزینش دقیق‌ترین و مناسب‌ترین معادل برای curation است. لذا باب واژه‌گزینی و معادل‌سازی مناسب برای curation همچنان باز است. از طرف دیگر، به دلیل اینکه مدیریت داده جزئی از اجزا و اعمال گزینش محسوب می‌شود و اصطلاحی قابل درک برای محققان، گزینش‌کنندگان و کتابداران و متخصصان اطلاعات است، لذا از عبارت ترکیبی گزینش و مدیریت داده‌های پژوهشی معادل research data management and curation استفاده شده است که به نظر می‌رسد عبارت روشن‌تر و جامع برای توصیف مفاهیم این حوزه باشد. بحث مفاهیم و اصطلاح‌شناسی در بخش دوم همین مقاله ارائه شده است.

^۲ Choudhury

^۳ Data-intensive

^۴ E-science

^۵ Big science

^۶ Borgman

^۷ Zimmerman

^۸ Kohn et al

^۹ Digital Curation Centre (DCC)

^{۱۰} International Journal of Digital Curation

^{۱۱} Digital Preservation Coalition

^{۱۲} Organisation for Economic Co-operation and Development (OECD)

^{۱۳} Joint Information Systems Committee (JISC)

توجه سیاستگذاران علم و فناوری، مجریان و مدیران قرار گیرد.

علیرغم پژوهش‌های گسترده‌ایی که در خارج از ایران در زمینه ابعاد و مسائل گزینش و مدیریت داده‌های پژوهشی صورت گرفته است، اما به نظر می‌رسد که در داخل کشور از جانب متخصصان علوم آرشیو، موزه، مدیریت سیستم‌های اطلاعاتی و داده، سیاست‌گذاری علمی و به خصوص علم اطلاعات و دانش‌شناسی چندان مورد توجه واقع نشده و کار پژوهشی خاصی در این حوزه صورت نگرفته است.^۴ افخمی (۱۳۸۹) از آن به عنوان شکاف موجود در یکی از عرصه‌های مهم مدیریت کلان پژوهش و فناوری در کشور یاد می‌کند. از این رو، هدف مقاله حاضر تبیین مبانی و مفاهیم، مدل‌ها و چهارچوب‌ها، ابعاد و مسائل گزینش و مدیریت داده‌های پژوهشی با استفاده از روش تحلیل محتوی متون پژوهشی و علمی است که به شیوه کتابخانه‌ایی گردآوری شده‌اند. این مقاله از آن جهت که می‌کوشد به صورت نسبتاً جامع به مبحث گزینش و مدیریت داده‌های پژوهشی بپردازد، بنابراین انتظار می‌رود که در جهت آشنایی بیشتر پژوهشگران ایرانی به خصوص پژوهشگران علم اطلاعات و دانش‌شناسی با مبانی نظری و عملی این حوزه و انجام پژوهش‌های جدی و مستمر در آن مفید باشد. همچنین می‌تواند برای سیاستگذاران علم و فناوری، مدیران دانشگاه‌ها و سازمان‌های علمی، کتابخانه‌ها و کتابداران که می‌خواهند در راستای سیاست‌گذاری و راه‌اندازی نظام‌های گزینش و مدیریت داده‌های پژوهشی گام بردارند راهگشا باشد.

روش

این پژوهش با رویکرد کیفی و با استفاده از روش تحلیل محتوی به بررسی و مطالعه متون پژوهشی و علمی حوزه گزینش و مدیریت داده‌های پژوهشی با هدف شناسایی و تبیین مبانی و مفاهیم، مدل‌ها و چهارچوب‌ها و ابعاد و مسائل گزینش و مدیریت داده‌های پژوهشی پرداخته است. برای این منظور پایگاه‌های اطلاعاتی علمی خارجی با استفاده از کلیدواژه‌های «*data curation*» و کلیدواژه‌های مرتبط از قبیل «*research data curation*»، «*digital curation*»، «*research data management*» در فیله‌های عنوان، کلیدواژه و چکیده در فاصله زمانی ۲۰۰۰ تا

تامین اعتبار و ولکام تراست^۱ در بریتانیا، موسسه ملی سلامت^۲ در ایالات متحده، شورای تحقیق استرالیا و بنیاد تحقیق آلمان^۳. به عنوان نمونه بنیاد تحقیق آلمان، پیشنهادات و رهنمودهایی درباره نحوه مدیریت داده پژوهشی، از جمله مستندسازی دقیق نتایج تحقیق و ذخیره پیوسته داده اولیه، تعاریف داده علمی اولیه، مفاهیم سازماندهی، ابرداده و استانداردها، مدیریت حقوق، فراهم‌آوری داده اولیه و نیز کیفیت توصیف داده ارائه می‌کند.

مروری بر مطالعات و پژوهش‌های نظری و عملی حوزه گزینش و مدیریت داده‌های پژوهشی نشان می‌دهد که گستره مسائل و موضوعات این حوزه بسیار وسیع، نوظهور و در مرحله ابتدایی تکامل خود هستند. دامنه موضوعات پژوهشی در گزینش و مدیریت داده‌های پژوهشی طیف وسیعی از مسائل از قبیل آموزش و تربیت دانشجویان، کتابداران و متخصصان در زمینه گزینش و مدیریت داده‌های پژوهشی؛ نقش کتابداران و کتابخانه‌های دانشگاهی و تحقیقاتی در پشتیبانی از فرایند و اجرای گزینش داده در سطح دانشگاه‌ها، رشته‌ها و مراکز تحقیقاتی؛ مطالعه و شناسایی نیازها، ملزومات، نگرش و رفتار و رویه‌های گزینش و مدیریت داده‌های پژوهشی محققان و اعضای هیئت علمی در سطح رشته، موسسه خاص یا رشته‌های مختلف و میان موسسات گوناگون؛ تدوین و اجرای مدل‌ها، زیرساختار و چهارچوب‌های مختلف گزینش و مدیریت داده‌های پژوهشی در سطح یک موسسه، رشته یا اجتماع علمی بزرگ؛ رویکردها و فعالیت‌های سیاست‌گذاری و اجرای گزینش و مدیریت داده پژوهشی در سطح جهانی و ملی؛ شناسایی چالش‌ها و مشکلات گزینش و مدیریت داده پژوهشی از قبیل تامین اعتبار، ثبات مالی و نیروی انسانی متخصص و دیگر مسائل فنی پیچیده از قبیل زیرساخت ساخت‌افزایی و نرم‌افزاری و معماری سیستم، معماری ابرداده، شناسه‌گرهای شیء دیجیتال، معماری مخزن داده؛ و مسائل حقوقی مانند حق مولف داده، مالکیت داده، حقوق مالکیت فکری؛ و کسب اعتبار برای نشر و اشتراک‌گذاری داده و پی‌گیری منشاء و استفاده مجدد از داده از طریق استناد داده و مسائل دیگر را در بر می‌گیرد. لذا، برای تحقق و راه‌اندازی خدمات یا زیرساختار گزینش و مدیریت داده‌های پژوهشی موثر در سطوح مختلف، از زیرساختار جهانی گرفته تا سطح یک رشته خاص، باید همه این موارد و مسائل دیگر مورد

^۴ این مقاله برگرفته از مبانی و ادبیات رساله دکتری نویسنده با عنوان «ارائه مدلی برای خدمات گزینش و مدیریت داده پژوهشی در مراکز دانشگاهی» است که در سطح دکتری انجام می‌شود.

^۱ Wellcome Trust

^۲ National Institutes of Health (NIH)

^۳ Deutsche Forschungsgemeinschaft (DFG)

جدول ۱. پایگاه‌های اطلاعاتی جستجو شده

Taylor	Springer	Proquest	Wiley	Jstor	Sage	ScienceDirect	Emerald	نام پایگاه داده
۱۵	۴	۱۰	۱۱۳	۱	۲۸	۹۹	۳۴	Research data curation
۱۰۲	۸۸	۲۶۴	۱۱۹	۶	۳۷	۲۸	۷۱	Research data management
۸۶	۴۹۷	۴۶۰	۳۰۵	۲۴	۴۴	۲۳۹	۵۸	Data curation
۷۱	۷۹	۱۰۷	۷۹	۵۰	۹۶	۴۳	۶۰	Digital curation

یافته‌ها

پس از مطالعه و تحلیل مدارک و متون، موضوعات و مسائل اصلی استخراج شدند که در نمودار ۱، ارائه شده است. در ادامه مقاله به بحث و بررسی هریک از این موضوعات خواهیم پرداخت.

داده‌های پژوهشی: مفهوم و انواع آن

ادبیات نظری گزینش داده نشان می‌دهد که ماهیت و رفتار داده به شدت متکی بر بافتی است که در آن خلق می‌شود. از این رو داده تابع هنجارها، روش‌شناسی‌ها، خط‌مشی‌ها و سنت‌های رشته‌ها و حوزه‌های علمی است و در نتیجه تعریف آن از رشته‌ای به رشته دیگر و از محقق با محقق دیگر متفاوت است و نمی‌توان در متون گزینش داده تعریف مشترکی از داده و خصیصه‌های آن یافت. در جدول ۱، تعاریف داده (پژوهشی) که طیفی از دیدگاه‌های سازمان‌ها، پروژه‌ها، مدل‌ها و مقالات نظری را در برمی‌گیرد، ارائه شده است.

داده پژوهشی داده‌ای است که محقق برای انجام تحقیق در اختیار دارد و به عبارت دیگر درونداد تحقیق است از قبیل پرسشنامه و هم‌برونداد و نتیجه‌ای است که از انجام تحقیق حاصل شده است. داده پژوهشی داده‌ای است که در بافت و بستر پژوهش علمی خلق و یا گردآوری می‌شود. لذا در مقاله حاضر، داده پژوهشی به عنوان همه اطلاعات گردآوری شده، مشاهده شده یا خلق شده برای اهداف تحلیل و اعتبارسنجی نتایج تحقیق اصیل در نظر گرفته می‌شود؛ بنابراین این اصطلاح: نوبت‌بوک‌های آزمایشگاه، تحلیل‌های اولیه و پیش‌نویس‌های مقالات علمی، برنامه‌های تحقیق آینده، مرور همکاران، یا ارتباطات شخصی با همکاران یا اشیاء فیزیکی (به طور مثال نمونه‌های آزمایشگاه، گونه‌ایی از باکتری‌ها و آزمایش حیوانات) را پوشش نمی‌دهد» (آی‌سی‌دی، ۲۰۰۷).

داده بر اساس مولفه‌های مختلف طبقه‌بندی شده است. مورد ملی علوم (۲۰۰۵) داده را بر حسب ماهیت داده (ماهیت داده می‌تواند متفاوت باشد، برای مثال، تصاویر، صدا، ویدئو، انیمیشن، نرم‌افزار، الگوریتم‌ها، مدل‌ها یا نتایج شبیه‌سازی)،

۲۰۱۷ جستجو شدند. انتخاب سال ۲۰۰۰ به این دلیل بود که ظهور این اصطلاح و آغاز تحقیقات جدی در این حوزه به اوایل دهه ۲۰۰۰ مربوط می‌شود. جستجوی مدارک محدود بودند به مقالات، فصول کتاب، پایان‌نامه‌ها، مقالات کنفرانس‌ها و گزارش پروژه‌ها. علاوه بر جستجو در پایگاه‌های اطلاعاتی، موتور جستجوی گوگل و گوگل اسکالر نیز جستجو شد. در جستجوی مقالات و دیگر مدارک، معیارهای انجام داوری توسط همتایان و امکان دسترسی تمام متن لحاظ شده بود و مقالات جستجو شده در قالب مروری، مطالعات موردی و پژوهشی اصیل بودند. نتایج جستجو در جدول ۱، ارائه شده است. پس از جستجو و گردآوری مدارک بازیابی شده، ابتدا تمامی مدارک یکپارچه‌سازی شدند و مدارک تکراری حذف شدند. سپس مدارک باقی مانده از نظر عنوان و چکیده بر اساس اهداف تحقیق بررسی شدند و مدارکی که به لحاظ عنوان نامرتب و ناسازگار به نظر می‌رسیدند حذف شدند و در مرحله بعد چکیده‌های مدارک باقی مانده مورد بررسی قرار گرفت و مدارکی که به نظر مناسب می‌رسیدند برای تحلیل نهایی انتخاب و بقیه موارد کنار گذاشته شدند. علاوه بر پایگاه‌های اطلاعات علمی خارجی، پایگاه‌های اطلاعاتی علمی داخل از جمله پایگاه اطلاعات علمی جهاد دانشگاهی، درگاه جامع علوم انسانی، پایگاه مجلات تخصصی نور و گوگل به زبان فارسی جستجو شد که تنها تعداد انگشت‌شماری (افخمی (۱۳۸۹)، پورعباسی، آقایی و حاجیان حیدری (۱۳۹۴)، صراف‌زاده (۱۳۹۴)، رضائی شریف‌آبادی (۱۳۹۴)، مجیدی و همکاران (۱۳۹۵)، مجیدی، نقشینه و اسمعیلی گیوی (۱۳۹۵)، وزیر، نقشینه و نوروزی چاکلی (زودآیند، ۱۳۹۶)) به موضوع گزینش و مدیریت داده‌های پژوهشی یا یکی از ابعاد آن پرداخته‌اند. در نهایت ۱۳۲ مدرک علمی شناسایی شده و مورد تجزیه و تحلیل نهایی قرار گرفت. از این تعداد، ۸۶ مقاله مجلات علمی، ۳۲ مورد فصل کتاب، پایان‌نامه، گزارش یا سند دانشگاهی، متن مربوط به شوراها و موسسات سیاست‌گذار بین‌المللی و ۱۴ مورد مربوط به مقالات کنفرانس و کارگاه‌ها و همایش‌ها بود.

¹ OECD



نمودار ۱. مسائل و ابعاد شناسایی‌شده در متون گزینه‌ش و مدیریت داده‌های پژوهشی

عبارتند از داده تجربی، مدل‌ها یا شبیه‌سازی‌ها، مشاهدات، داده استخراج‌شده و داده مرجع یا متعارف^۲ (داده‌ایی که داده دیگر را بر اساس استاندارد خاص توصیف می‌کند یا داده در قالب استاندارد بازنمون شده است). از طرف دیگر، هوکس و بویجی^۳ (۲۰۰۵) داده را بر اساس هدف استفاده به داده اولیه یا دست اول (داده اولیه یا خام برای یک هدف خاص با رویه خاص گردآوری شده است) و داده ثانویه یا دست دوم (اگر داده اولیه برای مطالعات و تحقیقات دیگر مجدداً مورد استفاده قرار گیرند به آنها داده ثانویه گفته می‌شود) طبقه‌بندی می‌کند. پژوهشگر داده باید چند نکته را در مورد تعاریف و طبقه‌بندی‌های داده (پژوهشی) مد نظر قرار دهد: تعاریف و تقسیم‌بندی‌ها نشان می‌دهند که تعاریف اغلب از منظر فنی ارائه شده‌اند و عملی هستند. بافت اجتماعی که در آن داده وجود دارد مورد غفلت واقع شده است؛ زیرا مشاهداتی که برای یک دانشمند به عنوان یافته‌های تحقیق تلقی می‌شوند، ممکن است برای دیگری در مقام پیشینه تحقیق قلمداد شود. داده‌ایی که به منزله شواهد کافی برای یک هدف مورد نظر است می‌تواند برای دیگران ناکافی به نظر برسد. همچنین

سطح قابلیت تولید مجدد (داده می‌تواند برحسب منشاء خود در سه طبقه، مشاهده‌ایی، محاسباتی، یا داده تجربی تقسیم‌بندی شود. داده مشاهده‌ایی شامل نتایج یک موقعیت خاص است و نمی‌تواند دوباره گردآوری شود. برای مثال، مشاهده مستقیم حرارات اقیانوس یا مشاهدات آب و هوا. داده محاسباتی نتیجه اجرای یک شبیه‌سازی یا مدل رایانه‌ایی است. این داده‌ها می‌توانند، در صورت مستند کردن نرم‌افزار و سخت‌افزار و داده درون‌داد، مجدداً تولید شوند. داده تجربی نتیجه مطالعات آزمایشگاهی یا تجربیات میدانی هستند. در تئوری، این داده‌ها می‌توانند مجدداً به صورت دقیق تولید شوند، اما در عمل، نمی‌توانند) و سطح پردازش (داده خام توسط سنسورها و ابزارها گردآوری می‌شود و سپس بر اساس هدف تحقیق اصلاح و تحلیل می‌شوند. داده استخراج‌شده توسط فعالیت‌های پردازشی و گزینه‌ش تولید می‌شوند) طبقه‌بندی می‌کند. بورگمن (۲۰۰۷) داده رکوردها را (به طور مثال از زندگی خصوصی و عمومی، کسب و کار و دولت) به سه طبقه بالا اضافه کرده است. شبکه اطلاعات تحقیق^۱ (۲۰۰۸) انواع داده را بر اساس منبع داده به پنج دسته تقسیم‌بندی می‌کند که معنی و محتوی تقسیم‌بندی مورد ملی علوم را دارد و

^۲ Canonical or reference data

^۳ Hox & Boeije

^۱ Research Information Network

جدول ۲. تعاریف داده (پژوهشی)

تعریف	نویسندگان/سازمان‌ها
داده «بازنمود قابل تفسیر مجدد اطلاعات در یک شیوه رسمی مناسب برای ارتباطات، تفسیر یا پردازش» است. نمونه‌هایی از داده عبارت‌اند از: توالی بیت‌ها، جدولی از اعداد، کاراکترهایی در یک صفحه، ضبط صداها.	مدل مرجع سیستم اطلاعاتی آرشیو باز ^۵ ، ۲۰۰۲ بورد ملی علوم ^۶ ، ۲۰۰۵
داده به‌عنوان هرگونه «اطلاعات...شامل متن، اعداد، تصاویر، ویدئو یا فیلم‌ها؛ صوت، نرم‌افزار، الگوریتم‌ها، معادلات، پویانمایی‌ها، مدل‌ها، شبیه‌سازی‌ها و غیره» در نظر گرفته می‌شود.	سازمان توسعه و همکاری اقتصادی ^۷ ، ۲۰۰۷
داده پژوهشی «رکوردهای حقیقی (نمره‌های عددی، رکوردهای متنی، تصاویر و صداها) هستند که به‌عنوان منابع دست اول برای تحقیق علمی استفاده شده و معمولاً در اجتماع علمی برای اعتبارسنجی یافته‌های تحقیق ضروری هستند. مجموعه داده پژوهشی یک بازنمون نظام‌مند و مفصل از موضوع بررسی شده ارائه می‌کند. تعریف خط‌مشی پیشنهاد می‌کند که هر ماده استفاده شده به‌منزله یک ماده اولیه برای تحقیق می‌تواند به‌عنوان داده پژوهشی در نظر گرفته شود، مانند متن‌های منتشرشده، مصنوعات یا داده پردازش نشده. تعریف سازمان توسعه و همکاری اقتصادی از داده شامل هر نوع منبعی است که برای محققان مفید است.	پروژه دیتا وان ^۸ ، ۲۰۱۲
داده پژوهشی همه اطلاعات گردآوری شده، مشاهده شده یا خلق شده برای اهداف تحلیل و اعتبارسنجی نتایج تحقیق اصیل در نظر گرفته می‌شود. بااین حال، داده‌ها محصول سازوکار علمی هستند و خودشان بیش از پیش به‌عنوان یک برون‌داد ارزشمند تحقیق درک می‌شوند.	نیروی‌کار ریون آبی درباره اقتصاد پایدار برای یک سیاره دیجیتال ^۹ ، ۲۰۱۰
داده به‌عنوان درون‌داد اولیه تحقیق و نیز نتایج دست اول آن تحقیق محسوب می‌شود. این تعریف نشان می‌دهد که داده ممکن است دو نوع باشد: داده تولید شده بیرون از یک پروژه تحقیق به‌عنوان درون‌داد و داده تولیدشده در داخل یک حوزه تحقیقاتی معین.	

صحت‌شده و تایید شوند. ابزارهایی که از ابرداده‌های شیء پشتیبانی می‌کنند از قبیل، فهرست‌های مستند، اصطلاحنامه^۵، دیکشنری داده^۶ و دیگر منابع مستندسازی، باید حفظ، بر روزرسانی و در فاصله‌های منظم توزیع شوند و در عین حال امنیت و دسترسی به آنها نیز در نظر گرفته شود. همه این مسائل رشته گزینش داده را تشکیل می‌دهد^۷ (زوریخ، ۱۹۹۵).
واژه گزینش از لغت لاتین کیورتنس^۸ گرفته شده است که به مفهوم مراقبت و مدیریت است (کوچران^۹، ۲۰۱۴). اصطلاح گزینش داده در ادبیات این حوزه با تعاریف متعددی توصیف شده است. دلیل آن هم این است که مفهوم گزینش عمیقاً وابسته به بافت و کاربردهای محلی است. لذا تعریف و تفسیر آن متأثر از ملزومات قلمرو و بافتی است که در آن گزینش انجام می‌شود (بیگری^{۱۰}، ۲۰۰۸؛ شوریش^{۱۱}، ۲۰۱۲). این معانی مختلف به سردرگمی به خصوص در بحث‌های میان رشته‌ای منجر شده است. برای مثال تحقیق مولینز^{۱۲} (۲۰۱۰) نشان داد که معنی گزینش برای برخی از اعضای هیئت علمی مهندسی بر موضوعات کنترل کیفی و اعمال مرور در انتخاب داده برای ورود به پایگاه داده اشاره می‌کند. این معنی از گزینش نسبت به تعریفی که معمولاً توسط کتابداران استفاده می‌شود بسیار خاص است (مولینز^{۱۳}، ۲۰۱۰). تعاریفی که توسط علوم کتابداری

ممکن است داده‌ایی که مورد تجزیه و تحلیل قرار گرفته، برای تحقیق دیگر داده خام باشد (نیلسن و یورلند^{۱۴}، ۲۰۱۴). از سوی دیگر، داده پژوهشی ناهمگن و بافت‌مدار است و می‌تواند اشکال متفاوتی برحسب خاستگاه، مسئله تحقیق و رشته‌های دانشگاهی اختیار کند. از این رو، داده پژوهشی باید همیشه در ارتباط با رشته علمی مربوطه مورد بررسی قرار گیرد؛ بنابراین پژوهشگر داده باید درک درستی از سنت‌ها، روش‌شناسی‌های تحقیق و ماهیت رشته‌ها بدست آورد.

گزینش داده: مفهومی بحث‌برانگیز

هرچند گزینش^{۱۵} اصطلاح جدیدی نیست و به خصوص در کار موزه و هنر به خوبی تثبیت شده است، اما ایده گزینش در ارتباط با داده پژوهشی نسبتاً جدید است و در دهه ۱۹۹۰ معمول شد و برای اولین بار در سال ۱۹۹۵ در مقاله «دیانا جورینج^{۱۶}»، درباره آینده مدیریت مجموعه‌ها، استفاده شد (پالمر و همکاران^{۱۷}، ۲۰۱۳). او از گزارش وزارت انرژی سال ۱۹۹۴، مفهوم گزینش‌کننده پایگاه داده را به حوزه کاملاً جدیدی برای کار اطلاعاتی در موزه‌ها، کتابخانه‌ها و محیط‌های آزمایشگاه علمی سنتی بسط داد و استدلال کرد که: «مجموعه داده‌ها باید در طول زمان نسبت به کیفیت بلندمدت و پایدار و ربط بررسی شده و منابع جدید داده شناسایی و ارزیابی شوند. تغییرات یا روزآمدسازی‌ها در داده‌ها باید تعیین

⁵Thesauri

⁶Data dictionary

⁷Curatus

⁸Cochran

⁹Beagrie

¹⁰Mullins

¹ Nielsen & Hjørland

² Curation

³ Zorich

⁴ Palmer et al

جدول ۳. برخی از تعاریف گزینش داده

منبع	تعریف گزینش داده
پاتل و بال ^۱ ، ۲۰۰۸	مدیریت فعال داده دیجیتال و نتایج تحقیق در طول حیات علمی و پژوهشی خود به منظور استفاده کنونی و آینده.
شریوزو و کراگین ^۱ ، ۲۰۰۸	مدیریت فعال و مستمر داده در طول چرخه حیات خود به نفع و سودمندی برای پژوهش، علوم و آموزش است، که شامل ارزیابی و انتخاب، بازنمایی، و سازماندهی این داده‌ها برای دسترسی و استفاده در طول زمان است. فعالیت‌های گزینش داده امکان کشف، بازیابی، حفظ کیفیت، ارزش افزوده فراهم می‌کنند و امکان استفاده مجدد از داده را در طول زمان فراهم می‌کند و این حوزه جدید شامل تعیین اعتبار، آرشیو، مدیریت، حفاظت، بازیابی و بازنمایی است.
لورد و مک دونالد ^۱ ، ۲۰۰۳	مدیریت فعال اطلاعات، شامل برنامه‌ریزی؛ عمل درگیر شدن در مراقبت از داده دیجیتال بعد از استفاده اولیه از آن؛ ایجاد ارزش افزوده برای داده؛ مدیریت، بهبود و تقویت داده.
لورد و همکاران، ۲۰۰۴	فعالیت مدیریت و ارتقای استفاده از داده از نقطه خلق آن، برای اطمینان از اینکه داده برای هدف کنونی مناسب است و برای کشف و استفاده مجدد در دسترس است. (... سطوح بالای نظارت همچنین حفظ پیوندها با یادداشت‌ها و مواد منتشر شده دیگر را در بر می‌گیرد.
کمیته مشترک سیستم‌های اطلاعاتی (جیسک)، ۲۰۰۳	همه فرایندهای مورد نیاز برای خلق داده و مدیریت آن و ایجاد ارزش افزوده برای تولید منبع جدید اطلاعات و دانش.
بال، ۲۰۱۰	گزینش داده فرایند انتخاب، نرمال‌سازی، یادداشت‌گذاری ^۱ ، و ترکیب داده از مجلات، گزارش‌ها یا پایگاه‌های داده شخص ثالث در داخل یک پایگاه داده در یک موضوع خاص به جهت بروزنگه داشتن و مرتبط نگه داشتن داده است.

حفظ ارزش بالقوه و بلندمدت داده است. مدیریت فعال اغلب نیازمند مشارکت گزینش‌کنندگان و نیز تولیدکنندگان داده (خلق‌کنندگان) است. گزینش داده پژوهشی صرفاً به معنی گزینش منابع نیست، بلکه همچنین گزینش «دانش» است و شامل تعاملات بین تولیدکنندگان، گزینش‌کنندگان و مصرف‌کنندگان (کاربران) داده است (آنگووار^۴، ۲۰۰۹).

بسیاری از تعاریف گزینش داده، بر «ارزش افزوده» به عنوان جنبه مهمی از گزینش تأکید می‌کنند. اعمال گزینش، ارزش افزوده برای داده فراهم می‌کنند تا اینکه داده آسان‌تر کشف‌شده و در دسترس باشد و از لحاظ محتوی غنی‌شده و به راحتی مجدداً استفاده شود. گزینش داده نباید یک فعالیت یکباره تلقی شود بلکه یک فعالیت فرایندی است که پیش از خلق داده آغاز می‌شود و تا دفع داده ادامه می‌یابد و موجب می‌شود که داده مجدداً وارد چرخه حیات تحقیقات شده و به کشفیات و تولید دانش جدید بیانجامد. از این رو، گزینش فرایند ساده‌ای نیست و صرفاً به عنوان یک هدف تلقی نمی‌شود، بلکه ابزاری برای گردآوری، سازماندهی، اعتبارسنجی و حفاظت داده برای پرداختن به چالش‌های بزرگی است که جامعه با آنها روبروست (چاودری و هانیش^۵، ۲۰۰۹). گزینش صرفاً یک تعهد و امر فنی نیست، بنابراین موفقیت آن تا حد زیادی به مهارت و دانش و سواد داده محققان، تحلیل و تصمیم‌گیری ذینفعان، ثبات مالی و اقتصادی، آموزش و تغییر فرهنگ سازمانی و تربیت گزینش-

و اطلاع‌رسانی درباره گزینش داده استفاده می‌شوند نیز متفاوت هستند. در برخی اوقات اصطلاحات گزینش داده و گزینش دیجیتال به جای هم استفاده می‌شوند. برخی از تعاریف گزینش داده، کارکردهای حفاظت و آرشیو را شامل می‌شوند (لورد و همکاران^۱، ۲۰۰۴؛ یوآی‌یوسی^۲، ۲۰۱۱)، در حالی که تعاریف دیگر این کارکردها را صریحاً از هم جدا می‌کنند و آنها را به صورت مستقل از گزینش داده تعریف می‌کنند (مک‌گاورن^۳، ۲۰۰۹). نبود یک تعریف پذیرفته‌شده جهانی هم در داخل و هم بیرون از حوزه کتابداری و اطلاع‌رسانی مانعی است که باید به آن توجه شود. در جدول ۲، برخی از تعاریف گزینش داده ارائه شده است.

بررسی و تحلیل تعاریف نشان می‌دهد که جنبه‌ها، ابعاد و ویژگی‌های هسته و مشترکی در میان تعاریف ارائه‌شده وجود دارند که به درک روشن و صحیح ما از گزینش داده کمک می‌کنند. از قبیل، گزینش داده پیوستاری از فعالیت‌ها است که استفاده از داده را در حال و آینده پشتیبانی می‌کند. فعالیت‌های اساسی و پایه گزینش سازماندهی، فراهم کردن دسترسی و حفاظت هستند. گزینش و فعالیت‌های آن امکان کشف، بازیابی، کیفیت و مدیریت ارزش افزوده و استفاده مجدد از داده پژوهشی را تسهیل می‌بخشند. گزینش، مدیریت فعال داده دیجیتالی است. این مدیریت فعال (مراقبت) داده در طول چرخه حیات، کلید قابلیت تولیدمجدد، قابلیت استفاده مجدد و

¹Lord et al²The University of Illinois at Urbana-Champaign (UIUC) Graduate School of Library and Information Science³McGovern⁴Angevaar⁵Choudhury&Hnisch

جدول ۴. دیدگاه‌های نویسندگان درباره روابط مفهومی میان گزینش داده و دیگر اصطلاحات مرتبط

دیدگاه	نویسنده
در کتابداری و اطلاع‌رسانی تعریف حفاظت اخیراً به شامل شدن آرشیو و گزینش دیجیتال گسترش یافته است. این تعاریف هنوز در حال تکامل هستند و معمولاً به جای هم استفاده می‌شوند که به برقراری ارتباطات دقیق و روشن زبان می‌رساند. در علوم کتابداری و اطلاع‌رسانی تعاریف گزینش داده متعدد است. معمولاً اصطلاحات گزینش داده و گزینش دیجیتال به جای هم استفاده می‌شوند.	چاودری و هینیش ^۱ (۲۰۰۹)، بیگری ^۲ (۲۰۰۶)
گزینش و نظارت معانی متفاوتی دارند، هرچند که این اصطلاحات برخی اوقات به جای هم استفاده می‌شوند. گزینش داده و نظارت داده درباره ماهیت داده، چرخه‌های حیات و روابط با محیط‌های علمی دیدگاه‌های متفاوتی دارند. گزینش داده با حفاظت دیجیتال و برخی اوقات با آرشیو تلفیق می‌شود. آرشیو و حفاظت ممکن است زیرمجموعه فعالیت‌های مرتبط با گزینش داده باشند یا فی نفسه حوزه‌های تحقیق و عمل حرفه‌ای.	ژاکوب و ورلی ^۱ ، ۲۰۰۹؛ کراستی، بیکر و هالکولا ^۱ ، ۲۰۰۶ لورد و مک دونالد ^۳ ، ۲۰۰۳
به صورت فزاینده، گزینش دیجیتال به یک مفهوم چترمانندی تبدیل می‌شود که شامل حفاظت دیجیتال، گزینش داده، مدیریت رکوردهای الکترونیکی، و مدیریت دارایی دیجیتال است. این مسئله بوسیله تحلیل چکیده‌های مقالات مجله بین‌المللی گزینش دیجیتال نیز تایید شده است. در حال حاضر قلمرو گزینش دیجیتال حوزه ترکیبی است که حوزه‌های مختلف مرتبط با مدیریت و حفاظت اشیای دیجیتالی و ایجاد ارزش افزوده را در طول زمان و چرخه حیات خود در بر می‌گیرد.	یاکل ^۱ ، ۲۰۰۷؛ نیو ^۱ ، ۲۰۱۴؛ رای ^۱ ، ۲۰۰۹
برنامه‌ریزی مدیریت داده، گزینش داده، گزینش دیجیتال و حفاظت دیجیتال مستقل از هم اما مفاهیم مرتبطی هستند. این مفاهیم کلیدی نباید به جای هم استفاده شوند. مفاهیم کلیدی باید زمانی که تحقیق مدیریت و گزینش داده میان رشته‌های چندگانه انجام می‌شود، به درستی شناسایی، روشن و تعریف شود.	اسمیت ^۱ ، ۲۰۱۴، ۲۰۱۲

گزارش جیسک (۲۰۰۳) درباره زیرساخت‌های موردنیاز برای پشتیبانی از علم الکترونیکی در سال ۲۰۰۳، یکی از تلاش‌های اولیه برای مشخص کردن تفاوت‌های مفهومی بین حفاظت، آرشیو و گزینش (جدول ۴) در ارتباط با پژوهش دیجیتالی بود. این گزارش، گزینش داده را به عنوان فعالیت «مدیریت و بهبود استفاده از داده از نقطه خلق آن، برای اطمینان از اینکه برای هدف کنونی مناسب و برای کشف و استفاده مجدد قابل دستیابی است» تعریف می‌کند. بر اساس این گزارش «آرشیو» لازمه «گزینش» و «حفاظت»، «یک جنبه از آرشیو» است (لورد و همکاران، ۲۰۰۴).

گیارلو^۵ (۲۰۱۳) اعتقاد دارد که گزینش داده، مفهوم حفاظت دیجیتال را شامل می‌شود. تفاوت ظریفی بین گزینش داده و حفاظت بلندمدت وجود دارد. گزینش نه تنها بر حفاظت و نگهداری گروهی از مجموعه داده‌ها دلالت دارد بلکه همچنین بر خلق ارزش افزوده و تولید دانش جدید مرتبط است. مرکز گزینش دیجیتال بریتانیا نیز چنین برداشتی از گزینش داده دارد و بیان می‌کند که گزینش دیجیتال شامل نگهداری، حفاظت و ایجاد ارزش افزوده برای داده پژوهشی دیجیتال در سراسر چرخه حیات داده است (مرکز گزینش دیجیتال، ۲۰۱۰).

تمایز مفهومی بین گزینش داده و گزینش دیجیتال نامحسوس ولی مهم است. اصطلاح «گزینش دیجیتال» برای اولین بار در سال ۲۰۰۱ در بریتانیا، به منزله عنوان سمیناری درباره آرشیوهای دیجیتال، کتابخانه‌ها و علم الکترونیکی استفاده شد

کنندگان ماهر بستگی دارد. در نهایت باید گفت که رسالت گزینش داده فراتر از استفاده مجدد از داده پژوهشی است و به تولید، تایید داده و دانش جدید کمک می‌کند.

مفاهیم مرتبط با گزینش داده (پژوهشی)

گزینش داده به جهت اینکه حوزه نسبتاً جدیدی است، اصطلاحات آن از ثبات مفهومی لازم برخوردار نیستند و مرز روشی میان مفاهیم مرتبط با گزینش داده وجود ندارد. ادبیات این حوزه نشان می‌دهد که اصطلاحات گزینش داده، گزینش دیجیتال^۱، حفاظت دیجیتال^۲، مدیریت داده^۳، نظارت بر داده^۴ گاهی به جای هم استفاده می‌شوند و تمایزات مفهومی روشی میان آنها وجود ندارد. در جدول ۳، برخی از دیدگاه‌های نویسندگان و صاحب‌نظران در زمینه وجود همپوشانی و تفاوت مفهومی میان این اصطلاحات ارائه شده است.

همان‌طور که دیدگاه‌های برخی از نویسندگان در جدول (۳) نشان می‌دهد، علیرغم اذعان به تفاوت‌های مفهومی، اما تمایزات و مرزهای روشی میان این مفاهیم وجود ندارد و گاهی به جای هم استفاده می‌شوند. با این حال، برخی از نویسندگان بر این عقیده‌اند که برای انجام هرگونه تحقیق درباره گزینش داده و نظریه‌سازی در این زمینه، مفاهیم اصلی و مرتبط باید به درستی و به روشنی تعریف شوند (از جمله، اسمیت، ۲۰۱۴).

¹ Digital Curation
² Digital Preservation
³ Data Management
⁴ Data Stewardship

⁵Giarlo

جدول ۵. مفاهیم گزینش، حفاظت و آرشیو (لورد و همکاران ۲۰۰۴)

گزینش	فعالیت مدیریت و ارتقای استفاده از داده از نقطه خلق آن برای اطمینان از اینکه داده برای هدف کنونی مناسب است و برای کشف و استفاده مجدد قابل دستیابی است. برای مجموعه داده‌های پویا، گزینش ممکن است به معنای غنی‌سازی مستمر یا بروزرسانی برای حفظ تناسب آن برای یک هدف مشخص باشد. بالاترین سطوح گزینش همچنین شامل حفظ پیوندها با یادداشت‌گذاری و دیگر مواد منتشر شده خواهد بود.
آرشیو	فعالیت گزینشی است که اطمینان می‌دهد داده به دقت انتخاب و ذخیره می‌شود و در دسترس قرار می‌گیرد و تمامیت فیزیکی و منطقی آن در طول زمان حفظ می‌شود، و شامل امنیت و اعتبار است.
حفاظت	فعالیتی در داخل آرشیو است که در آن فقره‌های خاص داده در طول زمان نگهداری می‌شوند تا اینکه بتوانند در دسترس قرار گیرند و بر اساس تغییرات ایجاد شده در فناوری مجدداً قابل درک باشند.

تحقیق جدید داده را از نوه‌هدف‌گذاری^۴ می‌کند (پالم و همکاران، ۲۰۱۳). مونز و رنر^۵ در تبیین تمایز بین گزینش داده و گزینش دیجیتال و حفاظت بیان می‌کنند که «گزینش داده به چالش حفاظت از اطلاعات دیجیتال در طول دوره تحقیق می‌پردازد به نحوی که معنی و سودمندی آن به عنوان درونداد بالقوه تحقیق دیگر حفظ شود» (مونز و رینر، ۲۰۱۱).

از طرف دیگر، گزینش داده فراتر از مدیریت داده عمل می‌کند و شامل خدمات اضافی به منظور حفاظت و ایجاد ارزش افزوده در داخل چرخه حیات پروژه تحقیق و بعد از اتمام آن است، به طور مثال استفاده مجدد از داده را میسر می‌سازد (پارتلو، سیمونز و کارلسون^۶، ۲۰۱۵). نظارت داده شامل فعالیت‌هایی است که مستقیماً به حفظ تمامیت و دسترسی به داده مربوط می‌شود و به اینکه چه داده‌ایی توسط خالقان ذخیره می‌شود، کجا ذخیره و حفظ می‌شود، چگونه توصیف می‌شود و چگونه می‌تواند کشف شود و دسترسی یابد و مجدداً استفاده شود، اطلاق می‌شود (یورک، گاتمن و برمن^۷، ۲۰۱۶). همترازی با فرایند تحقیق ویژگی اساسی است که گزینش داده را از نظارت داده متمایز می‌کند. نظارت داده درباره مدیریت منبع مشترک است که اغلب توسط یک شخص یا یک گروه معین تعبیه شده است (بیکر و یارمی^۸، ۲۰۰۹). نظارت داده برای حفاظت و دوام مجموعه‌های داده ضروری است، اما آن فقط یک کارکرد «مدیریت کردن داده» است که بر حفاظت ثابت و ایستای داده در طول زمان دلالت دارد (ژاکوب و ورلی^۹، ۲۰۰۹)؛ اما گزینش، با دسترس‌پذیری و استفاده از داده در آینده، ارتقاء، بسط و توسعه و بهبود محصولات داده برای استفاده مجدد در آن سوی اجتماع علمی در ارتباط است (پالم و همکاران، ۲۰۱۳).

(آبوت^۱، ۲۰۰۸). تعاریف متعددی برای گزینش دیجیتال ارائه شده است، سه مورد از آنها که بیشتر مورد توجه قرار گرفته‌اند، عبارتند از: ۱. اصطلاح گزینش دیجیتال به صورت فراینده برای اعمال مورد نیاز برای حفظ داده پژوهشی دیجیتال و دیگر مواد دیجیتالی در طول چرخه حیات کامل خود است که در طول زمان توسط نسل‌های کاربران کنونی و آینده استفاده می‌شود. فرایندهای آرشیو دیجیتال و حفاظت دیجیتال ملزومات این تعریف هستند، آن همچنین همه فرایندهای مورد نیاز برای خلق داده خوب، مدیریت و ایجاد ارزش افزوده برای داده در جهت تولید منابع جدید اطلاعات و دانش را شامل می‌شود» (بیگری^۲، ۲۰۰۴)، ۲. مفهوم چترمانندی که مفاهیم حفاظت دیجیتال، گزینش داده، مدیریت رکورد‌های الکترونیکی و مدیریت دارایی دیجیتال را شامل می‌شود (یاکل، ۲۰۰۷؛ لی و تیبو^۳، ۲۰۰۷؛ پول^۳، ۲۰۱۶)، ۳. عمل نظارت که قابلیت تولید مجدد و استفاده مجدد داده دیجیتال معتبر و دیگر دارایی دیجیتال را فراهم می‌کند (لی و تیبو، ۲۰۰۷).

تمرکز گزینش داده بر «نفع و سودمندی» داده برای پژوهش است و بر ماموریت سنتی کتابخانه‌های تحقیقاتی برای فراهم ساختن اطلاعات برای پشتیبانی از تحقیق و تولید دانش جدید تمرکز دارد. در مقابل، گزینش دیجیتال اصطلاحی است که عمدتاً با طیف گسترده‌ایی از مواد دیجیتالی مطابقت دارد. مرکز گزینش دیجیتال دیدگاه وسیعی از گزینش دیجیتال ارائه کرده است. گزینش دیجیتال اصطلاحی است که بیشتر با طیف گسترده‌ایی از مواد دیجیتالی مطابقت دارد و منحصرأ داده‌گرا نیست و نشان نمی‌دهد که چه چیزی گزینش می‌شود و ضرورتاً نشان نمی‌دهد که کدام اجتماعات می‌توانند به صورت هدفمند از فعالیت‌های گزینشی بهره ببرند. در مقابل، گزینش داده، مستقیماً به داده مرتبط است که توسط اجتماعات پژوهشی تولید و استفاده می‌شود و استفاده مجدد از داده را تسهیل می‌بخشد و برای برطرف نمودن نیازهای

⁴ Repurposing

⁵ Munoz & Renear

⁶ Partlo, Symons and Carlson

⁷ York, Gutmann and Berman

⁸ Baker and Yarmey

¹ Abbott

² Lee, Tibbo and Schaefer

³ Poole

مدل‌های چرخه حیات و چهارچوب‌های گزینش و مدیریت داده پژوهشی

استفاده از مدل‌های چرخه حیات برای هدایت فرایند گزینش و مدیریت داده‌های پژوهشی می‌تواند برای خالقان/تولیدکنندگان داده، آرشیویست‌ها، کتابداران و کاربران مفید باشد. مدل‌های چرخه حیات داده، ساختاری برای سازماندهی وظایف و فعالیت‌های مرتبط با گزینش و مدیریت داده در داخل یک پروژه یا سازمان فراهم می‌کنند. آنها همچنین ابزاری برای برقراری ارتباط وظایف گزینش داده با مخاطب مورد نظر از قبیل محققان، مدیران داده، گزینش‌کنندگان، متخصصان مخزن، کتابداران و مدیران پروژه که مسئول سازماندهی داده در یک پروژه یا آزمایشگاه هستند، ارائه می‌کنند. مدل‌های متعددی برای گزینش داده یا چرخه حیات گزینش ایجاد شده‌اند و هدف همه آنها فراهم کردن راهنمای موثر برای قلمرو عمومی یا حوزه‌های خاص است. در این مقاله به همه این مدل‌ها پرداخته نمی‌شود، بلکه مدل‌هایی معرفی می‌شوند که در فهم و درک گزینش داده و خدمات آن جایگاه برجسته-ایی دارند و در توسعه و خلق مدل‌های دیگر گزینش داده نقش مهمی داشته و به عنوان مرجع محسوب می‌شوند.

کارلسون (۲۰۱۴) مدل‌های چرخه حیات را بر اساس خالق/تولیدکننده و استفاده‌کننده/مخاطب داده به سه نوع مدل متفاوت طبقه‌بندی می‌کند: مدل‌های چرخه حیات فردمحور، سازمان‌محور و اجتماع‌محور. مدل‌های فردمحور برای یک پروژه خاص و اغلب در سطح انتزاعی نبوده ولی شامل اطلاعات جزئی مرتبط با پروژه هستند. مدل‌های سازمانی توسط سازمان‌هایی که خدمات ارائه می‌کنند خلق می‌شوند. این سازمان‌ها شامل دانشگاه‌ها، کتابخانه‌ها، مخازن داده، ناشران و غیره هستند. یک نمونه از آنها، مدل سازمانی نمودار مدیریت داده پژوهشی دانشگاه آکسفورد^۲ است. مدل‌های چرخه حیات سازمان‌محور در مقایسه با مدل‌های چرخه حیات فردمحور مراحل مختلف چرخه حیات داده را شامل شده و بر پروژه خاصی متمرکز نیستند. مدل‌های نوع سوم مدل‌های اجتماع‌محور هستند. آنها برای پشتیبانی از نیازهای یک اجتماع تحقیق خاص توسعه یافته‌اند. یک نمونه از مدل چرخه حیات اجتماع‌محور، مدل چرخه حیات گزینش مرکز گزینش دیجیتال^۳ است (هیگنز^۴، ۲۰۰۸) که مراحل مختلف گزینش

داده را به صورت مفصل توصیف می‌کند. مدل مرکز گزینش دیجیتال پیرامون داده (اشیای دیجیتال یا پایگاه‌های داده) و اعمال شکل گرفته است. آن اعمال را به اعمال چرخه حیات کامل (اطلاعات توصیف و بازنمایی، برنامه‌ریزی حفاظت، مشارکت و مشاهده اجتماع، گزینش و حفاظت)، اعمال دائمی (مفهوم‌سازی، خلق و دریافت، ارزیابی و انتخاب، جذب، حفاظت، ذخیره، دسترسی، استفاده و استفاده مجدد، تبدیل) و اعمال موقتی (دفع، ارزیابی مجدد، انتقال) تقسیم‌بندی می‌کند (ویزیک و دورکو^۵، ۲۰۱۵).

اگرچه مدل‌ها و چهارچوب‌های زیادی برای چرخه حیات تحقیق و گزینش دیجیتال/داده طراحی شده‌اند، ولی تنها یک مدل گزینش داده پذیرفته‌شده توسط سازمان بین‌المللی استانداردسازی^۶ وجود دارد و آن مدل مرجع سیستم اطلاعات آرشیو باز^۷ است. مدل مرجع شامل دو مفهوم مدل‌سازی است: مدل کارکردی و مدل اطلاعاتی. مدل کارکردی به عنوان یک استاندارد برای گزینش داده پذیرفته شده است و چهارچوبی برای درک مفاهیم ضروری آرشیو برای گزینش و کنترل دسترسی بلندمدت داده و اطلاعات فراهم می‌کند (تئودوریدو و همکاران^۸، ۲۰۱۰). مدل کارکردی یک استاندارد و رویکردی برای تضمین ثبات و پایداری داده تحت گزینش فراهم می‌کند. مدل کارکردی دارای شش کارکرد کلیدی است: جذب، ذخیره، آرشیو، مدیریت داده، اجرا، برنامه‌ریزی حفاظت و دسترسی. مدل کارکردی مجموعه هسته‌ای از مکانیسم‌هایی را توصیف می‌کند که باید در داخل مراکز داده، مخازن و آرشیوها ترکیب شوند تا بتوانند وظایف خود را در زمینه حفاظت بلندمدت داده و در دسترس ساختن آن برای کاربران تحقق بخشند. برخی از چهارچوب‌ها و مدل‌های گزینش داده بر پایه مدل مرجع توسعه یافته‌اند، از قبیل سکوی حفاظت دیجیتال تسینگو^۹ (ما و همکاران^{۱۰}، ۲۰۰۸) و مدل تعدیل‌شده جیسک^{۱۱} (بیگری، ۲۰۰۴). با این حال، مدل‌های گزینش داده متعددی وجود دارند که بر اساس مدل مرجع نیستند، اما سعی کرده‌اند که قابلیت ثبات داده را در بلندمدت تضمین کنند (لاگتون و پلسی^{۱۲}، ۲۰۱۳).

در میان بسیاری از مدل‌هایی که ارائه شده است، هشت مدل

⁵ Wissik and Durco

⁶ International Organization for Standardization (ISO)

⁷ Open Archival Information System (OAIS) Reference Model

⁸ Theodoridou et al

⁹ Tsinghua Digital Preservation Platform (THDP)

¹⁰ Ma et al

¹¹ JISC modified OAIS functional model

¹² Laughton and Plessi

جدول ۶. مدل‌های گزینش داده و ابعاد آنها

مدل	سازمان/نویسنده/عمومی یا حوزه خاص	ابعاد / مراحل چرخه حیات
<i>Digital Curation Centre (DCC) Lifecycle model</i>	مرکز گزینش دیجیتال / در بافت یک اجتماع علمی	اعمال کامل چرخه حیات شامل برنامه‌ریزی حفاظت، اطلاعات توصیفی و بازنمایی، گزینش و حفاظت، اعمال موقتی: مفهوم‌سازی، خلق، دسترسی و استفاده، ارزشیابی و انتخاب، حفظ، جذب کردن، عمل حفاظت، ارزشیابی مجدد، ذخیره، دسترسی و استفاده مجدد، انتقال
<i>Data Documentation Initiative(DDI) Combined Life Cycle Model</i>	اتلاف مستندسازی داده / در بافت یک اجتماع علمی	مطالعه مفهوم، گردآوری داده، پردازش داده، آرشیو داده، توزیع داده، کشف داده، تحلیل داده، هدف‌گذاری مجدد.
<i>Australian National Data Service(ANDS) Data Sharing Verbs</i>	توسط دولت استرالیا از طریق برنامه راهبردی زیرساختار تحقیق مشارکتی ملی / در بافت سازمانی	افعال اشتراک گذاری داده <i>ANDS</i> شامل خلق، ذخیره، توصیف، شناسایی، ثبت، کشف، دسترسی و بهره برداری است.
<i>DataONE Data Lifecycle</i>	بنیاد ملی علوم / داده محیطی و زمین	مدل دیتاوان از افعالی مشابه چرخه حیات داده <i>ANDS</i> استفاده می‌کند که شامل برنامه‌ریزی، گردآوری، اطمینان یافتن، توصیف، حفاظت، کشف، ترکیب، و تحلیل است.
<i>UK Data Archive Data Lifecycle</i>	شورای تحقیق اجتماعی و اقتصادی بریتانیا / داده علوم اجتماعی و اقتصادی	خلق داده، پردازش داده، تحلیل داده، حفاظت داده، دسترسی به داده، و استفاده مجدد از داده.
<i>Capability Maturity Model for Scientific Management Data Research360</i>	کراوستون و کین ^۱ (۲۰۱۱) / در بافت یک پروژه تحقیق	فراهم آوری داده، پردازش و تضمین کیفیت؛ توصیف و بازنمایش داده؛ توزیع داده؛ و حفاظت / خدمات مخزن.
<i>Institutional Research Lifecycle Concept</i>	دانشگاه بث / در بافت یک پروژه تحقیق	برنامه‌ریزی و طراحی، گردآوری و جذب، تفسیر و تحلیل، مدیریت و حفاظت، انتشار و نشر، کشف و استفاده مجدد

براون (۲۰۰۸) خاطر نشان می‌کنند که گزینش، حفظ و آرشیو داده دیجیتال نقش و مسئولیت مهمی برای کتابخانه‌ها است؛ بنابراین لزوم تدوین برنامه‌های آموزشی در گزینش داده علمی بسیار روشن است. بسیاری از نویسندگان به شکاف موجود در آموزش دانشجویان علوم کتابداری و اطلاع‌رسانی در زمینه گزینش داده اشاره کرده‌اند (یاکل، ۲۰۰۷؛ کورال^۸، ۲۰۱۰؛ لیتل^۹، ۲۰۱۱؛ هیدورن، ۲۰۱۱؛ براون^{۱۰}، ۲۰۰۹). نویسندگان متعددی به بررسی برنامه‌های آموزشی که با کمک‌های مالی موسسه موزه و خدمات کتابخانه^{۱۱} در دانشگاه ایلینویز، دانشگاه کارولینای شمالی و دانشگاه آریزونا^{۱۲} راه‌اندازی شده بودند، پرداخته‌اند (یاکل، ۲۰۰۷؛ رای، ۲۰۰۹؛ فولتون و همکاران، ۲۰۱۱). حرفه‌مندان کنونی کتابخانه نیازمند آموزش برای یادگیری مهارت‌های جدید در مواجهه با گزینش و مدیریت داده‌های پژوهشی هستند. در جدول ۸، برخی از دانشگاه‌ها به همراه توصیف مختصری از دوره‌هایشان اشاره شده است.

چرخه حیات تحقیق و مدل گزینش دیجیتال مورد توجه ذینفعان گزینش دیجیتال بوده است: مدل چرخه حیات گزینش مرکز گزینش دیجیتال^۱، مدل آرمانی چرخه حیات فعالیت تحقیق علمی^۲، مدل چرخه حیات ترکیبی^۳، افعال به اشتراک‌گذاری داده^۴، چرخه حیات داده *DataONE*^۵، چرخه حیات تحقیق سازمانی *Research360*^۶، مدل بلوغ توانمندی برای مدیریت داده علمی^۷ و چرخه حیات داده آرشیو داده بریتانیا. در جدول ۶، مشخصات و ابعاد یا مراحل چرخه حیات مدل‌های گزینش داده مذکور ارائه شده است. علاوه بر مدل‌های مذکور، چهارچوب‌هایی در سطح جهانی، ملی و دانشگاه برای پشتیبانی و اجرای گزینش داده توسعه یافته‌اند. در جدول ۷، برخی از این چهارچوب‌ها به همراه مشخصات‌شان ارائه شده است.

آموزش گزینش و مدیریت داده پژوهشی - سوان و

¹ Digital Curation Center Curation Lifecycle Model

² I2S2 Idealized Scientific Research Activity Lifecycle Model

³ DDI Combined Life Cycle Model

⁴ ANDS Data Sharing Verbs

⁵ DataONE Data Lifecycle

⁶ Research360 Institutional Research Lifecycle

⁷ Capability Maturity Model for Scientific Data Management

⁸ Corral

⁹ Little

¹⁰ Brown

¹¹ Institute of Museum and Library Services (IMLS)

¹² University of Illinois, the University of North Carolina, and the University of Arizona

جدول ۷. چهارچوب‌های گزینش و مدیریت داده‌های پژوهشی

مدل	نویسنده/محقق	روش‌شناسی	توصیف طرح
گزینش داده در سیستم داده جهانی: چهارچوب پیشنهادی	لاکتون و پلسی ^۱ ، ۲۰۱۳	از مطالعه موردی برای امکان‌سنجی توسعه چهارچوب گزینش داده برای داده جهانی استفاده شد. چهار مورد با استفاده از نمونه گیری با حداکثر تنوع انتخاب و مصاحبه عمیق صورت گرفت.	این چهارچوب پیشنهادی عمدتاً بر اساس مدل کارکردی سیستم اطلاعات آرشیو باز است و برای داده آنالوگ و دیجیتال مناسب است.
ظهور رویکرد میکرو خدمات ^۱ برای زیرساختار گزینش دیجیتال	آبرامز، کوزنه و لوی ^۱ ، ۲۰۱۰	بر اساس مفهوم میکرو خدمات زیرساختار گزینش دیجیتال با چهار لایه در کتابخانه-های دانشگاه کالیفرنیا تهیه شد.	برنامه دانشگاه کالیفرنیا ارائه یک محیط گزینشی بود که در دامنه جامع باشد اما نسبت به خط‌مشی‌های محلی و رویه‌ها و ضرورت تغییر فناوری و انتظارات کاربر منعطف باشد. برای دستیابی به این هدف، دانشگاه عقیده دارد که باید از مرکزیت مخزن گزینش به عنوان یک مکان کاسته شود. رویکرد جدید دانشگاه برای زیرساختار گزینش دیجیتال بر اساس ایده واگذاری کارکردهای ضروری به مجموعه‌ای میکرو خدمات مستقل، اما میانگین‌پذیر بود که ارزش‌ها و راهبردهای گزینش را در بر می‌گیرد. چون هر یک از این خدمات کوچک هستند، آنها به راحتی توسعه، استقرار، حفظ و ارتقاء می‌یابند.
مدلی برای تحقیق گزینش داده در علوم خرد: موسسه کریستال مایع دانشگاه ایالت کنت	شاکری ^۱ ، ۲۰۱۴	از مطالعه موردی برای بررسی ملزومات گزینش داده در رشته خاص و رویکردها و رویه‌های محققان در شاخه‌ای از علم خرد استفاده شده است. از مصاحبه و پرسشنامه برخط برای گردآوری داده تجربی در راستای اهداف این مطالعه استفاده شد.	این مدل جهان‌بینی دانشمندی را که در این مطالعه شرکت کرده بودند منعکس می‌کند- شیوه‌هایی که درباره مدیریت داده فکر می‌کنند. ابرمقوله «دنیای داده دانشمندان کریستال مایع» دارای سه خوشه کوچکتر مفاهیم مرتبط بود: «رویه‌های مدیریت داده»، «رویکردهای محققان درباره خصوصیات داده»، و «رویکردهای محققان درباره مدیریت داده». هر خوشه خود به کدهای مرتبط جزئی‌تری تقسیم شده است. هر کد در داخل خوشه پدیده و مفاهیم خاص داخل داده گردوری شده برای این مطالعه را بازنمایی می‌کند.
رویه‌های داده مرتبط، انواع، و کارکردهای گزینش: چهارچوب استخراج شده تجربی	کراگین، پالمرو و چائو ^۱ ، ۲۰۱۰	این چهارچوب از طریق مطالعات تجربی رویه‌های داده در علوم زمین و زیست بدست آمده و با استفاده از موردهای استفاده به عنوان خدمات گزینش توسعه یافته است.	یک چهارچوب مفهومی عمومی از گزینش داده است که روابط و وابستگی‌های میان رویه‌های داده علمی، انواع داده تولیدشده و استفاده‌شده و فعالیت‌های گزینش مرتبط را ترسیم می‌کند. چهارچوب شامل سه رده عمده است: داده، رویه‌های داده، و گزینش. رده داده شامل انواع داده و معیارهای داده است. رویه‌های داده شامل فرایندهای تحقیق و فعالیت‌های مرتبط به کار محققان با داده است؛ و خدمات گزینش شامل کارکردهای و فعالیت‌های مرتبط است.
مدیریت داده تحقیق: چهارچوب مفهومی	پاتل ^۱ ، ۲۰۱۶	به ارائه یک چهارچوب مفهومی برای مدیریت داده پژوهشی در سطح سازمانی می‌پردازد.	چهارچوب به صورت روشن گردش کار چرخه حیات داده را در مراحل متعدد از خلق داده تا ذخیره، سازماندهی و اشتراک‌گذاری ارائه می‌کند. آن همچنین تلاش دارد که موضوعات حیاتی مدیریت داده پژوهشی را از قبیل حریم خصوصی داده، امنیت داده، حق مولف و مجوز داده مورد توجه قرار دهد. چهارچوب می‌تواند به موسسات در مدیریت چرخه حیات داده پژوهشی به شیوه موثر و سودمند کمک کند.
سیستم پشتیبانی مدیریت داده پژوهشی سازمانی (RDMSS)	ماکانی ^۱ ، ۲۰۱۵	این مطالعه یک بررسی مفهومی است.	چهارچوب ارائه‌شده به هیچ نوع فناوری خاصی وابسته نیست و با هر نوع سیستم ارتباطی رایانه‌ای که دانشگاه‌ها و محققان در بافت مدیریت داده و کشف دانش، استفاده و استفاده مجدد استفاده می‌کنند، بکار می‌رود. چهارچوب سیستم پشتیبانی مدیریت داده پژوهشی سازمانی به سه لایه تقسیم می‌شود: لایه رابط کاربر گرافیکی، لایه هسته کاربرد و لایه مدیریت داده و اشتراک‌گذاری. لایه رابط کاربر گرافیکی شامل همه عناصر مسئول برای تعامل کاربر و مصورسازی داده است. لایه هسته کاربرد همه الگوریتم‌های اجرای کاربر داده را که توسط درخواست‌های سطح بالا کنترل شده است، پوشش می‌دهد. لایه مدیریت داده و اشتراک‌گذاری دانش جنبه کلیدی چهارچوب سیستم پشتیبانی مدیریت داده پژوهشی سازمانی است. آن از دو لایه مجزا ساخته شده است، لایه‌های مدیریت داده و توزیع/اشتراک‌گذاری دانش. لایه مدیریت داده شامل همه فایل‌های داده برای اشتراک، ابزارهای مدیریت داده و نیز اطلاعات رشته خاص و نوع-بافت بین رشته‌ای است. لایه توزیع/اشتراک‌گذاری دانش شامل همه عناصر مسئول برای ارتباطات شبکه و مدیریت داده سطح پایین، به طور مثال ابزارهای استناد داده است.

نیازهای «سیل داده»^۲ کافی است یا نه. نتایج نشان داد که بسیاری از نویسندگان به شکاف موجود در آموزش گزینش داده برای دانشجویان کتابداری اشاره کرده‌اند. از ۵۲ مدرسه کتابداری در آمریکای شمالی، ۱۶ موسسه دوره‌های گزینش داده ارائه کرده‌اند. رشد تعداد مدارس ارائه‌کننده گزینش داده

هاریس-پیرس و لیو^۱ (۲۰۱۲) در مقاله خود به مطالعه کفایت آموزش گزینش داده در مدارس علوم کتابداری و اطلاع‌رسانی در امریکای شمالی پرداخته‌اند. هدف آنها این بود که ببینند آیا تعداد دوره‌های گزینش داده ارائه‌شده برای پرداختن به

² Data deluge¹ Harris-Pierce, Liu

جدول ۸. دوره‌های آموزش گزینش داده

نویسنده	دانشگاه	توصیف دوره آموزشی
یاکل، ۲۰۰۷	مدرسه اطلاعات دانشگاه میشیگان	این مدرسه تخصص جدیدی با عنوان «حفاظت اطلاعات» که بر گزینش دیجیتال تمرکز دارد، ارائه می‌کند. با استفاده از جایزه‌ای از بنیاد ملی علوم، سمیناری در سطح دکتری درباره گزینش داده برای پرداختن به اصول بنیادی، ملزومات و فنون گزینش داده توسعه یافته است. دوره به سه بخش تقسیم می‌شود: عناصر مورد نیاز برای گزینش داده پژوهشی، توسعه مخزن داده و مکانیزم‌های انفرادی و عناصر فنی مخزن داده. برنامه‌ای در سطح کارشناسی ارشد در مدیریت اطلاعات دیجیتال ^۱ توسعه داده است. هدف این برنامه توانمند کردن طیفی از حرفه‌مندان اطلاعاتی در جهت کسب سطح سواد ضروری در زمینه فناوری‌های پیشرفته مورد نیاز برای گزینش به خصوص برنامه‌های کاربردی مخزن و سیستم‌های مدیریت محتوی وب محور بود. شش دوره در برنامه درسی توصیف شده است: مقدمه‌ای بر مجموعه‌های دیجیتال، مقدمه‌ای بر فناوری کاربردی، مدیریت محیط دیجیتال، حفاظت مجموعه‌های دیجیتال، مجموعه‌های دیجیتال پیشرفته، و معماری <i>DigIn</i> .
فولتون و همکاران ^۱ ، ۲۰۱۱	مدرسه علوم کتابداری و منابع اطلاعات دانشگاه آریزونا ^۱ با کمک مالی موسسه خدمات کتابخانه و موزه	این برنامه بر گزینش داده علمی متمرکز است اما بعداً داده علوم انسانی را نیز شامل شد. این برنامه همچنین بر یادگیری تجربی تاکید دارد. برنامه گزینش داده ایلینویز دارای شش رده است: بنیادهای گزینش داده، حفاظت دیجیتال، تحلیل سیستم‌ها، تجربه میدانی و نیز دو رده مورد نیاز برای همه دانشجویان کارشناسی ارشد علوم کتابداری و اطلاعات با عناوین سازماندهی و دسترسی اطلاعات و کتابخانه‌ها، اطلاعات و جامعه. موسسه گزینش داده در داخل دانشکده اطلاعات ^۱ ایجاد شده است.
فولتون و همکاران، ۲۰۱۱؛ رای، ۲۰۰۹.	مدرسه تحصیلات تکمیلی علم اطلاعات و کتابداری ^۱ دانشگاه ایلینویز با استفاده از کمک مالی موسسه خدمات کتابخانه و موزه دانشگاه تورنتو	موسسه گزینش داده در داخل دانشکده اطلاعات ^۱ ایجاد شده است.
رای، ۲۰۰۹		

نشان می‌دهد که مدارس کتابداری در حال پاسخگویی به تقاضای حرفه‌مندان گزینش داده هستند. بیشتر مدارس کتابداری باید گزینش داده را به برنامه درسی خود اضافه کنند. هاریس و لیو معتقدند که مدارس کتابداری که در حال حاضر دوره‌های گزینش داده ارائه می‌کنند باید برای تعیین اهداف عالی دوره و دستاوردهای یادگیری با یکدیگر همکاری کنند.

مطالعه نیازها و رویه‌های گزینش و مدیریت داده پژوهشی

مطالعه نیازها و رویه‌های گزینش و مدیریت داده‌های پژوهشی محققان، اعضای هیئت علمی یکی از ملزومات اولیه و اصلی ایجاد زیرساختار و خدمات گزینش و مدیریت داده‌های پژوهشی موفق در سازمان‌ها و دانشگاه‌هاست. از آنجایی که گزینش و مدیریت داده‌های پژوهشی یک فرایند وابسته به بافت است، لذا آگاهی از رویه‌ها و نیازهای تولیدکنندگان و محققان برای تهیه و ارائه خدمات داده بسیار مفید و ضروری خواهد بود. یافته‌های این نوع مطالعات نشان می‌دهد که محققان چگونه داده‌های پژوهشی خود را تولید، مدیریت، ذخیره و نشر می‌دهند و به چه خدمات و زیرساختاری در جهت مدیریت بهتر این داده‌ها نیاز دارند. برای این منظور محققان از روش‌ها و رویکردهای مختلفی برای مطالعه رویه‌های گزینش داده استفاده کرده‌اند که در جدول ۹، به برخی از این مطالعات

اشاره شده است. همان‌طور که مشخص شد از روش‌های مختلفی مانند پیمایش، مصاحبه، تطبیقی و ترکیبی از روش‌های کمی و کیفی برای بررسی رفتارها، نگرش‌ها، نیازها، آگاهی محققان در زمینه فرایند گزینش و مدیریت داده‌های پژوهشی استفاده شده است. برخی از یافته‌های ارزشمند این مطالعات عبارتند از: محققان انواع مختلف داده در اندازه‌های متفاوت تولید می‌کنند. اغلب محققان از ابرداده و استانداردهای ابرداده استفاده نمی‌کنند. بالاتکلیفی و تردید در زمینه مالکیت داده-های پژوهشی وجود دارد. مدیریت داده‌ها بر اساس روش‌های قدیمی و شهودی انجام می‌شود و بر اساس تجربیات برتر نیست. محققان اغلب از خدماتی که در مورد مدیریت داده ارائه می‌شود ناآگاه هستند. رفتارها و نیازهای ذخیره داده محققان در طول چرخه حیات تحقیق متفاوتند. در زمینه اشتراک‌گذاری داده‌های پژوهشی و استفاده مجدد از آنها میان محققان رشته‌های مختلف تفاوت وجود دارد و این برخاسته از فرهنگ‌های داده چندگانه در میان رشته‌هاست. اغلب محققان نیازمند آموزش درباره نحوه انبار و ذخیره داده در مخازن داده هستند. بیشتر شیوه‌های حفظ و نگهداری داده‌ها محلی و توسط خود محققان صورت می‌گیرد. در کل، ملزومات و نیازهای گزینش و مدیریت داده محققان

اشاره شده است. همان‌طور که مشخص شد از روش‌های مختلفی مانند پیمایش، مصاحبه، تطبیقی و ترکیبی از روش‌های کمی و کیفی برای بررسی رفتارها، نگرش‌ها، نیازها، آگاهی محققان در زمینه فرایند گزینش و مدیریت داده‌های پژوهشی استفاده شده است. برخی از یافته‌های ارزشمند این مطالعات عبارتند از: محققان انواع مختلف داده در اندازه‌های متفاوت تولید می‌کنند. اغلب محققان از ابرداده و استانداردهای ابرداده استفاده نمی‌کنند. بالاتکلیفی و تردید در زمینه مالکیت داده-های پژوهشی وجود دارد. مدیریت داده‌ها بر اساس روش‌های قدیمی و شهودی انجام می‌شود و بر اساس تجربیات برتر نیست. محققان اغلب از خدماتی که در مورد مدیریت داده ارائه می‌شود ناآگاه هستند. رفتارها و نیازهای ذخیره داده محققان در طول چرخه حیات تحقیق متفاوتند. در زمینه اشتراک‌گذاری داده‌های پژوهشی و استفاده مجدد از آنها میان محققان رشته‌های مختلف تفاوت وجود دارد و این برخاسته از فرهنگ‌های داده چندگانه در میان رشته‌هاست. اغلب محققان نیازمند آموزش درباره نحوه انبار و ذخیره داده در مخازن داده هستند. بیشتر شیوه‌های حفظ و نگهداری داده‌ها محلی و توسط خود محققان صورت می‌گیرد. در کل، ملزومات و نیازهای گزینش و مدیریت داده محققان

جدول ۹. روش‌ها و هدف مطالعات نیازها و رویه‌های گزینش و مدیریت داده‌های پژوهشی

روش/هدف	جامعه مورد بررسی	محقق/نویسنده
پیمایش / شناسایی نیازها و رویه‌های گزینش و مدیریت داده‌های پژوهشی	محقق / هیئت علمی در سطح یک دانشگاه یا چند دانشگاه	آکرز و دوتی ^۱ ، ۲۰۱۳؛ بیگری، بیگنه و رولند ^۱ ، ۲۰۰۹؛ هنتی و همکاران ^۱ ، ۲۰۰۸؛ پارهام و همکاران ^۱ ، ۲۰۱۲؛ وایتمایر و بوک و ساتن ^۱ ، ۲۰۱۵؛ تنویپر و همکاران، ۲۰۱۱؛ آکرز و دوتی، ۲۰۱۳
شیوه کمی و تطبیقی / مطالعه رویه‌های مدیریت داده بر اساس تفاوت‌های رشته‌ای	محققان رشته‌های مختلف دانشگاهی	استینه‌هارت و همکاران، ۲۰۱۲
پیمایش / تعیین آمادگی محققان برای روبرو شدن با نیازهای برنامه مدیریت داده بنیاد ملی علوم	محققان بنیاد ملی علوم	اسکاراموزینو و همکاران، ۲۰۱۲
پیمایش / مطالعه رفتارها و نگرش و آگاهی و نیازها درباره گزینش و مدیریت داده	اعضای هیئت علمی دانشکده علم و ریاضی دانشگاه پلی تکنیک کالیفرنیا	کارلسون، ۲۰۱۲؛ کراگین و همکاران، ۲۰۱۰؛ استولیا و همکاران ^۱ ، ۲۰۱۳؛ پترز و دریاد ^۱ ، ۲۰۱۱
رویکرد مصاحبه عمیق با گروه‌های کانونی / شناسایی نیازها و رویه‌های مدیریت داده‌های پژوهشی	محققان دانشگاه هاوستون	وان تیول و میشالک ^۱ ، ۲۰۱۵؛ بیگری، بیگری و رولند، ۲۰۰۹؛ شاکری، ۲۰۱۴؛ زیلسکی و لورنزو، ۲۰۱۲
پروتکل‌های مصاحبه / شناسایی نیازها و رویه‌های مدیریت داده-های پژوهشی	محققان دانشگاه	لی و استولیا ^۱ ، ۲۰۱۷؛ لی و استولیا، ۲۰۱۴؛ مارتینز- یوریب و مک دونالد ^۱ ، ۲۰۰۹
ترکیب رویکردهای پیمایش و مصاحبه / شناسایی نیازها و رویه-های مدیریت داده‌های پژوهشی	محققان دانشگاه پوردو	مکلور و همکاران ^۱ ، ۲۰۱۴
مصاحبه / توسعه سیستم مدیریت داده دیجیتال جدید بر اساس شناسایی رویه‌ها و نیازهای محققان	کارکنان مخازن سازمانی در دانشگاه‌های مختلف	دیکن ^۱ ، ۲۰۱۲
مصاحبه / بررسی رویه‌های گزینش داده در مخازن سازمانی مصاحبه، گروه‌های کانونی و یادداشت‌های روزانه / بررسی ملزومات خدمات مخزن دیجیتالی برای مدیریت داده گروه کانونی / مطالعه ماهیت مجموعه داده، رویه‌های مدیریت داده، و نیازهای محققان در پشتیبانی از گزینش داده، اشتراک گذاری، امکان سنجی پذیرش ابزار مصاحبه پروفایل گزینش داده دانشگاه پوردو	هیئت علمی، دانشمندان و دستیاران پژوهش دانشگاه یالتی کلورادو	ویت و همکاران ^۱ ، ۲۰۰۹
مصاحبه‌های نیمه ساختاریافته / شناسایی مشترکات رویکردهای مدیریت داده، برای مستند کردن زیرساختار مدیریت داده کانونی و برای بررسی تجربیات و نگرش‌ها نسبت به اشتراک‌گذاری داده پژوهشی	محققان رشته‌های مختلف کشاورزی	
مصاحبه / شناسایی نیازهای هیئت علمی درباره کشف، دسترسی، حفاظت، و استفاده مجدد از داده پژوهشی خود	اعضای هیئت علمی دانشگاه پوردو	

در بین رشته‌ها و دانشگاه‌ها متفاوت است. محققان علوم انسانی، اجتماعی و علوم طبیعی نه تنها مجموعه داده‌هایی خلق می‌کنند که در اندازه و محتوی متفاوت هستند، بلکه از فرهنگ‌های پژوهشی و اجتماعی با نگرش‌ها و تجربیات متفاوت در زمینه آرشیو و اشتراک‌گذاری داده برخوردار هستند؛ بنابراین گزینش داده موثر نیازمند درک تمایزات رشته‌ای و توسعه خدماتی است که برای اجتماعات محققان متفاوت دانشگاهی مناسب باشد.

پشتیبانی از گزینش و مدیریت داده‌های پژوهشی: نقش کتابخانه‌های دانشگاهی و کتابداران

مطالعات گذشته نشان داده‌اند که بسیاری از پژوهشگران از دانش لازم برای مدیریت موثر داده‌هایشان برخوردار نیستند

پشتیبانی از گزینش و مدیریت داده‌های پژوهشی: نقش کتابخانه‌های دانشگاهی و کتابداران

مطالعات گذشته نشان داده‌اند که بسیاری از پژوهشگران از دانش لازم برای مدیریت موثر داده‌هایشان برخوردار نیستند

¹ Macdonald and Martinez-Urbe

² Oxford's Scoping Digital Repository Services for Research Data Management

³ RIN's Patterns of Information Use and Exchange: Case Studies of Researchers in the Life Sciences

بریتانیا^۹ راهبرد خود را برای سال ۲۰۱۱-۲۰۱۴ منتشر کرد و راهبرد چهارم، مدل‌سازی نقش کتابخانه در مدیریت داده پژوهشی بود (کتابخانه‌های تحقیقاتی بریتانیا، ۲۰۱۱). در نوامبر ۲۰۱۱، با تأمین اعتبار اتحادیه اروپا، پروژه «فرصت‌هایی برای مبادله داده»، انجام شد و گزارشی راجع به فرصت‌هایی برای پشتیبانی از رکورد علمی بسیار ترکیب‌شده و به هم مرتبط ارائه کرد. نتایج نشان داد که مدل‌های نشر جدید برای پرداختن به نگرانی‌های خاص، نیازمند کتابخانه‌ها و مراکز داده هستند. ریلی^{۱۰} (۲۰۱۲) برای روشن ساختن نقش کتابخانه در پشتیبانی از مبادله داده، یک بررسی درباره کتابخانه‌ها در شبکه انجمن کتابخانه‌های تحقیقاتی اروپا^{۱۱} انجام داد و یافته‌ها نشان می‌دهد که محققان از کتابخانه‌های تحقیقاتی تقاضا دارند که از آنها در زمینه مدیریت داده به خصوص داده آرشیو شده پشتیبانی کنند. اوسوالد و استراسمن^{۱۲} (۲۰۱۲) در آلمان یک بررسی از موقعیت کنونی فعالیت‌ها در حوزه‌های گزینش دیجیتال و حفاظت دیجیتالی داده پژوهشی در ۱۱ رشته انجام دادند. آنها گزارش کردند که دانشمندان و مراکز داده پژوهشی به تخصص کتابداران درباره کاربرد استانداردهای ابر داده و فراهم‌آوری خدماتی در ارتباط با داده پژوهشی گزینش شده توسط محققان نیاز دارند.

درک کتابداران از نقش‌های خود در خدمات داده پژوهشی - جامعه کتابخانه و کتابداری نیز درباره نقش‌های بالقوه خود درباره ارائه خدمات داده پژوهشی در حال و آینده بحث کرده است. برای مثال، در سپتامبر ۲۰۰۶، انجمن کتابخانه‌های تحقیقاتی کارگروهی برای بررسی نقش کتابخانه‌های دانشگاهی و تحقیقاتی در مورد نظارت داده^{۱۳} دیجیتال علمی و مهندسی سازماندهی کرد. در می ۲۰۰۸، گروه کاری داده کتابخانه دانشگاه کرنل^{۱۴} (استینهارت و همکاران، ۲۰۰۸)، مرور کلی از چشم‌انداز کنونی و موضوعات پیرامون گزینش داده انجام داد و پنج پیشنهاد در زمینه گزینش داده و فعالیت‌های آن برای کتابخانه دانشگاه کرنل ارائه کرده است که عبارتند از: جستجو و تقویت مشارکت‌ها با دیگر سازمان‌ها، فراهم کردن امکان دسترسی به محققان کرنل در حوزه‌های متعدد؛ ارزیابی نیازهای محلی و توسعه زیرساختار محلی و خط‌مشی‌های مرتبط؛ ایجاد یک نیروی کار

رشته‌های مختلف، کتابخانه‌های دانشگاهی می‌توانند از مهارت و دانشی که در زمینه گردآوری، سازماندهی و مدیریت محتوی تحقیق دیجیتالی دارند استفاده کنند و به محققان در امر گزینش و مدیریت داده کمک کنند. چاودری (۲۰۱۰) اعتقاد دارد که در عصر حاضر کتابخانه‌ها باید به «داده‌ها به عنوان مجموعه و خدمات، به کتابداران به عنوان دانشمندان داده؛ و مراکز داده به عنوان قفسه‌های کتابخانه جدید» بنگرند.

در حال حاضر کتابداران دانشگاهی با توسعه و ارائه زیرساختار و خدمات لازم برای پشتیبانی از گزینش و مدیریت داده‌های پژوهشی در پردیس‌های دانشگاهی (هیدرون، ۲۰۱۱؛ موناسترسکی^۱، ۲۰۱۳؛ اولندورف و کوچ^۲، ۲۰۱۲؛ استار و همکاران^۳، ۲۰۱۲؛ تنویپر و همکاران، ۲۰۱۴، ۲۰۱۳؛ سوهنر، استیوز و وارد^۴، ۲۰۱۰؛ رزنیچک-زالن، آدامیک و مک-گینتی^۵، ۲۰۱۲؛ کوکس و پنفیلد^۶، ۲۰۱۴؛ گینوچ^۷، ۲۰۱۶) درگیر شده‌اند. نقش کتابخانه‌ها و کتابداران را می‌توان از دو منظر مورد بررسی قرار داد: درک محققان از نقش کتابخانه در خدمات گزینش و مدیریت داده‌های پژوهشی و درک کتابداران از نقش‌های خود در خدمات گزینش و مدیریت داده‌های پژوهشی.

درک محققان از نقش کتابخانه در خدمات گزینش و مدیریت داده‌های پژوهشی - کتابخانه‌های دانشگاهی

به لحاظ تاریخی به عنوان گزینش‌کنندگان منابع تحقیقات عمل کرده‌اند و با انتخاب، سازماندهی و فراهم کردن منابع مرتبط از تحقیق پشتیبانی کرده‌اند. در نتیجه، بسیاری عقیده دارند که کتابداران دانشگاهی با تجربه کسب‌شده از فهرست‌نویسی، نمایه‌سازی سنتی و مهارت‌های سازمانی در کنار توسعه، تاسیس و نگهداری مخازن سازمانی، طبیعی است که نقش‌اشان به عنوان گزینش‌کننده داده مطرح شود. گابریچ^۸ (۲۰۰۹) در گزارش «آخرین مایل: نقش‌های رابط در گزینش داده پژوهشی علم و مهندسی» این ایده را مطرح کرد که رابط‌های موضوعی کتابخانه می‌توانند نقش خود را در مورد ارائه خدمات گزینش داده برای محققان و موسسات گسترش دهند. در سال ۲۰۱۱، کتابخانه‌های تحقیقاتی

¹ Monastersky

² Olendorf & Koch

³ Starr et al

⁴ Soehner, Steeves and Ward

⁵ Reznik-Zellen, Adamick and McGinty

⁶ Cox & Pinfield

⁷ Gynoch

⁸ Gabridge

⁹ Research Libraries UK

¹⁰ Reilly

¹¹ LIBER, the Association of European Research Libraries

¹² Osswald and Strathmann

¹³ Data stewardship

¹⁴ Cornell University

امکان استفاده جهانی، در دسترس بودن، بازبودن و اشتراک‌گذاری داده پژوهشی را در کشورهای عضو ارتقا دهد. یکی دیگر از سازمان‌هایی که در امر مدیریت داده‌های پژوهشی در سطح بین‌المللی فعالیت می‌کند، کمیته داده برای علوم و فناوری^۴ است. این کمیته یک سازمان بین رشته‌ای است که توسط شورای بین‌المللی علم^۵ در سال ۱۹۶۶ راه‌اندازی شده است. هدف کمیته داده برای علوم و فناوری، پشتیبانی و بهبود مدیریت، کیفیت، قابلیت اعتماد و دسترس‌پذیری داده پژوهشی است. این کمیته در حال تقویت مدیریت داده پژوهشی است (کمیته، ۲۰۱۴).

در سطح اروپا، مدیریت و زیرساخت‌های داده پژوهشی از توجه زیادی برخوردار بوده است. در این زمینه طرح‌ها، توصیه‌ها و رهنمودهایی ارائه شده است. برای مثال، طرح اروپایی برون مرزی، مبادله دانش^۶، در سال ۲۰۰۵ راه‌اندازی شد. در حال حاضر پنج سازمان با آن همکاری می‌کنند: مرکز فناوری اطلاعات برای علوم در فنلاند^۷، کتابخانه تحقیق الکترونیکی دانمارک^۸، بنیاد تحقیق آلمان^۹، کمیته مشترک سیستم‌های اطلاعاتی بریتانیا^{۱۰} و سازمان سرف هلند^{۱۱}. مبادله دانش بر فعالیت‌های متعددی متمرکز است. نقشه راه علم اروپا^{۱۲} در سال ۲۰۱۳ توسط ساینس یوروپ^{۱۳} منتشر شد. ساینس یوروپ یک انجمنی از سازمان‌های تامین مالی تحقیق و سازمان‌های انجام تحقیق است که در سال ۲۰۱۱ تاسیس شد (کوردت^{۱۴}، ۲۰۱۴).

طرح‌ها و فعالیت‌های متعددی در حوزه مدیریت داده پژوهشی در آمریکا اجرا شده است. برای مثال، در مارس ۲۰۰۷، «چشم‌انداز زیرساخت‌های سایبری برای کشف قرن ۲۱» توسط بنیاد ملی علوم (۲۰۰۷) منتشر شد. علاوه بر این، نیروی کار خط‌مشی‌های داده مورد ملی علوم^{۱۵}، گزارشی درباره مدیریت و اشتراک‌گذاری داده پژوهش دیجیتال (بورد ملی علوم، ۲۰۱۱) منتشر و درباره اهمیت این موضوع بحث کرد. علاوه بر این، توصیه‌هایی توسط بنیاد ملی علوم^{۱۶} ارائه شده است.

با توان پرداختن به چالش‌های جدید گزینش داده و توسعه زیرساخت‌های سایبری؛ ایجاد یک گروه اجرایی گزینش داده و به رسمیت شناختن یک گروه کاری داده. سوان و براون (۲۰۰۸) سه نقش اصلی بالقوه برای کتابخانه‌ها در تحقیق داده‌بر شناسایی کرده‌اند: افزایش آگاهی محققان از داده؛ فراهم کردن آرشیو و خدمات حفاظت برای داده از طریق مخازن سازمانی در موسسه و توسعه حوزه حرفه‌ای جدید در قالب کتابداری داده. گولد^۱ (۲۰۰۷) مرور و تحلیلی از تئوری‌های کنونی درباره نقش‌های کتابداران و کتابخانه‌ها در خدمات داده ارائه کرده است. نویسنده پیشنهاد می‌کند که کتابداران می‌توانند در علم الکترونیکی نقش‌های زیر را برعهده بگیرند: انتخاب، فراهم‌آوری و اجازه‌نامه^۲ داده و مجموعه داده؛ خلق ابرداده برای کشف و توصیف مجموعه‌های داده؛ کمک به کاربران با یافتن داده مرتبط برای تحقیق‌اشان؛ کار با اجتماعات علمی و ناشران دانشگاهی برای مشورت و کمک به توسعه استانداردهای نشر و سیستم‌ها و تهیه مخازن بلندمدت برون‌داد علمی.

خط‌مشی‌ها و تجربیات جهانی و ملی برای گزینش و مدیریت داده پژوهشی

لزوم گزینش و مدیریت داده پژوهشی در سطح بین‌المللی، سازمان‌های تحقیقاتی، سازمان‌های تامین‌کننده بودجه تحقیقات و نیز موسسات علمی به رسمیت شناخته شده است و در این زمینه توصیه‌ها، اسناد راهبردی و رهنمودهایی منتشر کرده‌اند. طرح جهانی اتحاد داده پژوهشی^۳ (اتحاد داده پژوهشی، ۲۰۱۴) که در سال ۲۰۱۳ راه‌اندازی شد، به موضوع مدیریت داده پژوهشی می‌پردازد. هدف اتحاد داده پژوهشی، تشکیل پیوندهای اجتماعی و فنی است. چشم‌انداز اصلی اتحاد داده پژوهشی تسهیل اشتراک‌گذاری باز و مبادله داده پژوهشی توسط محققان و نوآوران است. این امر باید بدون در نظر گرفتن فناوری‌های استفاده‌شده در بین رشته‌ها و کشورها صورت گیرد.

سازمان همکاری و توسعه اقتصادی سند «اصول و رهنمودهای سازمان توسعه و همکاری اقتصادی برای دسترسی به داده پژوهشی با بودجه دولتی» را منتشر کرده است (اوا سی دی، ۲۰۰۷)، این سند تاثیر عمده‌ای بر مدیریت داده پژوهشی بر سطح ملی داشته است. به عنوان مثال، اصول این سازمان از کشورهای عضو خواست که باید رهنمودهایی را اجرا کنند که مکمل قوانین و خط‌مشی‌های موجود باشد و

⁴ Committee on Data for Science and Technology (CO-DATA)

⁵ International Council for Science (ICSU)

⁶ Knowledge Exchange

⁷ IT centre for Science in Finland

⁸ Denmark's Electronic Research Library

⁹ German Research Foundation

¹⁰ Joint Information Systems Committee UK

¹¹ Dutch SURF

¹² Science Europe Roadmap

¹³ SCIENCE EUROPE

¹⁴ Curdt

¹⁵ National Science Board

¹⁶ National Science Foundation (NSF)

¹ Gold

² Data Licensing

³ Research Data Alliance (RDA)

پژوهشی توسط دانشگاه‌های متعددی در سطح دنیا به رسمیت شناخته شده است. به خصوص در بریتانیا، آمریکا و نیز دانشگاه‌های استرالیا به صورت گسترده از خطمشی مدیریت داده استفاده می‌شود. برای مثال با مراجعه به صفحه‌خانگی دانشگاه‌های ادینبورگ^{۱۳}، دانشگاه آکسفورد^{۱۴}، دانشگاه تنس^{۱۵}، دانشگاه جان هاپکینز^{۱۶}، دانشگاه ملبورن^{۱۷}، دانشگاه نیوکاسل^{۱۸}، می‌توان به خطمشی مدیریت داده پژوهشی آنها دسترسی یافت. لازم به ذکر است که بیشتر خطمشی‌های داده دانشگاه‌های بریتانیا توسط مرکز گزینش دیجیتال تهیه شده‌اند.

رضائی شریف‌آبادی (۱۳۹۴) در مقاله خود با عنوان «مدیریت داده‌های پژوهشی: مطالعه موردی داده‌های زبانی» ضمن بیان اهمیت مدیریت داده‌های پژوهشی برای انجام تحقیقات با کیفیت و بیان تجربیات جهانی در این حوزه، به معرفی و تجربیات به دست آمده از تولید داده‌های زبانی و راه‌اندازی و مدیریت «مرجع دادگان زبان فارسی» پرداخته است.^{۱۹} نکاتی چون اشتراک‌گذاری، داده‌های زبانی در پایگاه‌های اشتراک داده معتبر، مشخص کردن وضعیت مالکیت معنوی و شرایط استفاده داده، تهیه مستندات مناسب برای توصیف داده و مشخص کردن مستندی که کاربران باید هنگام استفاده از داده به آن ارجاع دهند، استفاده از منابع اولیه مناسب برای تولید داده، ساختار بندی داده بر مبنای قالب‌های استاندارد و مشترک، کسب اطمینان از کیفیت داده با استفاده از روش‌های مختلف ارزیابی، ایجاد ابزارهای مناسب برای پردازش و نمایش داده و استفاده از ساختارهای مناسب برای فایل داده از جمله پیشنهاداتی هستند که برای مدیریت داده‌های زبانی در این مقاله ارائه شده است.

وزیری، نقشینه و نوروزی چاکلی (زودآیند، ۱۳۹۶) با بررسی اسناد بالادستی داخلی در زمینه اشتراک‌گذاری داده‌های پژوهشی از جمله نقشه جامع علمی کشور، برنامه ششم توسعه، برنامه پنجم و قوانین دسترسی به اطلاعات به این

در بریتانیا، در حوزه مدیریت داده پژوهشی طرح‌ها و سرمایه‌گذاران تحقیق متعددی فعال هستند، برای مثال، کمیته مشترک سیستم‌های اطلاعاتی، شورای تحقیق بریتانیا^۱ و مرکز گزینش دیجیتال^۲. برای مثال، شورای تحقیق بریتانیا «خطمشی و اصول اجرای کنترل تحقیق» (شورای تحقیق بریتانیا، ۲۰۱۳) را که بر حفاظت و مدیریت داده پژوهشی متمرکز است در سال ۲۰۰۹ منتشر کرد. از طرف دیگر، کمیته مشترک سیستم‌های اطلاعاتی^۳، گزارش «مدیریت و به اشتراک‌گذاری داده: راهنمای تجربیات برتر»^۴ برای محققان «جیسک، ۲۰۰۹» را در سال ۲۰۰۹ منتشر کرد. علاوه بر این، بسیاری از تامین مالی‌کنندگان تحقیق در بریتانیا خطمشی‌های داده منتشر کرده‌اند و از پژوهشگران برنامه مدیریت داده درخواست می‌کنند. برای مثال، شورای تحقیق محیط طبیعی^۵ اولین دست‌نامه خطمشی داده را در سال ۱۹۹۶ منتشر و در سال ۲۰۱۰ ویرایش کرد. به همین ترتیب، ولکام تراست^۶ یک خطمشی درباره مدیریت داده و اشتراک‌گذاری در سال ۲۰۱۰ منتشر کرده است (ولکام تراست، ۲۰۱۰).

اصول انجام مسئولانه تحقیق در استرالیا^۷ در سال ۲۰۰۷ با مشارکت شورای ملی تحقیق سلامت و پزشکی^۸، شورای تحقیق استرالیا و دانشگاه‌های استرالیا منتشر شد (شورای ملی تحقیق سلامت و پزشکی، ۲۰۰۷). این برنامه، رهنمودهایی برای مدیریت داده پژوهشی و مواد دست اول فراهم کرده و مسئولیت‌های موسسات و محققان را شناسایی می‌کند. طرح دیگر در استرالیا، خدمات داده ملی استرالیا^۹ است که طرح ملی پشتیبانی از مدیریت داده پژوهشی در سطح استرالیا محسوب می‌شود. این طرح با هدف تاسیس و فراهم کردن زیرساختار ملی داده پژوهشی راه‌اندازی شده است. هدف این طرح تضمین و پشتیبانی از همه داده‌های پژوهشی است تا به خوبی مدیریت شوند، قابل دسترس باشند تا مجدداً استفاده گردد (بورتون و همکاران^{۱۰}، ۲۰۱۲؛ گرونوگان و ترلور^{۱۱}، ۲۰۱۳؛ ترلور، ۲۰۰۹؛ ترلور و ویلکینسون^{۱۲}، ۲۰۰۸).

اهمیت و ضرورت اجرای خدمات گزینش و مدیریت داده‌های

¹³ University of Edinburgh

¹⁴ University of Oxford

¹⁵ University of Tennessee

¹⁶ Johns Hopkins University

¹⁷ University of Melbourne

¹⁸ University of Newcastle

^{۱۹} مرجع دادگان زبان فارسی در اسفندماه سال ۱۳۹۱ به عنوان بستری برای اشتراک-

گذاری داده‌های زبان فارسی در دسترس عموم کاربران وب با آدرس www.dadegan.ir راه‌اندازی شد. در این پایگاه اینترنتی هم‌اکنون تعداد ۴۳ داده معرفی شده است و تقریباً به همین تعداد داده‌های دیگر نیز شناسایی شده است که مراحل آماده‌سازی برای انتشار در مرجع را طی می‌کنند داده‌ها در قالب متنی، تصویری و صوتی هستند.

¹ Research Council UK (RCUK)

² Digital Curation Centre

³ Joint Information Systems Committee (JISC)

⁴ Best practice

⁵ Natural Environment Research Council (NERC)

⁶ Wellcome Trust

⁷ Australian Code for the Responsible Conduct of Research

⁸ National Health and Medical Research Council

⁹ Australian National Data Service

¹⁰ Burton et al

¹¹ Groenewegen & Treloar

¹² Treloar & Wilkinson

همکاران^۶، ۲۰۰۷)، پیچیدگی و ماهیت متنوع داده در کنار رویه‌های داده متنوع در میان رشته‌ها (بورگمن، ۲۰۱۲)، نبود گزینش‌کنندگان ماهر و رویه‌های گزینش داده توسعه‌نیافته (هیدرون و همکاران، ۲۰۰۷)، آگاهی فرهنگی، آموزش محققان و توسعه برنامه‌های آموزشی مهارت‌های کارکنان (کارلسون، ۲۰۱۲؛ پول، ۲۰۱۶؛ جانکه و آشر^۷، ۲۰۱۲؛ ویکت، هو و تومر^۸، ۲۰۱۲؛ کریمر، ۲۰۱۵)، امنیت و مالکیت داده (گیارلو، ۲۰۱۳؛ کریمر، ۲۰۱۵)، حقوق مالکیت فکری (استینهارت و همکاران، ۲۰۰۸؛ کریمر، ۲۰۱۵)، تضمین کیفیت داده (کریمر، ۲۰۱۵)، تضمین ثبات فنی، مالی و نیروی انسانی (سیمبولن^۹، ۲۰۱۳؛ کنان و مارکابواسکیت^{۱۰}، ۲۰۱۵؛ ویکت، هو و تومر، ۲۰۱۲؛ استینهارت و همکاران، ۲۰۰۸؛ شن و وارول^{۱۱}، ۲۰۱۳؛ وایت، ۲۰۰۹)، مدیریت خطر (پول، ۲۰۱۶)، حفظ حریم خصوصی (گاتمن و همکاران^{۱۲}، ۲۰۰۴؛ استینهارت و همکاران، ۲۰۰۸)، اشاره کرد. هر موسسه‌ای باید این چالش‌ها را در راه‌اندازی و اجرای گزینش و مدیریت داده پژوهشی مورد توجه قرار دهد.

بحث و نتیجه‌گیری

در عصر حاضر، داده پژوهشی به عنوان یک دارایی ارزشمند شناخته می‌شود و از پتانسیل بالایی برای تجمیع و استفاده مجدد جهت کشف دانش جدید و یافتن راه‌حل‌های قوی و ابتکارانه برای حل معضلات و چالش‌های پیچیده محیطی و جوامع انسانی برخوردار است. لکن، اگر با شیوه و رویکرد مناسب شناسایی، گردآوری و مدیریت نشود، در طول زمان بر اثر اتفاقات مختلف از بین خواهد رفت و تلاش برای تولید مجدد آن بسیار زمان‌بر، پرهزینه و در مواردی غیر ممکن خواهد بود؛ بنابراین، لازم است که داده‌های پژوهشی از تولید تا مصرف به نحو درست و مناسب مدیریت شوند. این مدیریت چرخه تولید تا مصرف داده‌های پژوهشی گزینش و مدیریت داده‌های پژوهشی نامیده می‌شود. گزینش داده فرایند مستمری است که تحت تاثیر مسائل پیچیده فنی، سازمانی و فرهنگی است. گزینش به عنوان یک پدیده، نیازمند پشتیبانی سازمانی از بالا به پایین و تغییرات فرهنگی پایین به بالا است که بر رفتار مدیریت داده‌های افراد تاثیر می‌گذارد. گزینش

نتیجه رسیدند که «در اسناد بالادستی کشور، نمی‌توان به صورت مشخص مساله اشتراک داده‌های علمی و پژوهشی را رصد کرد؛ اما می‌توان این مساله را به صورت کلی‌تر در قالب مدیریت پژوهش، توسعه شبکه علمی کشور، تبادل و به اشتراک‌گذاری رایگان اطلاعات به منظور ایجاد سامانه‌های اطلاعاتی و کاهش تولید و نگهداری اطلاعات تکراری، توسعه شبکه ملی اطلاعات و مراکز داده داخلی و آگاهی‌های علمی عمومی در نظر گرفت». مطالعه اسناد بالادستی نگارنده این مقاله هم نشان داد که هرچند در اسناد بالادستی به صورت صریح درباره گزینش و مدیریت داده‌های پژوهشی سخنی به میان نیامده است ولی بحث‌های مربوط به پرهیز از دوباره کاری در گردآوری داده‌های پژوهشی، امنیت و نگهداری از داده‌ها و اطلاعات و اشتراک‌گذاری آنها در جهت استفاده سایرین، توسعه شبکه ملی اطلاعات، توسعه مراکز داده داخلی از عمده مسائلی هستند که به صورت ضمنی بر اهمیت و ضرورت توسعه نظام‌های مدیریت داده‌های پژوهشی اشاره می‌کنند. لذا می‌توان با استناد به اسناد بالادستی طرح‌های عملی در زمینه اجرای گزینش و مدیریت داده‌های پژوهشی در تمامی سطوح از قبیل دانشگاه، پژوهشگاه، اجتماع علمی خاص به اجرا گذاشت.

چالش‌های گزینش و مدیریت داده‌های پژوهشی

علیرغم تلاش‌های بسیاری که در امر توسعه نظری و عملی گزینش و مدیریت داده انجام شده است، اما همچنان گزینش داده برای اجرای مناسب با چالش‌هایی روبرو است. از جمله می‌توان به موارد زیر اشاره کرد: کمبود بودجه، پرسنل و تجهیزات، نبود حمایت گسترده سازمانی (کریمر و همکاران^۱، ۲۰۱۲)، شکاف در دانش و مهارت‌های کارکنان کتابخانه و نبود اعتماد به نفس در نقش مورد انتظار (کورل و کنان و افضل^۲، ۲۰۱۳)، فقدان آموزش‌های لازم برای کتابداران، درخواست‌های نامعین محققان (براون^۳، ۲۰۱۰)، عدم علاقه‌مندی در هیئت علمی، فقدان ارزش‌های مشترک در پردیس و عدم تمایل کارکنان کتابخانه به بازآموزی برای مدیریت داده (سوهنر، استیوز و وارد^۴، ۲۰۱۰)، نبود آموزش، غفلت از تجربیات برتر برای مدیریت داده پژوهشی و نیز راهنمایی و پشتیبانی ناکافی (سالانز و لیک^۵، ۲۰۱۴)، ارتباطات ناکافی بین دانشمندان رشته‌های مختلف و مدیران مسئول برای اجرای سیستم مدیریت داده (موکچل و

⁶ Muckschel et al

⁷ Jahnke and Asher

⁸ Wickett, Hu and Thomer

⁹ Simbulan

¹⁰ Kennan & Markauskaite

¹¹ Shen & Varvel

¹² Guttman et al

¹ Creamer, Morales, Crespo, Kafel, and Martin

² Corral, Keenan, and Afzal

³ Brown

⁴ Soehner, Steeves & Ward

⁵ Sallans & Lake

داده فعالیتی نیست که تنها در کتابخانه‌ها یا هر نوع موسسه یا سازمان به انزوا کشیده شود، بلکه یک امر مشارکتی است و نیازمند همکاری و مشارکت تمامی عناصر تحقیق است. هرچند از دهه ۱۹۶۰ تلاش‌های مختلفی در راستای حفاظت و نگهداری دارایی‌های دیجیتال در سطح سازمانی و موسسات دانشگاهی آغاز گردید ولی استفاده از اصطلاح گزینش برای داده‌های پژوهشی از دهه ۱۹۹۰ رایج شد. گزینش داده، حوزه پژوهشی و آموزشی نسبتاً جدیدی در داخل علوم اطلاعات و دانش‌شناسی و حوزه‌های مرتبط از جمله، آرشیو، موزه و مدیریت سیستم‌های اطلاعاتی و داده است. این مفهوم از دهه ۲۰۰۰ در متون و ادبیات آموزشی و پژوهشی علوم کتابداری و اطلاع‌رسانی و آرشیو وارد شده و از سال ۲۰۰۴ به طور قابل توجهی رشد یافته است (وبر، پالمر و چائو، ۲۰۱۲). اصطلاح گزینش بر مدیریت دارایی‌های دیجیتال در طول چرخه حیات خود و استفاده مجدد از آن در بلندمدت به منظور کشف دانش جدید و حل مسائل علمی و پژوهشی تاکید دارد. لذا، گزینش دیدگاهی جامع‌نگر و ایده‌آل‌گرایانه دارد و در دامنه خود کارکردهای حفاظت و آرشیو را نیز شامل می‌شود. شمول کارکردهای حفاظت و آرشیو به عنوان اجزایی از گزینش مورد قبول واقع شده است ولی کماکان مناقشات و بحث‌ها در این زمینه در میان محققان حوزه‌های مختلف در جریان است.

علاوه بر این، اصطلاح گزینش در بافت‌های مختلف خارج از کتابخانه‌ها و موزه‌ها نیز استفاده می‌شود، به عنوان نمونه در علوم زیست‌شناسی به عنوان جزئی از مدیریت پایگاه داده و نشر و در برنامه علم الکترونیکی بریتانیا^۲ به عنوان یک جنبه از زیرساختار اطلاعاتی معرفی شده است (تیلور، ۲۰۰۱). لذا تنوع معنایی و از طرف دیگر وابستگی گزینش به بافت و فقدان واژگان هماهنگ و مشترک برای تبیین کارکردها و فعالیت‌های گزینش داده مانع ایجاد گفتمان سازنده در این حوزه شده است. این مسئله به اختلاف و ابهام در کاربرد اصطلاحات پایه، از قبیل گزینش داده، گزینش دیجیتال و علم داده منجر شده است. مجادلات درباره اینکه این حوزه چه باید نامیده شود و چگونه باید کار حرفه‌مندان داده از دیگر حرفه‌مندان اطلاعات متفاوت باشند، فضایی باز برای تحقیق و بررسی است (وبر، پالمر و چائو، ۲۰۱۲).

علیرغم تعاریف متعددی که برای گزینش داده ارائه شده است، ولی لورد و مک‌دونالد (۲۰۰۳) و مرکز گزینش دیجیتال تعریف

داده فعالیتی نیست که تنها در کتابخانه‌ها یا هر نوع موسسه یا سازمان به انزوا کشیده شود، بلکه یک امر مشارکتی است و نیازمند همکاری و مشارکت تمامی عناصر تحقیق است.

معنی گسترده‌تری از گزینش ارائه می‌کنند. به عقیده آنها گزینش «بافت گسترده‌تری از فقط آرشیو کردن را پوشش می‌دهد و همچنین مراقبت از رکوردها در داخل بافت و محیط علمی را در بر می‌گیرد. آنچه اصطلاح بر آن دلالت می‌کند - یک چرخه است». اجزای مورد تاکید تعریف لورد و مک‌دونالد شامل مدیریت و تقویت داده از نقطه خلق خود، اطمینان از مناسبت داده برای اهداف معاصر و در دسترس قرار دادن داده برای کشف و استفاده مجدد هستند. فعالیت‌های اصلی گزینش داده پژوهشی شامل کشف، شناسایی، انتخاب، فراهم‌آوری، تایید، تحلیل، مدیریت، آرشیو، انتشار و استناد می‌باشند. گزینش و فعالیت‌های آن امکان کشف، بازیابی، تضمین کیفیت و ایجاد ارزش افزوده و استفاده مجدد از داده پژوهشی را تسهیل می‌بخشند.

فرایند گزینش شدیداً بافت-محور است، لذا مطالعه مسائل، اعمال و رویه‌های گزینش داده پژوهشی به تحلیل بافتی چندوجهی نیاز دارد. گزینش داده پژوهشی موضوع چندوجهی است که نیازمند حضور و هماهنگی میان فناوری‌ها، ساختارهای سازمانی، دانش و مهارت‌های انسانی است (مایرنیک و همکاران، ۲۰۱۲). معمولاً هدف اغلب ارزیابی‌های ملزومات و نیازهای گزینش داده معطوف به اعضای هیئت علمی و محققان داخل یک رشته، موسسه دانشگاهی یا اجتماع علمی است که دانش و بینش عمیق و سودمندی درباره نگرش، رفتارها و رویه‌های گزینش داده محققان ارائه می‌کنند. هرچند انجام چنین تحقیقاتی برای اجرا و تثبیت نظام گزینش و مدیریت داده‌های پژوهشی بسیار ارزشمند است، اما چنین مطالعاتی اغلب با نرخ‌های پاسخ پایین انجام شده‌اند و لذا دستیابی به نتایج گسترده و تعمیم‌پذیری آنها را با مشکل روبرو ساخته‌اند. مطالعات موردی عمیق و ارزشمندی درباره داده‌های پژوهشی مانند پروفایل‌های داده پوردو^۵ انجام گرفته است که درک عمیقی از رویه‌های داده در سطح رشته ارائه می‌کند که بر اساس مصاحبه‌های مفصل با تک تک اعضای هیئت علمی فردی حاصل شده است. با این حال، با توجه به ماهیت توصیفی و انفرادی مطالعات موردی، برخی از یافته‌های این تحقیقات مختص رشته یا دانشگاه خاص هستند و از این رو، بهره‌برداری از نتایج آنها و قابلیت کاربرد آن در بیرون از حوزه مورد نظر متفاوت خواهد بود.

هرچند بسیاری از کتابخانه‌های دانشگاهی در حال ارائه خدماتی برای پشتیبانی از گزینش و مدیریت داده پژوهشی

¹Weber, Palmer and Chao

² United Kingdom E-Science Programme

³ Taylor

⁴ Mayernik et al

⁵ Purdue Data Profiles

خدمات و برای چه مخاطبانی ارائه خواهد شد؛ اگر هدف طراحی مدلی برای یک پروژه خاص باشد از مدل انفرادی استفاده می‌شود، این مدل درک بسیار مفصلی از کاری که انجام می‌شود ارائه می‌کند؛ اگر هدف ارائه خدمات برای سطح وسیع‌تری از دامنه پروژه واحد باشد، در این صورت از مدل سازمان‌محور استفاده می‌شود. مدل‌های چرخه حیات سازمان‌محور ترکیبی از چرخه‌های حیات داده‌محور را بازنمایی می‌کنند و بنابراین، یک بازنمایی وسیع‌تری از مراحل یا فعالیت‌هایی که در بین پروژه‌های تحقیق مشترک هستند ارائه می‌کنند. دقت و شرح جزئیات مدل چرخه حیات سازمان‌محور احتمالاً بر تعداد و طیفی از پروژه‌های تحقیق، حوزه‌ها یا رشته‌هایی که ارائه می‌شود بستگی دارد. مدل‌های چرخه حیات سازمان‌محور می‌توانند گسترش یابند و طیف گسترده‌ای از انواع داده تولیدشده با انواع روش‌های تحقیق را پوشش دهند. این مدل‌ها اجتماع‌محور هستند و از انتزاع بالایی برخوردارند و دقت و جزئیات مدل در آنها کاهش می‌یابد، ادغام تجربیات برتر با استانداردهای اجتماع (به این معنی که بدون توجه به دامنه، مدل چرخه حیاتی که توسعه می‌یابد باید بر اساس تجربیات برتر شناسایی شده در مدل‌های چرخه حیات مبتنی بر اجتماع شکل گیرد. مدل‌های چرخه حیات مبتنی بر اجتماع می‌توانند کمک کنند که استانداردهای موجود، رویه‌ها و خدماتی که مورد توجه قرار می‌گیرند و نیز انواع فعالیت‌هایی که معمولاً این خدمات را تشکیل می‌دهند شناسایی شوند) و بازنمایی فعالیت‌های دنیای واقعی (شناسایی دقیق و روشن نیازها و رویه‌های داده‌محققان و سازگاری مدل‌ها با این ملزومات که به ایجاد خدمات مدیریت داده موثر منجر خواهد شد) گزینش و مدیریت داده‌های پژوهشی را مورد توجه جدی قرار دهند (کارلسون، ۲۰۱۴).

ما در مراحل اولیه پاسخ نظام‌مند و واقعی برای چالش‌های گزینش و مدیریت داده‌های پژوهشی هستیم. بدون تردید گزینش و مدیریت داده‌های پژوهشی با چالش‌هایی در عرضه‌های فناوری، حقوقی، سازمانی، فرهنگی، قانونی و امنیتی روبرو خواهد بود. خدمات گزینش داده با چالش‌های مختلفی از قبیل نبود آگاهی و آموزش میان محققان در زمینه جوانب و ابعاد گزینش، نبود زیرساخت‌های فنی و فرهنگی در موسسات دانشگاهی، مسائل حقوقی داده از قبیل حق مولف و مالکیت داده، حفظ حریم خصوصی و محرمانگی داده، نبود فرهنگ اشتراک‌گذاری داده در میان محققان و اعضای هیئت علمی دانشگاه‌ها، کمبود بودجه و عدم پشتیبانی مالی از سوی سازمان‌ها، عدم احساس ضرورت به گزینش و مدیریت داده‌های پژوهشی، وقت‌گیر بودن و هزینه‌بر بودن کار گزینش

هستند، اما مسیرهایی که طی کرده‌اند، منحصر به فرد و بر اساس نیازها و فشارهای خاص محیطی خود بوده است (کوپر و همکاران^۱، ۲۰۱۳). گزارش اولیه مطالعه موردی شش دانشگاه (دانشگاه پوردو^۲، دانشگاه کالیفرنیا سان دیگو^۳، دانشگاه کرنل^۴، دانشگاه جان هاپکینز^۵، دانشگاه ایلینویز شیکاگو^۶، موسسه فناوری ماساچوست^۷)، تفاوت‌های موجود در هر موسسه را در زمینه نحوه پشتیبانی از علم الکترونیکی، از قبیل توسعه برنامه‌های پشتیبانی در سطح هر موسسه، نقش کتابخانه، انواع خدمات ارائه‌شده و راهبردهای توسعه نیروی کار ماهر برای پشتیبانی از نیازهای محققان علوم را برجسته ساخته است (سوهنر، استیوز و وارد، ۲۰۱۰). توسعه یک رویکرد تیمی با اتکاء به مجموعه مهارت‌های از قبل موجود نیز می‌تواند به عنوان یک راهبرد برای دانشگاه‌هایی که خدمات داده پژوهشی اجرا می‌کنند در نظر گرفته شود. از سوی دیگر، بازاریابی و ایجاد ارتباط با ذینفعان کلیدی نیز اصل مهمی در اجرای موفق فرایند گزینش داده پژوهشی در سطح سازمانی است. از این طریق حمایت مدیران ارشد، آمادگی و تعهد سازمانی و فرهنگی برای پذیرش گزینش و مدیریت داده‌های پژوهشی جلب می‌شود. تثبیت موقعیت کتابخانه برای گزینش داده پژوهشی تنها با انبار مجموعه داده‌ها تحقق نمی‌یابد، بلکه گزینش‌کنندگان داده باید سعی کنند که ماهیت داده، نحوه استفاده از آن در آینده، جذب، ذخیره، دسترسی به آن را درک کنند و در نهایت به بررسی و ضبط عوامل فرهنگی، اجتماعی و فنی متعددی که می‌توانند بر گزینش داده تاثیر بگذارند، اقدام کنند.

مدل‌های چرخه حیات منابع و مواد حیاتی برای کار گزینش فراهم می‌کنند. در مواردی که یک سازمان می‌خواهد از مدل چرخه حیات توسعه یافته یک سازمان دیگر استفاده کند باید مراقب باشد که مدل انتخاب‌شده به دقت رویه‌های تحقیق و داده گروه‌هایی که برای آنها خدمات گزینش داده ارائه خواهد شد منعکس کند. کارلسون میان انواع متعدد مدل‌های چرخه حیات داده تمایز قائل می‌شود، از قبیل انفرادی، اجتماع‌محور و سازمانی. طراحان خدمات در توسعه یا پیروی از یک یا چند مدل چرخه حیات برای گزینش و مدیریت داده‌های پژوهشی باید مسائلی از قبیل دامنه (به این معنی که چه سطحی از

¹ Kouper et al

² Purdue University

³ University of California-San Diego

⁴ Cornell University

⁵ Johns Hopkins University

⁶ University of Illinois-Chicago

⁷ Massachusetts Institute of Technology

References

- Abbott, D. (2008). What is Digital Curation? DCC Briefing Papers: Introduction to Curation. Edinburgh: Digital Curation Centre. Handle: 1842/3362. Available at: <http://www.dcc.ac.uk/sites/default/files/documents/resource/briefing-papers/what-is-digital-curation.pdf>.
- Abrams, S., Kunze, J., Loy, D. (2010). An Emergent Micro-Services Approach to Digital Curation Infrastructure. *The International Journal of Digital Curation*. 1(51); 5172-5186.
- Afkhami, R. (2010). Management of Production to consume cycle of Research Data: Challenges and Opportunities. National Conference on Research and Technology management. Tehran: Imam Sadiq University. (Persian)
- Akers, K.G., & Doty, J. (2013). Research data management practices and perspectives: Differences among the arts and humanities, social sciences, medical sciences, and basic sciences. *International Journal of Digital Curation*. 8(2); 5-26.
- Angevaere, I. (2009). Taking care of digital collections and data: "Curation" and organizational choices for research libraries. *LIBER Quarterly*. 19 (1); 1-12.
- Baker, K.S. and Yarmey, L. (2009). Data stewardship: Environmental data curation and a web-of-repositories. *International Journal of Digital Curation*. 4(2); 12-27.
- Ball, A. (2010). Review of the State of the Art of the Digital Curation of Research Data. Other. University of Bath. (Unpublished). <http://opus.bath.ac.uk/18713/1/erim1rep091103ab10.pdf>
- Ball, A. (2012). Review of Data Management Lifecycle Models. Other. University of Bath. (Unpublished). Available at: <http://opus.bath.ac.uk/28587/1/redm1rep120110ab10.pdf>
- Beagrie, N. (2004). The continuing access and digital preservation strategy for UK Joint Information Systems Committee (JISC). *D-Lib Magazine*. 10(7/8).
- Beagrie, N. (2006). Digital Curation for Science, Digital Libraries, and Individuals. *International Journal of Digital Curation*. 1(0).
- Beagrie, N. (2008). Digital curation for science, digital libraries, and individuals. *International Journal of Digital Curation*. 1(1); 3-16.
- Beagrie, N., Beagne, R., & Rowlands, I. (2009). Research Data Preservation and Access. The Views of Researchers. *Ariadne*. 60
- Blue Ribbon Task Force on Sustainable Digital Preservation and Access (2010). Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information. Final Report, February 2010. Retrieved May 24, 2010 from http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf

و مدیریت داده‌های پژوهشی، کیفیت داده، پایداری داده، امنیت داده و چالش‌های دیگر روبرو است باید در هنگام عملیاتی کردن گزینش و مدیریت داده‌های پژوهشی مورد توجه قرار گیرند.

در کل، حوزه گزینش و مدیریت داده‌های پژوهشی حوزه نسبتاً جدید و پویا با مسائل و موضوعات خاص خود می‌باشد. این حوزه دارای مسائل و موضوعات پژوهشی فعال، پویا، پیچیده و چندوجهی است که پژوهشگران مختلفی از رشته‌های علمی مذکور و حتی پژوهشگران از حوزه‌های علمی دیگر را به خود درگیر کرده است. هرچند در خارج کارهای زیادی در زمینه توسعه مبانی و مفاهیم، مدل‌ها و چهارچوب‌های اجرایی، شناسایی چالش‌ها، آموزش و تربیت نیروی انسانی، شناسایی رفتارها، نیازها و رویه‌ها و ملزومات گزینش و مدیریت داده‌های پژوهشی در سطح بین‌المللی و ملی و در قلمروهای دانشگاه‌ها، سازمان‌های علمی صورت گرفته است، اما در ایران چندان مورد توجه قرار نگرفته است؛ بنابراین، گزینش و مدیریت داده‌های پژوهشی را می‌توان از دو جنبه مورد ملاحظه قرار داد: جنبه نظری و جنبه عملی. از جنبه نظری شایسته است که متخصصان رشته‌هایی چون علم اطلاعات و دانش‌شناسی، سیاست‌گذاری علم و فناوری، آرشیو، متخصصان علم الکترونیکی و سیستم‌های مدیریت اطلاعات و داده با هدف مشارکت جهانی در ادبیات نظری این حوزه، پژوهش‌های نظری برای توسعه ابعاد مختلف گزینش و مدیریت داده‌های پژوهشی انجام دهند. در بعد عملی باید دو رویکرد را همزمان مد نظر قرار داد. رویکرد کلان و خرد. همان‌طور که در سطور آغازین مقاله اشاره شد، مدیریت داده‌های پژوهشی یک مسئله کلان در عرصه مدیریت پژوهش و فناوری در کشورهاست، لذا دولت باید به سیاست‌گذاری و نظارت و ارزیابی سیاست‌ها و برنامه‌های مدیریت کلان داده‌های پژوهشی در کشور و شناخت معضلات و پرکردن خلاءهای قانونی، مالی، فناوری، فرهنگی و سازمانی اقدام نماید. در بحث خرد دانشگاه‌ها و مراکز علمی هستند که باید ضمن راه‌اندازی دوره‌های آموزش رسمی و تربیت نیروی متخصص داده، خط‌مشی‌ها، راهبردها و بسترهای لازم برای پیاده‌سازی نظام گزینش و مدیریت داده‌های پژوهشی را در سطح دانشگاه، رشته یا اجتماع خاص فراهم کنند.

- Framework. Proceedings of the Association for Information Science and Technology. 47 (1); 1-2.
- Creamer, A., Morales, M.E., Crespo, J., Kafel, D., & Martin, E.R. (2012). An assessment of needed competencies to promote the data curation and management librarianship of health sciences and science and technology librarians in New England. *Journal of eScience Librarianship*. 1 (1); 18-26.
- Creamer, A. (2015). Current Issues and Approaches to Curating Student Research Data. *Bulletin of the Association for Information Science and Technology*. 41 (6); 22-25.
- Crowston, K., & Qin, J. (2011). A capability maturity model for scientific data management: Evidence from the literature. *Proceedings of the American Society for Information Science and Technology*. 48 (1); 1-9.
- Curdt, C. (2014). Design and Implementation of a Research Data Management System: The CRC/TR32 Project Database (TR32DB). PhD Thesis, Universität zu Köln.
- Data ONE (2012). DataONE Education Module: Analysis and Workflows. Retrieved May 29, 2015, from [http://www.dataone.org/sites/all/documents/L10_Analysis Workflows.pptx](http://www.dataone.org/sites/all/documents/L10_Analysis%20Workflows.pptx)
- Diekmann, F. (2012). Data Practices of Agricultural Scientists: Results from an Exploratory Study. *Journal of Agricultural & Food Information*, 13(1):14-34.
- Digital Curation Centre. (2010). Data dimensions: disciplinary differences in research data sharing, reuse and long term viability. Retrieved from <http://www.dcc.ac.uk/sites/default/files/documents/publications/SCARP-Synthesis.pdf>.
- Digital Curation Centre. (2013). Digital Curation. Retrieved from <http://www.dcc.ac.uk/>
- Fulton, B., Botticelli, P. and Bradley, J. (2011). DigIn: a hands-on approach to a digital curation curriculum for professional development. *Journal of Education for Library & Information Science*. 52(2):95-109.
- Gabridge, T. (2009). The last mile: liaison roles in curating science and engineering research data. *Research Library Issues*. (265); 15-21.
- Giarlo, M. (2013). Academic libraries as quality hubs. *Journal of Librarianship and Scholarly Communication*. 1 (3); 1-10.
- Gold, A. (2007). Cyberinfrastructure, data and libraries, part 2: libraries and the data challenge: roles and actions for libraries. *D-Lib Magazine*. 13 (9/10).
- Groenewegen, D., Treloar, A. (2013). Adding Value by Taking a National and Institutional Approach to Research Data: The ANDS Experience. *International Journal of Digital Curation*. 8 (2); 89-98.
- Grynoch, T. (2016). Implementing research data management services in a Canadian context. *Dalhousie Journal of Interdisciplinary Management*. 12; 1-24.
- Guttman, M, Schürer K, Donakowski and Hilary Borgman, C.L. (2009). The future is now: a call to action for the humanities. *Digital Humanities Quarterly*. 3(4).
- Borgman, C.L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*. 63.
- Brown, D.J. (2009). International Council for Scientific and Technical Information (ICSTI) Annual Conference – Managing Data for Science. *Information Services & Use*. 29(4); 103-21.
- Brown, E. (2010). I know what you researched last summer: How academic librarians are supporting researchers in the management of data curation. *The New Zealand Library & Information Management Journal*. 52(1); 55-69.
- Burton, A., Groenewegen, D., Love, C., Treloar, A., Wilkinson, R. (2012). Making Research Data Available in Australia. *IEEE Intelligent Systems*. 27 (3); 40-43.
- Carlson, J. (2012). Demystifying the data interview: Developing a foundation for reference librarians to talk with researchers about their data. *Reference Services Review*. 40(1); 7-23.
- Carlson, J. (2014). The Use of Lifecycle Models. In *Research Data Management: Practical Strategies for Information Professionals*, by Joyce Ray, 63-86. West Lafayette: Purdue University Press.
- Choudhury, G.S. (2010). Data curation – an ecological perspective. *College & Research Libraries News*. 71(4).
- Choudhury, S, Hanisch, . (2009). The Data Conservancy: Building a Sustainable System for Interdisciplinary Scientific Data Curation and Preservation. Presentation at the PV 2009 conference in Madrid, Spain.
- Cochran, K T. (2014). Access in Curation: Outreach at the Anasazi Heritage Center. Thesis. Northern Arizona University. Master of Arts In Applied Anthropology.
- CODATA. (2014). CODATA: Who are we? <http://www.codata.org/about/who>.
- Corrall, S. (2010). Educating the academic librarian as a blended professional: a review and case study. *Library Management*. 31 (8/9); 567-93.
- Corrall, S., Kennan, M.A., & Afzal, W. (2013). Bibliometrics and research data management: Emerging trends in library research support services. *Library Trends*. 61 (3); 636-674.
- Cox, A.M. and Pinfield, S. (2014). Research data management and libraries: Current activities and future priorities. *Journal of Librarianship and Information Science*. 46(4); 299-316.
- Cragin, M.H., Palmer, C.L., Carlson, J.R. and Witt, M. (2010). Data sharing, small science and institutional repositories. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*. 368(1926); 4023-38.
- Cragin, Melissa H., Carole L. Palmer, Tiffany C. Chao. (2010). Relating Data Practices, Types, and Curation Functions: An Empirically Derived

- on Digital Libraries(pp. 375–376).New York, NY:ACM.
- Loughton, P and Plessi, T du. (2013).Data Curation in the world Data System: Proposed Framework. *Data Science Journal*.12;56-70.
- Lee DJ, Stvilia,B. (2017).Practices of research data curation in institutional repositories: A qualitative view from repository staff. *PLoS ONE*. 12(3);e0173987.
- Lee,C.A.,& Tibbo, H. (2007).Digital curation and trusted repositories: Steps toward success. *Journal of Digital Information*.8 (2).
- Lee,dong joon, stvilia, besiki. (2014).Data curation practices in institutional repositories:An exploratory study: Proceedings of the Association for Information Science and Technology.51(1);1–4.
- Little,G. (2011).We are all digital humanists now.*The Journal of Academic Librarianship*.37(4);352-4.
- Lord,P.,Macdonald,A.,Lyon, L.,&Giaretta, D. (2004).From data deluge to data curation. (S. J. Cox, Ed.)*E-Science All Hands Meeting 2004* (pp. 371-375). Nottingham, England. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.111.7425&rep=rep1&type=pdf>.
- Majidi A,Naghshineh N, Esmaeili Gavi M R,Hashemi, M R. (2017).Research Data Curation: From Storage to Discovery. First International Interactive Information Retrieval Conference.University of Tehran. Kish International Campus. International Kish International Conference Center.22-23 February 2017. (Persian)
- Majidi A,Naqshineh N, Esmaeili Gavi M R. (2017).data curation services in academic centers.national conference of big data management in information and knowledge. Azad Islamic university.babol branch. 25 February 2017. (Persian)
- Ma,N., Li,C.Jiang,A.,& Xing,C. (2008).Design and implementation of open source based digital preservation experimental platform (THDP). Conference proceedings of the 9th International Conference for Young Computing Societies. Conducted by the China Computer Federation held in Hunan.
- Macdonald,S.and Martinez-Uribe,L. (2008).A new role for academic librarians:data curation. *El Profesional de la Informació n*.17 (3); 273-80.
- Macdonald, S. and Martinez-Uribe, L. (2010).Collaboration to data curation: harnessing institutional expertise.*New Review of Academic Librarianship*. (16);4-16.
- Makani,Joyline. (2015).Knowledge management, research data management, and university scholarship. *VINE*.45 (3);344 –359.
- Martinez-Uribe,L.,& Macdonald,S. (2009).User engagement in research data curation. In Proceedings of the 13th European conference on Research and advanced technology for digital libraries (pp. 309–314). Berlin, Heidelberg: Springer-Verlag.
- Beedham. (2004).The selection,Appraisa, and Retention of Digital social science data.*Data Science Journal*. (3);209-221.
- Harris-Pierce, R.L. and Liu, Y.Q. (2012).Is data curation education at library and information science schools in North America adequate?.*New Library World*.113 (11/12); 598-613.
- Heidorn,P.B. (2011).The emerging role of libraries in data curation and E-Science. *Journal of Library Administration*.51 (7); 662–72.
- Heidorn, P.Bryan, Helen R. Tobbo,G.Sayeed Choudhury, Christopher Greer, Richard Marciano. (2007).Identifying Best Practices and Skills for Workforce Development in Data Curation.Proceedings of the American Society for Information Science and Technology.44(1);1-3.
- Henty, M.,Weaver,B.,Bradbury,S., & Porter, S. (2008).Investigating data management practices in Australian universities. Retrieved July 7, 2016, from <http://hdl.handle.net/1885/47627>.
- Higgins,S. (2008).The DCC curation lifecycle model. *International Journal of DigitalCuration*.1 (3); 134-140.
- Hox,J.J.,Boeije,H.R. (2005).Data collection, primary versus secondary. In: KEMPF-LEONARD, K. (Ed.), *Encyclopedia of social measurement*. Elsevier, Amsterdam, The Netherlands.593-599.
- Jacobs,C A. & Worley, S J. (2009).Data Curation in Climate and Weather: Transforming Our Ability to Improve Predictions through Global Knowledge Sharing*The International Journal of Digital Curation*.2(4);68-79.
- JISC. (2009).Managing and sharing data:A best practices guide for researchers.UK Data Archive. Retrieved from <http://www.data-archive.ac.uk/media/2894/managingsharing.pdf>.
- JISC. (2003).JISC Circular 6/03 (Revised).An invitation for expressions of interest to establish a new Digital Curation Centre for research into and support of the curation and preservation of digital data and publications.
- Kahn,M,Higgs,R,Davidson,J and Jones S. (2014).Research Data Management in South Africa:How We Shape Up.*Australian Academic & Research Libraries*.45(4);296–308.
- Karasti, H, Baker, K S. &Halkola, E. (2006).Enriching the Notion of Data Curation in E-Science: Data Managing and Information Infrastructuring in the Long Term Ecological Research (LTER) Network. *Computer Supported Cooperative Work*.15;321–358
- Kennan,Mary Anne, Markauskaite, Lina. (2015).Research Data Management Practices:A Snapshot in Time.*The International Journal of Digital Curation*.10(2);69-95.
- Kouper,I.,Akers, K.G.,Nicholls,N.H., & Sferdean,F.C. (2013).A roadmap for data services. In Proceedings of the 13th ACM/IEEE-CS Joint Conference

- Osswald, A. and Strathmann, S. (2012). The role of libraries in curation and preservation of research data in Germany: findings of a survey In IFLA General Conference and Assembly, available at: <http://conference.ifla.org/past/2012/116-osswald-en.pdf>
- Palmer, Carole L., Nicholas M. Weber, Trevor Muñoz, and Allen H. Renear. (2013). Foundations of Data Curation: The Pedagogy and Practice of 'Purposeful Work' with Research. *Data. Archive Journal*, (3).
- Parham, Wells, S., Bodnar, J., & Fuchs, S. (2012). Supporting tomorrow's research: Assessing faculty data curation needs at Georgia Tech. *College and Research Libraries News*. (73); 10-13..
- Partlo, K., Symons, D. and Carlson, J.D. (2015). Revolutionary or evolutionary? Making research data management manageable. In Eden BL, Ed. *Creating Research Infrastructures in the 21st Century Academic Library: Conceiving, Funding, and Building New Facilities and Staff*, Rowman and Littlefield, Lanham, MD, pp: 175-201.
- Patel, M., & Ball, A. (2008). Challenges and issues relating to the use of representation information for the digital curation of crystallography and engineering data. *International Journal of Digital Curation*. 3(1); 76-88.
- Peters, C., & Dryden, A. R. (2011). Assessing the academic library's role in campus-wide research data management: A first step at the University of Houston. *Science & Technology Libraries*. 30 (4); 387-403.
- Poole, Alex H. (2016). The conceptual landscape of digital curation. *Journal of Documentation*. 72(5); 961-986.
- Pourabbasi J, Aghaie A, Heidary M H. (2015). A Service-oriented Architecture for Generating and Sharing Science and Knowledge in the E-science Environment. *Iranian Journal of Processing & Management*. 31 (1); 265-292. (Persian)
- Ray, J. (2009). Sharks, digital curation, and the education of information professionals. *Museum Management & Curatorship*. 24 (4); 357-68.
- RCUK (2011). RCUK common principles on data policy. Swindon: RCUK (Available at: <http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx>).
- Research Data Alliance (2014). About RDA. <https://www.rd-alliance.org/about>.
- Reference Model for an Open Archival Information System (2002). Recommendation for Space Data System Standards: Consultative Committee for Space Data Systems Secretariat, Program Integration Division (Code M-3), National Aeronautics and Space Administration. Retrieved from <http://public.ccsds.org/publications/archive/650x0b1.pdf> on 4 October 2015.
- Reilly, S. (2012). The role of libraries in supporting data exchange. In 78th IFLA General Conference and Assembly. 11-17 August Helsinki Finland.
- Martinez-Urbe, L. (2008). Findings of the scoping study and research data management workshop. <http://tinyurl.com/55fxgw>.
- Mayernik, Matthew S., G. Sayeed Choudhury, Tim DiLauro, Elliot Metsger, Barbara Pralle, Mike Rippin, Ruth Duerr. (2012). The Data Conservancy Instance: Infrastructure and Organizational Services for Research Data Curation. *D-Lib Magazine*. 18(9/10).
- McGovern, N. (2009). Digital Preservation Management: Implementing Short-term Strategies for Long-term Problems. [workshop] ICPSR, May 2009, Ann Arbor, MI.
- McLure, Merinda., Allison V. Level., Catherine L. Cranston., Beth Oehlerts., and Mike Culbertson. (2014). Data Curation: A Study of Researcher Practices and Needs. *Portal: Libraries and the Academy*. 14 (2); 139-164.
- Monastersky, R. (2013). Publishing frontiers: The library reboot. *Nature*. 495: 430-432.
- Mullins, J.L. (2010). The Challenges of E-Science Data-set Management and Scholarly Communication for Domain Sciences and Engineering: a Role for Academic Libraries and Librarians, In Marcum, D.B. and George, G. (Ed.), *The Data Deluge: Can Libraries Cope with E-Science?* ABC-CLIO, Santa Barbara, CA. pp: 33-42.
- National Science Board. (2005). Long-lived digital data collections Enabling research and education in the 21st century. Arlington, VA: National Science Board Committee on Programs and Plans, NSB-05-40. Retrieved from http://www.nsf.gov/pubs/2005/nsb0540/nsb0540_1.pdf.
- National Science Foundation. (2007). Cyberinfrastructure vision for 21st century discovery. Retrieved from <http://www.nsf.gov/pubs/2007/nsf0728/nsf0728.pdf>.
- NHMRC. (2007). Australian Code for the Responsible Conduct of Research, Australian Government, National Health and Medical Research Council, Canberra, ACT, Australia. <http://www.nhmrc.gov.au/guidelines/publications/r39>.
- Niu, Jinfang (2014) Appraisal and Selection for Digital Curation. *International Journal of Digital Curation*. 9(2); 65-82.
- Nielsen, Hans Jørn, Hjørland, Birge. (2014). Curating research data: the potential roles of libraries and information professionals. *Journal of Documentation*, 70(2): 221 - 240.
- OECD. (2007). OECD principles and guidelines for access to research data from public funding. Retrieved from <http://www.oecd.org/dataoecd/9/61/38500813.pdf>
- Olendorf, R. & Koch, S. (2013). Beyond the low hanging fruit: Data services and archiving at the University of New Mexico. *Journal of Digital Information*, 13.

- Retrieved from http://old.arl.org/bm~doc/escience_report2010.pdf
- f
- Starr, J., Willett, P., Federer, L., Horning, C. & Bergstrom, M.L. (2012). A collaborative framework for data management services: The experience of the University of California. *Journal of eScience Librarianship*.1(2); 109-114
- Steinhart, G., Chen, E., Arguillas, F., Dietrich, D., & Kramer, S. (2012). Prepared to plan? A snapshot of researcher readiness to address data management planning requirements. *Journal of eScience Librarianship*.1(2);63-78
- Steinhart, G., Saylor, J., Albert, P., Alpi, K., Baxter, P., Brown, E., Chiang, K., Rikert, J.C., Hirtle, P., Jenkins, K., Lowe, B., McCue, J., Ruddy, D., Silterra, R., Solla, L., Stewart-Marshall, Z. and Elaine, L. (2008). Digital research data duration: overview of issues, current activities and opportunities for the Cornell University Library. available at: <http://hdl.handle.net/1813/10903>.
- Stvilia, B., Charles, C., Shuheng, W., Adam, W., Dong, J., Lee, K., Burnett, Gary, B., Michelle M.K, Paul F.M . (2013). Studying the data practices of a scientific community, JCDL'13 Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries.
- Swan, A. and Brown, S. (2008). The skills, role and career structure of data scientists and curators: an assessment of current practice and future needs. available at: http://eprints.soton.ac.uk/266675/1/Data_skills_report_final_draft.doc
- Taylor, J. (2001). The UK E-Science Programme. Presented at E-Science London Meeting 27th July 2001.
- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A.U., Wu, L., Read, E. Frame, M. (2011). Data sharing by scientists: Practices and perceptions. *PLoS One*.6 (6);e21101.
- Tenopir, C., Sandusky, R.J., Allard, S., & Birch, B. (2013). Academic librarians and research data services: Preparation and attitudes. *IFLA Journal*.39 (1); 70-78.
- Tenopir, C., Sandusky, R.J., Allard, S., & Birch, B. (2014). Research data management services in academic research libraries and perceptions of librarians. *Library and Information Science Research*.36(2);84-90.
- The University of Illinois at Urbana-Champaign (UIUC), Graduate School of Library and Information Science. (2011). Master of Science: Specialization in Data Curation. Available at: http://www.lis.illinois.edu/academics/programs/ms/data_curation.
- Theodoridou, M., Tzitzikas, Y., Doerr, M., Marketakis, Y. and Melessanakis, V. (2010). Modeling and querying provenance by extending CIDOC CRM. *Distributed and Parallel Databases*.27(20);169-210.
- Available at: <https://www.ifla.org/past-wlic/2012/116-reilly-en.pdf>.
- Rezaei Sharif Abadi, M. (2015). Research Data Management: A Case Study of Linguistic Data. The 3rd National Conference on Linguistics and Persian Language Education; The Perspective of Language Studies in the 21st Century. Mashhad: Ferdowsi University of Mashhad. available at: transparency4iran.ir/post/480. (Persian)
- Reznik-Zellen, R.C., Adamick, J. & McGinty, S. (2012). Tiers of research data support services. *Journal of eScience Librarianship*.1(1); 27-35.
- Sallans, A., Lake, S. (2014). Data Management Assessment and Planning Tools. In: RAY, J. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*. Purdue University Press, West Lafayette, IN, USA, 87-107.
- Sarrafzadeh, M. (2015). Research big Data management: A New Role for Academic Libraries. *Review of the Information and Communication Book*.2(6); 265-274. (Persian)
- Scaramozzino, J.M., Ramirez, M L., & McGaughey, K.J. (2012). A study of faculty data curation behaviors and attitudes at a teaching-centered university. *College and Research Libraries*.73(4); 349-365.
- Shakeri, S. (2014). A Model for Data Curation Research in Small Sciences. *Proceedings of the American Society for Information Science and Technology*.51(1);1-5.
- Shen, Y., Varvel, Jr V E. (2013). Developing Data Management Services at the Johns Hopkins University. *The Journal of Academic Librarianship*.39(6); 552-557.
- Shorish, Y. (2012). Data curation is for everyone! The case for master's and baccalaureate institutional engagement with data curation. *Journal of Web Librarianship*.6 (4); 263-273.
- Shreeves, S L. and Cragin, M H. (2008). Introduction: Institutional Repositories: Current State and Future. *Library Trends*.57(2); 89-97.
- Simbulan, M S R. (2013). Transitioning from Data Storage to Data Curation: The Challenges Facing an Archaeological Institution. *Issues in Informing Science and Information Technology (IISIT)*. (10); 489-499.
- Smith II, P L. (2012). Exploring Data Curation and Management Programs, Projects, and Services through Metatriangulation. *JCDL'12*.8(2);1-3. Available at: <http://www.ieee-tcdl.org/Bulletin/v8n2/papers/smith.pdf>
- Smith, M. (2014). Data Governance: Where Technology and Policy Collide. In *Book: Research Data Management: Practical Strategies for Information Professionals*, 45-63 (Chapter). Purdue University Press.
- Soehner, C., Steeves, C., & Ward, J. (2010). E-science and data support services: A study of ARL Member Institutions. *Association of Research Libraries*.

- Society for Information Science and Technology.38(5);14-19.
- Wissik, T, ur o, M. (2015). Research Data Workflows:From Research Data Lifecycle Models to Institutional Solutions.In CLARIN 2015 Selected Papers, Linköping Electronic Conference Proceedings, Annual Conference 2015, October 14–16, 2015, Wroclaw, Poland <http://www.ep.liu.se/ecp/123/008/ecp15123008.pdf>
- Witt, M., Carlson, J., Brandt, D. and Cragin, M. (2009).Constructing data curation profiles.International Journal of Digital Curation.4 (3); 93-103.
- Yakel, E. (2007). Digital curation.OCLC Systems & Services: International digital library perspectives.23 (4/2); 335-340
- York, J., Gutmann, M. and Berman, F. (2016). What Do We Know About The Stewardship Gap? https://deepblue.lib.umich.edu/bitstream/handle/2027.42/122726/StewardshipGap_Final.pdf?sequence=1&isAllowed=y.
- Zimmerman, A S. (2008). New knowledge from old data: The role of standards in the sharing and reuse of ecological data. Science, Technology & Human Values.33 (5); 631–652.
- Zorich, D M. (1995). Data management:Managing electronic information: Data curation in museums. Museum Management and Curatorship.14 (4); 430-432.
- Treloar A. (2009). Design and Implementation of the Australian National Data Service. International Journal of Digital Curation.4 (1); 122-137.
- Treloar A., Wilikinson R. (2008). Access to Data for eResearch: Designing the Australian National Data Service Discovery Services. International Journal of Digital Curation.2 (2); 151-158.
- Van Tuyl, S., & Michalek, G. (2015). Assessing Research Data Management Practices of Faculty at Carnegie Mellon University. Journal of Librarianship and Scholarly Communication.3 (3), eP1258.
- Vaziri I, Naghshineh N, Norouzi Chakly A. (forthcoming 2017). Sharing Research Data: Global and National Approaches. Iranian Journal of Processing & Management.
- Weber, N M., Palmer, C L., and Chao, T C. (2012). Current Trends and Future Directions in Data Curation Research and Education.Journal of Web Librarianship.6(4);305–320.
- WELLCOME TRUST, (.2010). Policy on data management and sharing, <http://www.wellcome.ac.uk/About-us/Policy/Policy-andposition-statements/WTX035043>.
- Whitmire, A.L, Boock, M and Sutton, S C. (2015). Variability in academic research data management practices.Program.49(4);382-407.
- Wickett, K M., Hu, X and Thomer, A. (2012). The RDAP12 Summit:Challenges and Opportunities for Data Management.Bulletin of the American

Archive



Study of the Foundation, Models and Issues of Research Data Curation and Management in Scientific and Academic Environments

Akbar Majidi: PhD student in Information and Knowledge Science, University of Tehran, Tehran, Iran (Corresponding author). akbarmajidi@ut.ac.ir

Nader Naghshineh: PhD in Information and Knowledge Science, Assistant Professor, University of Tehran, Tehran, Iran. nnaghsh@ut.ac.ir

Mohammad Reza Esmaili Givi: PhD in Systems Management, Assistant Professor, University of Tehran, Tehran, Iran. s.givi@ut.ac.ir

Mahmood reza hashemi: PhD in Computer Science, Assistant Professor, University of Tehran, Tehran, Iran. rhashemi@ut.ac.ir

Abstract

Background and Aim: The purpose of this paper is to study, identifying and discuss the foundation and concepts, models and frameworks, dimensions and challenges of research data curation and management in scientific and academic environments.

Methods: This article is a review article and library method was used to collect scientific and research texts in this field. In this research, external and internal scientific databases, as well as web resources, were searched with the keywords "data curation", "digital curation", "research data management", "research data curation" and their equivalent in Persian. After removing duplicate sources and unrelated sources, a total of 132 sources were selected and their content evaluated and analyzed.

Results: The analysis of the literature revealed that the curation and management of research data is a emerging area with complex issues and different dimensions which included of a wide range of educational, organizational, cultural, technological, legal, and security issues. Since the 2000s, This area has been at the forefront of governments, funding organizations, universities and has been formed around it a strong research community of researchers, especially in the field of information and knowledge science, archives and information systems. Today, the curation and management of research data is considered as one of the main components of the large-scale science and technology policy and needs government support and legislative and policy-making institutions. Different models and frameworks have been drawn up at various levels of the national, institution or community for understand the dimensions of the curation and management of research data and its implementation. The study of literature also has shown that libraries and librarians have the ability and competence to take on the roles and responsibilities of curation and management of research data at universities and scientific institutes. However, the research data curation and management in implementation and practice faces with challenges such as privacy, data ownership, intellectual property rights, lack of data sharing by researchers, lack of proper data management infrastructure, lack of awareness and cognition and knowledge of the process of research data curation and management, and so on.

Conclusion: Despite extensive research abroad, this area has not received much attention in Iran. Therefore, this article addressed relatively comprehensive the subject and dimensions of the research data curation and management, and it can be useful for researchers, policy makers of science and technology, librarians for research and implement research data curation and management services.

Keywords: Data curation, Research Data Management, Research Data, Life Cycle Models, Data Curation Life Cycle Models, Digital Curation.