

توانمندی‌ها و قابلیت‌های موردانتظار از کتابدار داده برای ایفای نقش در مدیریت داده‌های

پژوهشی بر اساس چرخه‌ی حیات داده زیست‌پزشکی هاروارد: یک مرور دامنه

زهرة احتشامی^۱، اعظم شاه‌داغی^۲، محمدجواد منصورزاده^۳

چکیده

زمینه و هدف: کتابدار داده‌ی کارآمد و واجد توانمندی‌ها و قابلیت‌های لازم، از مهمترین مولفه‌ها در مدیریت داده‌های پژوهشی است. هدف از پژوهش حاضر استخراج توانمندی‌ها و قابلیت‌های مورد انتظار برای کتابدار داده به‌منظور ایفای نقش در مدیریت داده‌های پژوهشی است.

روش بررسی: این مطالعه از نوع مرور دامنه بوده و برای ارایه ساختارمند یافته‌ها از مدل چرخه‌ی حیات داده‌های زیست‌پزشکی هاروارد استفاده شده است. به‌منظور دستیابی به متون مرتبط، استراتژی جستجو با استفاده از کلیدواژه‌های مرتبط در پایگاه‌های اطلاعاتی PubMed، Scopus، Web of Science و همچنین سایر پایگاه‌های اطلاعاتی معتبر داخلی و خارجی در بازه زمانی ۵ سال اخیر انجام شد. مقالات پژوهشی اصیل فارسی و انگلیسی که به‌نحوی به توانمندی‌ها و قابلیت‌های مورد انتظار برای کتابدار داده در مدیریت داده‌های پژوهشی پرداخته بودند، جامعه پژوهش حاضر را تشکیل دادند.

یافته‌ها: از ۵۰۶۴ مدرک یافت شده، ۱۹۶ مدرک جهت بررسی متن کامل به مرحله بعد وارد شد. پس از بررسی متن کامل این مدارک، در نهایت ۱۷ مطالعه وارد پژوهش شد. در مجموع ۹۲ توانمندی و قابلیت در ۲۳ فرایند از ۷ مرحله چرخه‌ی حیات داده‌های زیست‌پزشکی هاروارد، ۱۶ مورد مربوط به مرحله اول، ۱۶ مورد مرحله دوم، ۷ مورد مرحله سوم، ۱۵ مورد مرحله چهارم و پنجم، ۱۲ مورد مرحله ششم، ۸ مورد مرحله هفتم، و ۱۸ مورد مربوط به توانمندی‌ها و قابلیت‌های عمومی بود. مطابق یافته‌ها، بیشترین مطالعات به توانمندی‌ها و قابلیت‌های مورد نیاز برای مرحله دوم یعنی «گردآوری و ایجاد» و کمترین مطالعات به مرحله هفتم یعنی «انتشار و استفاده‌ی مجدد» پرداخته‌اند. در خصوص انجام فرایند «مدیریت تصاویر» در مرحله سوم و «پیش‌چاپ و انتشار» در مرحله هفتم در هیچ‌یک از مطالعات توانمندی و قابلیت‌های مورد اشاره قرار نگرفته بود.

نتیجه‌گیری: نتایج این پژوهش نشان داد که در میان مراحل مختلف چرخه‌ی حیات داده، مرحله‌ی «گردآوری و ایجاد» بیشترین میزان توجه را داشته است. همچنین کتابدار داده باید علاوه بر توانمندی‌ها و قابلیت‌های تخصصی و حرفه‌ای، ویژگی‌های عمومی لازم را هم داشته باشد. توصیه می‌شود که یافته‌های این پژوهش به‌منظور تربیت کتابدار داده برای طراحی برنامه‌های آموزشی کوتاه‌مدت و بلندمدت مدنظر قرار گیرد.

واژه‌های کلیدی: مدیریت داده‌های پژوهشی، کتابدار داده، چرخه حیات داده زیست‌پزشکی هاروارد

دریافت مقاله: ۱۴۰۳/۵/۲۷

پذیرش مقاله: ۱۴۰۳/۹/۱۲

* نویسنده مسئول:

اعظم شاه‌داغی؛

دانشکده پیراپزشکی دانشگاه علوم پزشکی شهید بهشتی

Email:

shahbodaghi@sbmu.ac.ir

۱ دانشجوی کارشناسی ارشد کتابداری و اطلاع‌رسانی پزشکی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

۲ استادیار گروه کتابداری و اطلاع‌رسانی پزشکی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

۳ دکتری علوم کتابداری و اطلاع‌رسانی پزشکی، مرکز تحقیقات پیشگیری از بیماری‌های متابولیک، پژوهشکده علوم غدد درون‌ریز و متابولیسم، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

مقدمه

داده‌های پژوهشی منابع ارزشمندی هستند که در صورت ذخیره‌سازی و به اشتراک‌گذاری مناسب می‌توانند در فعالیت‌های علمی و پژوهشی آینده قابل استفاده باشند (۱). ارزشمندی منابع داده‌های مستلزم مدیریت صحیح این منابع است (۲). از آنجایی که استفاده‌ی مجدد از داده‌های پژوهشی می‌تواند هزینه‌های تحقیق را به آسانی کاهش دهد و از اتلاف وقت، تلاش و انرژی در جمع‌آوری مجدد داده‌ها جلوگیری کند (۳)، مدیریت این منابع ارزشمند یک ضرورت انکارناپذیر است. مدیریت این منابع ارزشمند، باید از مرحله‌ی پیش از انجام پژوهش آغاز شود و تا بعد از اتمام آن ادامه یابد (۲) و تمام فعالیت‌ها و اقدامات مربوط به حفظ و نگهداری، دسترسی و استفاده‌ی طولانی مدت از داده‌ها را در برگیرد (۴). شکی نیست که مدیریت داده‌ها یک چالش بزرگ برای سازمان‌های پژوهشی است. حجم عظیمی از داده‌های پژوهشی در حال حاضر به اشکال مختلف و با سرعت در دانشگاه‌ها و موسسات پژوهشی تولید می‌شود و از نظر حجم، تنوع و سرعت تولید، چالش‌هایی را ایجاد می‌کند. بنابراین نیاز به توسعه‌ی سیاست‌ها، زیرساخت‌ها و خدمات با هدف کمک به محققان در ایجاد، جمع‌آوری، دستکاری، تجزیه و تحلیل، انتقال، ذخیره و حفظ مجموعه داده‌ها احساس می‌شود (۵).

جهت توصیف و درک بهتر مدیریت داده‌های پژوهشی می‌توان از چرخه‌ی حیات داده‌های پژوهشی استفاده کرد. این چرخه، تمامی مراحل از ایجاد داده‌ها تا حفظ و به اشتراک‌گذاری آن‌ها را در بر می‌گیرد (۶). مدل‌های متنوعی برای چرخه‌ی حیات داده‌های پژوهشی ارائه شده است؛ اما به نظر می‌رسد که چرخه‌ی حیات داده‌های زیست‌پزشکی هاروارد (۷) به دلیل جامعیت، جدید بودن و استفاده در حوزه‌ی زیست‌پزشکی، برای تبیین بهتر مراحل مدیریت داده‌های پژوهشی پزشکی مناسب‌تر باشد (۸). این چرخه شامل مراحل برنامه‌ریزی و طراحی، جمع‌آوری و ایجاد، تجزیه و تحلیل، ذخیره‌سازی و مدیریت، ارزیابی و آرشیو، اشتراک‌گذاری و انتشار و دسترسی و استفاده‌ی مجدد است (۷).

با نگاهی به چرخه‌ی حیات داده‌های زیست‌پزشکی هاروارد می‌توان دریافت که کتابخانه‌ها می‌توانند با کسب توانمندی‌ها و قابلیت‌هایی در تمامی این مراحل نقش مؤثری ایفا کنند. کتابخانه‌ها و نهادهای تحقیقاتی مراکز اصلی حمایت از پژوهش و منابع پژوهشی هستند و تجارب با ارزشی در حوزه‌ی سازماندهی، نگهداری، کنترل و حفظ اطلاعات دارند (۲). همچنین به خاطر جدید بودن موضوع مدیریت داده‌های پژوهشی، بسیاری از کتابخانه‌های دانشگاهی جهان در این

زمینه خدمات ترویجی، آموزشی و مشاوره‌ای ارایه می‌کنند (۹). این مراکز به تدریج در حال توسعه‌ی خدمات خود به منظور برآورده کردن الزامات مدیریت داده‌های پژوهشی هستند (۱۰)؛ در حال حاضر برخی از کتابخانه‌ها در توسعه سیاست‌ها و خدمات مدیریت داده‌های پژوهشی سازمان‌ها مشارکت دارند و این امر را بخش مهمی از نقش آینده‌ی خود می‌دانند (۱۱).

کتابدارانی که به صورت تخصصی در مدیریت داده‌ها آموزش می‌بینند و در کتابخانه‌ها به اشکال مختلف وظایف مرتبط با مدیریت داده‌های پژوهشی را انجام می‌دهند، کتابدار داده نامیده می‌شوند (۱۵-۱۲). کتابدار داده، متخصص کتابداری است که به جای «مجموعه منابع» با «مجموعه‌ی داده» سروکار دارد (۱۶). با حرکت به سمت فضای مجازی و تغییر شکل داده‌ها از فرمت کاغذی به دیجیتال، کتابداران داده به پژوهشگران کمک می‌کنند تا فرمت‌های جدید داده را کشف و تحلیل کنند و آن‌ها را در یافتن داده‌های موردنیازشان از منابع داده‌های موجود یاری می‌رسانند (۱۶). آن‌ها می‌توانند از لحظات آغازین یک پژوهش علمی در کنار محققان قرار بگیرند و به عنوان یکی از اعضای تیم تحقیق در فرایندهای مختلف مدیریت داده‌های پژوهشی به محققان یاری برسانند (۱۷). با توجه به ضرورت روزافزون مدیریت داده‌های پژوهشی، شاید به جرات بتوان گفت که کتابدار داده‌ی کارآمد و واجد شرایط، مهمترین مؤلفه در مدیریت داده‌های پژوهشی است. برای تربیت چنین کتابدارانی، به‌ویژه به منظور مدیریت داده‌های پژوهشی در حوزه‌های علوم پزشکی، لازم است تا سیاست‌گذاران و برنامه‌ریزان نظام‌های آموزش دانشگاهی و دوره‌های بازآموزی از توانمندی‌ها و قابلیت‌های مورد انتظار کتابداران داده آگاه شوند. این داده‌ها همچنین می‌تواند در راستای جذب نیروی انسانی مورد نیاز برای مدیریت داده‌های پژوهشی، به تصمیم‌سازان و دست‌اندرکاران مراکز تحقیقاتی علوم پزشکی یاری برساند. این پژوهش بر آن است تا با استفاده از چرخه‌ی حیات داده‌های زیست‌پزشکی هاروارد، در قالب یک مرور دامنه، به بررسی مطالعاتی بپردازد که به نحوی به تشریح این توانمندی‌ها و قابلیت‌ها مبادرت کرده‌اند.

روش بررسی

این مطالعه که به روش مرور دامنه انجام شده است، به منظور استخراج توانمندی‌ها و قابلیت‌های مورد انتظار کتابدار داده به منظور ایفای نقش در مدیریت داده‌های پژوهشی، مطالعات اصیل معتبر بررسی شده است. جستجو

در پایگاه‌های اطلاعاتی در آبان ۱۴۰۲ و در محدوده زمانی ۵ سال اخیر انجام شد. استراتژی جستجوی مربوط به هر پایگاه در جدول ۱ آمده است:

جدول ۱: استراتژی جستجو

تعداد	پایگاه داده	استراتژی جستجو
۴۷۹	PubMed	((("data life cycle"[tiab] OR "data curation"[mesh] OR "data curation"[tiab] OR "data management"[mesh] OR "data management"[tiab] OR "data repository"[tiab] OR "data preservation"[tiab] OR "data anonymization"[tiab] OR "data sharing"[tiab] OR "data annotation"[tiab] OR "data archiving"[tiab] OR "data documentation"[tiab] OR "data security"[tiab] OR "data retention"[tiab] OR "data destruction"[tiab] OR "data polic*"[tiab]) AND (librar*[tiab] OR informationist[tiab])) OR ("data librar*"[tiab] OR "research librar*"[tiab] OR "research informationist*"[tiab]))
۳۵۳۳	Scopus	(TITLE-ABS-KEY ("data life cycle" OR "data curation" OR "data management" OR "data repository" OR "data preservation" OR "data anonymization" OR "data sharing" OR "data annotation" OR "data archiving" OR "data documentation" OR "data security" OR "data retention" OR "data destruction" OR "data polic*") AND TITLE-ABS-KEY (librar* OR informationist)) OR (TITLE-ABS-KEY ("data librar*" OR "research librar*" OR "research informationist*"))
۹۷۵	WOS	(TS= ("data life cycle" OR "data curation" OR "data management" OR "data repository" OR "data preservation" OR "data anonymization" OR "data sharing" OR "data annotation" OR "data archiving" OR "data documentation" OR "data security" OR "data retention" OR "data destruction" OR "data polic*") AND TS=(librar* OR informationist)) OR (TS=("data librar*" OR "research librar*" OR "research informationist*"))
۷۷		Google Scholar and other resources

شدند. مطالعات مروری یا مطالعات اولیه‌ای که امکان دسترسی به متن کامل آن‌ها میسر نشد، مطالعاتی که به زبان فارسی یا انگلیسی نبودند و یا مطالعاتی که از نظر میزان ربط، پتانسیل پاسخ‌گویی به پرسش‌های پژوهش را نداشتند، کنار گذاشته شدند.

در گام بعد به منظور پاسخ‌گویی به پرسش‌های پژوهش، مطالعات واجد شرایط، مورد تحلیل محتوا قرار گرفت. در این گام بر اساس نوع فرایند یا فرایندهایی از مدیریت داده‌های پژوهشی که مدنظر مطالعه بوده، توانمندی‌ها و قابلیت‌های مورد اشاره استخراج شد. در گام بعد تیم پژوهش با هدف ایجاد زبان مشترک به کدگذاری این داده‌ها مبادرت کردند. بدین ترتیب که فرایندهای ذکر شده در پژوهش‌های مورد بررسی با مراحل چرخه‌ی حیات داده‌ی زیست‌پزشکی هاروارد تطبیق داده شد. بدین ترتیب که فرایندهای ذکر شده در متون با مراحل چرخه‌ی حیات داده‌ی زیست‌پزشکی هاروارد که شامل مراحل برنامه‌ریزی و طراحی، گردآوری و ایجاد، تجزیه و تحلیل و همکاری، ذخیره‌سازی و مدیریت، ارزیابی و آرشیو، اشتراک‌گذاری و انتشار، و دسترسی و استفاده‌ی مجدد است، تطبیق داده شد (۷).

باید گفت که ابتدا توانمندی‌ها و قابلیت‌ها بر اساس مشابهت مفهومی، ذیل یک کد مفهومی، دسته‌بندی شدند و سپس رابطه‌ی آن‌ها با هر مرحله از چرخه‌ی حیات که مورد نظر متون مورد بررسی بوده است، مشخص شد. بدین ترتیب

به منظور دستیابی به متون، ابتدا کلیدواژه‌های مربوط به دو مفهوم «مدیریت داده‌های پژوهشی» و «کتابدار داده» شناسایی و سپس مفاهیم هم‌ارز آن‌ها به کمک ابزارهای کنترل زبان، متون، مشورت با استادان و متخصصان حوزه شناسایی شد. در گام بعد استراتژی‌های جستجو، با استفاده از برچسب‌هایی که امکان جستجوی هم‌زمان در عنوان، چکیده و کلیدواژه‌ها را فراهم می‌کرد، بر اساس مقتضیات هر یک از پایگاه‌های اطلاعاتی PubMed، Scopus، Web of Science تدوین شد. همچنین برای یافتن مطالعات انگلیسی مرتبط در Google scholar نیز از جستجوی آزاد استفاده شد و نتایج به دست آمده بر اساس میزان ربط، مرتب‌سازی شد. برای یافتن مطالعات فارسی مرتبط از Google scholar و پایگاه‌های داخلی مانند ایرانداک، پایگاه مرکز اطلاعات علمی جهاد دانشگاهی (SID)، نمایه‌نامه‌ی استنادی جهان اسلام و Magiran استفاده شد.

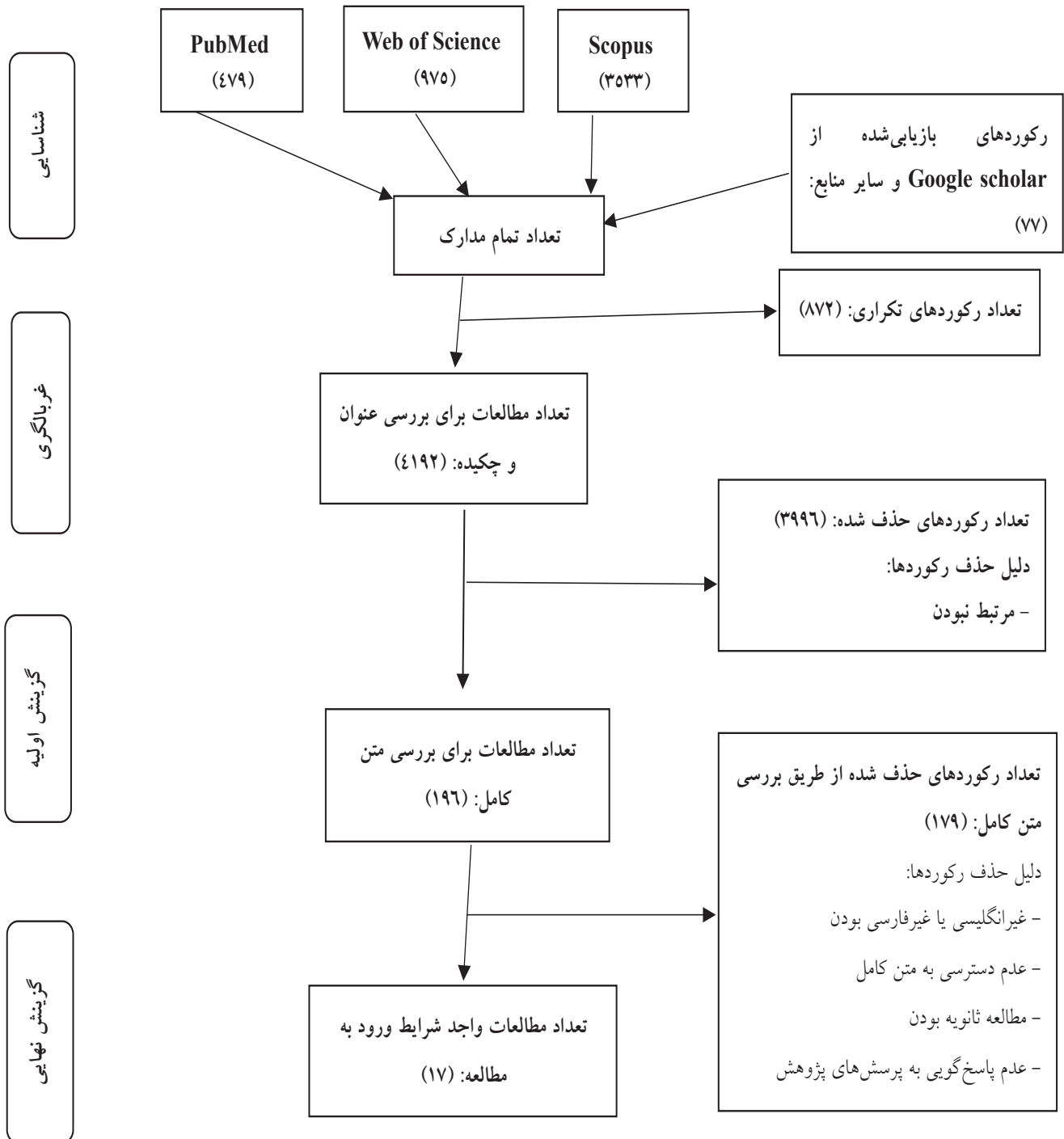
مطالعات بازیابی شده برای ساماندهی و نیز شناسایی و حذف موارد تکراری به نرم‌افزار Endnote وارد شدند. پس از حذف مقالات تکراری، ابتدا عنوان و چکیده مطالعات و در مرحله بعد، متن کامل آن‌ها بررسی گردید. در این مرحله مقالات به زبان‌های انگلیسی و فارسی که با استفاده از متدولوژی مشخص، به نحوی به بررسی توانمندی‌ها و قابلیت‌های مورد نیاز کتابداران برای ایفای نقش در مدیریت داده‌های پژوهشی پرداخته بودند، شناسایی و برای ورود به مطالعه و استخراج یافته‌های پژوهش واجد شرایط شناخته

در بیش از یک مرحله یا فرایند قرار گرفتند.

که ابتدا دو محقق جداگانه مشابهت‌های مفهومی را تعیین و کدگذاری کردند و سپس رابطه‌ی آن را با مراحل چرخه‌ی حیات داده‌های زیست‌پزشکی هاروارد مشخص کردند. سپس درصد توافق دو محقق محاسبه و در مواردی که عدم مطابقت بین نظر دو محقق مشاهده شد، نظر محقق سوم ملاک قرار گرفت. برخی از توانمندی‌ها و قابلیت‌ها محدود به فقط یک مرحله یا فرایند از چرخه‌ی حیات داده‌های پژوهشی نبودند؛ در چنین مواردی آن توانمندی‌ها و قابلیت‌ها

یافته‌ها

به‌منظور نشان دادن فرایند بررسی پایگاه‌های اطلاعاتی و غربالگری نتایج یافت شده برای دستیابی به مقالات واجد شرایط ورود به مطالعه، از نمودار جریانی PRISMA استفاده شده که در شکل ۱ نشان داده شده است.



شکل ۱: فرایند بررسی پایگاه‌های اطلاعاتی و غربالگری انجام‌شده

در مجموع تعداد ۵۰۶۴ مدرک از پایگاه‌های اطلاعاتی PubMed, Scopus, Web of Science و Google scholar و سایر منابع بازیابی شد. با استفاده از نرم‌افزار Endnote نسخه بیست‌ویکم، تعداد ۸۷۲ رکورد تکراری شناسایی و حذف شد و ۴۱۹۲ مدرک جهت بررسی عنوان و چکیده به مرحله بعد وارد شد. از این مدارک، تعداد ۳۹۹۶ مدرک در مرحله‌ی بررسی عنوان و چکیده با توجه به میزان ربط و معیارهای ورود و خروج از پژوهش حذف شدند و ۱۹۶ مدرک جهت بررسی متن کامل به مرحله بعد وارد شد. پس از بررسی متن کامل این مدارک، تعداد ۱۷ مقاله منتشرشده در مجلات وارد مطالعه شدند. مراحل بازیابی و انتخاب مدارک در شکل ۱ نشان داده شده است.

بر اساس معیارهای ورود و خروج، تعداد ۱۷ مقاله‌ی منتشرشده در بازه زمانی ۲۰۱۸ تا ۲۰۲۳ برای این مرور دامنه یافت شد. در جدول ۲، مشخصات هریک از مقالات نشان داده شده است.

جدول ۲: مشخصات مطالعات مورد بررسی

ردیف	عنوان مطالعه	نویسنده اول (نام خانوادگی، نام)	سال انتشار	کشور	منبع
۱	Partnering with health sciences libraries to address challenges in bioimaging data management and sharing	Silkotch, Christie	۲۰۲۳	آمریکا	۱۸
۲	Research data management services in academic libraries: A comparative study of South Asia and Southeast Asia	Sinha, Priyanka	۲۰۲۳	هند	۱۹
۳	Developing data services skills in academic libraries	Fuhr, Justin	۲۰۲۲	کانادا	۲۰
۴	Research data stewardship at the University of Hong Kong	Xiao, SiZhe	۲۰۲۲	چین	۲۱
۵	The landscape of research data management services in Malaysian academic libraries: librarians' practices and roles	Amanullah, Siti Wahida	۲۰۲۳	مالزی	۲۲
۶	پرکاربردترین مولفه‌های مدیریت داده‌های پژوهشی (موردپژوهی: کتابداران کتابخانه‌های دانشگاهی علوم پزشکی قطب ۷ کشور)	قهنویه، حسن	۲۰۲۲	ایران	۲۳
۷	Potential Roles for science librarians in research data Management: A Gap analysis	Bishop, Bradley Wade	۲۰۲۱	آمریکا	۲۴
۸	Practical considerations for a library's research data management services: The case of the National Institutes of health library	Kim, Soojung	۲۰۲۱	کره جنوبی	۲۵
۹	Conceptualizing data curation activities within two academic libraries	Lafferty-Hess, Sophia	۲۰۲۰	آمریکا	۲۶
۱۰	Research data management (RDM) services in libraries: Lessons for academic librarians in Nigeria	Nwabugwu, Mgbodichima Jummai	۲۰۲۰	نیجریه	۲۷
۱۱	The medical library association data services competency: A framework for data science and open science skills development	Federer, Lisa	۲۰۱۹	آمریکا	۲۸
۱۲	The role of a data librarian in academic and research libraries	Ohaji, Isaac K	۲۰۱۹	سودان	۱۳
۱۳	Awareness of research data management services at academic libraries in Jordan: Roles, responsibilities and challenges	Hamad, Faten	۲۰۱۹	عمان	۲۹
۱۴	Research data management and digital curation as a library activity	Chukanova, Svitlana	۲۰۲۰	اوکراین	۳۰
۱۵	Defining data librarianship: A survey of competencies, skills, and training	Federer, Lisa	۲۰۱۸	آمریکا	۱۲
۱۶	Enhancing collaboration across the research ecosystem: Using libraries as hubs for discipline-specific data experts	Ogier, Andrea	۲۰۱۸	آمریکا	۳۱
۱۷	Librarians' perspectives on the factors influencing research data management programs	Faniel, Ixchel M	۲۰۱۸	آمریکا	۳۲

در جدول ۲، مشخصات مطالعات مورد بررسی آمده است. همان‌طورکه در جدول نشان داده شده است، کشور متنوع نویسنده اول بیشتر مقالات، ایالات

به منظور نشان دادن قابلیت‌ها و توانمندی‌های اشاره‌شده در متون برای ایفای
 هاواراد و فرایندهای ذیل هر مرحله، جدول ۳ ترسیم شده است.

جدول ۳: یافته‌های ماصل از بررسی مطالعات

مرحله اصلی	فرایند	قابلیت‌ها و توانمندی‌ها	منبع
مرحله ۱: برنامه‌ریزی و طراحی (Plan and Design)	طرح مدیریت داده (Data management plan)	آشنایی با چگونگی تدوین طرح مدیریت داده	(۳۱-۲۹ و ۲۴ و ۲۲ و ۲۱ و ۱۸)
		توانایی کار با ابزارهای تدوین طرح مدیریت داده مانند: DMP online و DMP tool	(۳۰ و ۲۷ و ۱۸)
		آگاهی از استراتژی‌های توسعه و پشتیبانی مدیریت داده‌ها	(۲۹)
		آشنایی با سیاست‌ها و خط‌مشی‌های مربوط به اشتراک‌گذاری داده‌ها و آگاهی از الزامات حامیان مالی پژوهش‌ها و ناشران و مجلات برای پژوهشگران	(۳۲ و ۲۴)
	سیاست‌ها و انطباق با داده‌ها (Data policies and compliance)	آگاهی از سیاست‌ها و مسئولیت‌های اخلاقی و قانونی	(۲۹)
		توانایی مشارکت در سیاست‌گذاری داده‌ها	(۲۷ و ۲۳)
	نقش‌ها و مسئولیت‌ها (Roles and responsibilities)	آگاهی از سیاست‌ها و خط‌مشی‌های سازمان در خصوص مدیریت داده‌های پژوهشی	(۱۳)
		توانایی مشاوره به محققان در رعایت سیاست‌ها و تبعیت از خط‌مشی‌های مدیریت داده‌ها	(۲۶)
		توانایی تعریف خط‌مشی و دستورالعمل انجام مدیریت داده‌های پژوهشی	(۲۱)
		آشنایی با ذینفعان مختلف در فرایند مدیریت داده‌های پژوهشی	(۲۵)
	ساختارهای راهنما (Directory structures) و قرارداد نام‌گذاری فایل (File naming conventions)	آشنایی با مهارت‌های مورد نیاز برای وظایف مختلف مدیریت داده‌های پژوهشی	(۱۹)
		آشنایی با پروتکل‌های نام‌گذاری فایل‌ها	(۲۸)
		آشنایی با ساختار استاندارد فایل‌ها	(۳۰ و ۲۶ و ۲۳)
		توانایی قالب‌بندی داده‌ها	(۲۷ و ۲۵)
مرحله ۲: گردآوری و ایجاد (Collect and Create)	نوت‌بوک‌های آزمایشگاهی الکترونیکی (Electronic lab notebooks)	توانایی ارزیابی فرمت فایل	(۲۶)
		توانایی سازماندهی فایل‌ها	(۳۲ و ۲۳)
		آشنایی با ابزارها، روش‌ها، نرم‌افزارها و پلتفرم‌های مدیریت داده‌های پژوهشی	(۳۰ و ۲۷ و ۲۴ و ۱۸)
		آشنایی با ابزارهای همکاری مانند: Open Science Framework و protocols.io	(۷)
	ابزارها و نرم‌افزارهای مبتنی بر همکاری (Collaborative tools and software)	آشنایی با زبان‌های برنامه‌نویسی نظیر: جاوا اسکریپت، پایتون، R و ++C	(۲۵ و ۲۰)
		آشنایی با زبان نشانه‌گذاری (XML)	(۲۰)
	مستندسازی و فراداده (Documentation and metadata)	آشنایی با نرم‌افزارهای مدیریت منابع مانند: Mendeley و EndNote	(۲۷)
		آشنایی با ابزار تجزیه و تحلیل داده مانند: Watson, Zeppelin notebook, Jupiter notebook, Nvivo, REDCAP, Metacore, Qlucore studio	(۳۰ و ۲۲ و ۲۰ و ۱۸)
		آشنایی با چگونگی مستندسازی داده‌ها و فراداده‌ها	(۲۸ و ۲۱ و ۱۹ و ۲۶-۲۴ و ۳۰ و ۳۲)
		توانایی جمع‌آوری و فراهم‌آوری داده‌ها	(۲۸ و ۲۱)
تکرارپذیری (Reproducibility)	توانایی مصورسازی داده‌های پژوهشی	(۲۶ و ۲۲ و ۲۱)	
	آشنایی با استانداردهای فراداده	(۱۳)	
	توانایی استفاده از ابزارهای کاربردی فراداده برای بهبود بازیابی مجموعه داده‌ها	(۲۹ و ۱۳)	
	توانایی توصیف داده‌ها جهت بازیابی آسان‌تر در آینده	(۲۷ و ۲۶ و ۲۱)	
تکرارپذیری (Reproducibility)	توانایی مکان‌یابی مجموعه داده‌ها	(۲۹ و ۲۶ و ۱۸)	
	توانایی داده‌کاوی	(۳۱ و ۲۲)	
		ارایه مشاوره‌ی روش‌شناختی برای انجام پژوهش‌های دقیق و تکرارپذیر	(۱۹)

توانایی گردآوری، بازبینی و آماده‌سازی مجموعه داده‌های آماده تجزیه و تحلیل (توانایی ایجاد، قالب‌بندی، اعتباریابی، استانداردسازی، پاکسازی و مستندسازی مجموعه داده‌های آماده تجزیه و تحلیل)	مجموعه داده‌های آماده تجزیه و تحلیل (Analysis ready datasets)
آشنایی با ابزارهای تجزیه و تحلیل داده مانند: Zepelin notebook, Jupiter notebook, Nvivo و REDCAP, Metacore, Qlucore, Watson studio	مدیریت تصاویر (Image management)
ارایه آموزش تحلیل‌های آماری و روش‌های کدگذاری	کنترل نسخه (Version Control)
آشنایی با ابزارهای پردازش داده‌ها مثل: Data package creator	
آشنایی با پروتکل‌های نام‌گذاری فایل‌ها	امنیت داده (Data security)
آشنایی با نرم‌افزار کنترل نسخه	
توانایی نسخه‌سازی فایل داده‌ها	
توانایی برنامه‌ریزی برای حفظ امنیت و محرمانگی داده‌ها	
توانایی مدیریت داده‌های حساس	
آشنایی با ابزارها و تکنیک‌های دستکاری و آنالیز داده‌ها	
آگاهی از ذخیره‌سازی و بازیابی فراداده‌ها	
توانایی برنامه‌ریزی برای ذخیره‌سازی و نگهداری امن داده‌ها	
توانایی ایجاد آرشیو برای نگهداری و حفاظت داده‌ها و فراداده‌ها	نگهداری داده (Data retention)
آشنایی با راهنمای نگهداری منابع وب	
آشنایی با الزامات نگهداری داده‌ها	
توانایی آماده‌سازی زیرساخت‌های شبکه برای نگهداری داده‌ها	
محافظت و نگهداری طولانی مدت از فایل داده‌ها	
توانایی بررسی مجموعه داده‌های ذخیره شده برای یافتن و رفع مشکلات احتمالی	
آشنایی با فایل README	آرشیو و مدیریت رکوردها (Archives and records management)
ارایه مشاوره در زمینه ارزیابی و انتخاب داده‌ها برای حفظ و نگهداری	
توانایی ارزیابی داده‌ها برای حفظ و نگهداری	
توانایی ارزیابی داده‌ها برای انهدام و حذف دائمی داده‌ها	انهدام داده (Data destruction)
آشنایی با سیاست‌ها و قوانین اشتراک‌گذاری و استفاده‌ی مجدد از داده‌ها	اشتراک‌گذاری داده (Data sharing)
توانایی متقاعد کردن پژوهشگران برای اشتراک‌گذاری داده‌ها	
توانایی ایجاد دسترسی و استفاده‌ی مجدد از داده‌ها	
توانایی ارایه مشاوره‌ی به پژوهشگران در زمینه اشتراک‌گذاری داده‌ها	
آشنایی با سیاست‌ها، فرایندها و الزامات انتشار داده‌ها	دسترسی آزاد (Open Access)
آگاهی از سیاست‌های دسترسی آزاد و اشتراک داده‌های دسترسی آزاد	
آشنایی با منبع آنلاین خط‌مشی دسترسی آزاد Sherpa Romeo	
ارایه مشاوره در مورد استانداردها و واسطه‌گاه فراداده‌ها	قرارداد استفاده داده (Data use agreements)
آگاهی از سیاست‌های اشتراک‌گذاری و استفاده‌ی مجدد از داده‌ها	
آشنایی با ملاحظات قانونی و اخلاقی و اصول استناددهی به داده‌ها برای استفاده‌ی مجدد	مالکیت معنوی (Intellectual property)
مدیریت کتاب‌شناختی	
ارایه مشاوره و راهنمایی برای تسهیل استناد به داده‌ها مانند: ایجاد پلتفرم کتاب‌سنجی و تأثیر پژوهشی و ارایه معیارهای پژوهش	



(۱۸)	کمک به محققان برای تولید شناساگر مجموعه داده‌ها تا در دسترس باقی بمانند	برون‌دادهای علمی
(۲۹)	انتقال داده (از شی دیجیتال به واسپارگاه‌ها)	(Scholarly products)
(۲۳ و ۲۶)	ایجاد شناسه دیجیتال شی (DOI)	پیش چاپ و انتشار
-	-	(Preprints and publishing)
(۱۸ و ۲۴ و ۲۵ و ۲۷ و ۳۰)	آشنایی با واسپارگاه‌های داده‌های پژوهشی	واسپارگاه‌های داده
(۲۹ و ۳۱)	مدیریت واسپارگاه داده‌های سازمانی	(Data repositories)
(۲۹)	آگاهی از ابزار و تکنیک‌های راه‌اندازی واسپارگاه داده‌ها	
(۱۸ و ۲۱ و ۲۴)	آشنایی با چگونگی انتشار مجموعه داده در واسپارگاه	
(۲۱ و ۲۶)	توانایی اشتراک‌گذاری، دسترسی و استفاده مجدد از داده‌ها	
(۱۲ و ۱۳ و ۲۰)	توانایی برقراری ارتباط کلامی	
(۱۲)	مهارت‌های ارتباط نوشتاری	
(۱۲ و ۲۰)	توانایی انجام کار تیمی	
(۱۳ و ۲۰)	توانایی مدیریت پروژه	
(۱۲ و ۱۸ و ۲۸ و ۲۹ و ۳۱)	مهارت‌های آموزشی و اطلاع‌رسانی	
(۱۲)	مهارت‌های کتابخانه‌ای	
(۲۴)	آگاهی از فرایند پژوهش	
(۱۲ و ۱۸ و ۲۸ و ۳۱)	ایجاد آگاهی در بین محققان در مورد مدیریت داده‌های پژوهشی	
(۱۹)	پشتیبانی از شفافیت در جریان کار و یکپارچگی تحقیقات	
(۲۸ و ۲۹)	توانایی پشتیبانی از بهترین شیوه‌های تحقیق در طول چرخه‌ی حیات داده	
(۲۰)	آشنایی با نرم‌افزارهای ویرایشگر متن	
(۲۷)	آشنایی با ORCID Identifier	
(۲۷)	توانایی کار با Scopus author management	
(۲۷)	آشنایی با DOAJ.org	
(۲۹)	توانایی انجام مرور دامنه	
(۲۴)	توانایی مرور پیشینه	
(۲۷)	آشنایی با روش‌های تحقیق	
(۲۷)	آشنایی با شاخص‌های علم‌سنجی	

تصاویر» در ذیل مرحله سوم و فرایند «پیش‌چاپ و انتشار» در ذیل مرحله هفتم در هیچ‌یک از مطالعات، توانمندی و قابلیت مطرح نشده بود.

بحث

در این مرور دامنه، ۱۷ مطالعه به منظور استخراج توانمندی‌ها و قابلیت‌های مورد انتظار برای کتابدار داده به منظور ایفای نقش در مدیریت داده‌های پژوهشی، بررسی گردید و یافته‌های حاصل از آن بر اساس چرخه‌ی حیات داده‌های زیست‌پزشکی هاروارد گزارش شد. بیشترین توانمندی‌ها و قابلیت‌های مورد اشاره

همان‌گونه که در جدول ۳ نشان داده شده است، در مجموع ۹۲ توانمندی و قابلیت برای کتابدار داده از متون مورد بررسی استخراج شد که در ذیل ۲۳ فرایند از ۷ مرحله چرخه حیات داده‌های زیست‌پزشکی هاروارد قرار گرفت. بیشترین توانمندی‌ها و قابلیت‌های مورد اشاره در متون، ذیل مرحله دوم چرخه حیات یعنی «گردآوری و ایجاد» و کمترین میزان، ذیل مرحله هفتم یعنی «انتشار و استفاده‌ی مجدد» قرار گرفت. از میان فرایندها نیز بیشترین قابلیت‌ها و توانمندی‌های مورد اشاره در متون برای کتابدار داده، به فرایند «مستندسازی و فراداده» از فرایندهای مرحله دوم، اختصاص یافت. برای فرایند «مدیریت

اشاره کرده‌اند (۳۸) که به نظر می‌رسد کتابداران می‌توانند در تمامی این سطوح به دولت‌ها، سازمان‌ها و پژوهشگران یاری برسانند. با توجه به مزایای فراوان به اشتراک‌گذاری داده‌ها در راستای پیشبرد علم و پژوهش لازم است کتابداران به توانمندی‌های مرتبط با مدیریت اشتراک‌گذاری داده‌های پژوهشی مجهز باشند (۳۹). همچنین در برخی مطالعات نیز به اهمیت آشنایی کتابداران با الزامات مالی حامیان پژوهش‌ها اشاره شده است (۴۰ و ۳۵). برخی دیگر از پژوهشگران نیز به لزوم توانایی کتابداران در مشارکت به منظور سیاست‌گذاری داده‌ها توجه داشته‌اند (۴۲-۴۰ و ۳۷ و ۳۴). شاید یکی از مهمترین دلایل اهمیت مشارکت کتابدار داده در فرایند سیاست‌ها و انطباق با داده‌ها این باشد که با آگاهی و درک عمیق سیاست‌ها و خط‌مشی‌های تدوین‌شده بتوانند به نحو مناسبی به تدوین طرح مدیریت داده‌ها اقدام کنند.

«نقش‌ها و مسئولیت‌ها» سومین فرایند مرحله برنامه‌ریزی و طراحی است. در این فرایند، آشنایی با ذینفعان مختلف در فرایند مدیریت داده‌های پژوهشی (۲۵)؛ و آشنایی با مهارت‌های مورد نیاز برای وظایف مختلف مدیریت داده‌های پژوهشی (۱۹)؛ توانمندی‌ها و قابلیت‌های مورد انتظار کتابدار داده بوده است. برخی مطالعات مروری نیز به یافته‌های مشابه با یافته‌های این مرحله دست یافتند (۴۳ و ۳۸ و ۳۴). بنابراین در فرایند نقش‌ها و مسئولیت‌ها، کتابدار داده، در حقیقت وظیفه‌ی تقسیم وظایف و مسئولیت‌ها را در میان ذینفعان پروژه‌ی مدیریت داده‌ها دارد. برای این منظور در وهله‌ی اول باید با ذینفعان مختلف در پروژه‌ی مورد نظر آشنا باشد. در وهله‌ی بعد، از مهارت‌های مورد نیاز برای وظایف مختلف مدیریت داده‌های پژوهشی آگاه باشد تا بتواند مسئولیت‌های لازم را برای هر یک از ذینفعان تعیین کند.

فرایند «ساختارهای راهنما، قرارداد نام‌گذاری فایل» به‌عنوان آخرین فرایند مرحله برنامه‌ریزی و طراحی با هم مرتبط بوده و به‌همین دلیل به‌عنوان یک فرایند و در یک ردیف کنار هم قرار می‌گیرند. آشنایی با ساختار استاندارد فایل‌ها (۳۰ و ۲۶ و ۲۳)؛ و توانایی سازماندهی فایل‌ها (۳۲ و ۲۳)؛ از دیگر توانمندی‌ها و قابلیت‌های مورد اشاره در متون مورد بررسی بوده است. بنابراین کتابدار داده باید با پروتکل‌های نام‌گذاری فایل‌ها، و ساختار استاندارد آن‌ها آشنایی داشته باشد. همچنین کتابدار داده باید در سازماندهی فایل‌ها نیز توانمندی‌های لازم را داشته باشد، به‌گونه‌ای که دسترسی و استفاده‌ی دوباره از آن‌ها تسهیل شود.

• مرحله دوم: گردآوری و ایجاد

مربوط به مرحله دوم یعنی «گردآوری و ایجاد» و همچنین فرایند «مستندسازی و فراداده» این مرحله قرار گرفت. هیچ قابلیت و توانمندی برای فرایند «مدیریت تصاویر» در ذیل مرحله‌ی تجزیه و تحلیل و «پیش‌چاپ و انتشار» در ذیل مرحله‌ی «انتشار و استفاده‌ی مجدد» در متون یافت نشد. در ادامه به تفسیر و تحلیل توانمندی‌ها و قابلیت‌های یافت شده در متون برای کتابدار داده جهت ایفای نقش در هر یک از مراحل چرخه‌ی حیات داده‌های زیست‌پزشکی هاروارد مبادرت شده است.

• مرحله اول: برنامه‌ریزی و طراحی

مرحله برنامه‌ریزی و طراحی از مهمترین مراحل چرخه‌ی حیات داده زیست‌پزشکی هاروارد است؛ زیرا محصول آن طرح مدیریت داده است که وجود آن برای مدیریت داده‌های پژوهشی حیاتی است. در این مرحله، لازم است تمامی مراحل چرخه‌ی حیات داده، از ورود به سیستم تا بسته شدن پروژه و مدیریت منابع داده مد نظر قرار گیرد (۷). اولین فرایند این مرحله «طرح مدیریت داده» است. مطابق یافته‌های این پژوهش، آشنایی با چگونگی تدوین طرح مدیریت داده (۳۱-۲۹ و ۲۴ و ۲۲ و ۲۱ و ۱۸) و توانایی کار با ابزارهای تدوین آن (۳۰ و ۲۷ و ۱۸) از توانمندی‌ها و قابلیت‌های مورد انتظار کتابدار برای این فرایند است. همچنین در برخی مطالعات نیز به توانایی کتابدار داده برای کار با ابزارهای تدوین طرح مدیریت داده اشاره شده است (۳۷-۳۳). Chan و Wong در پژوهش خود به این نتیجه دست یافتند که با توجه به جدید بودن حوزه‌ی مدیریت داده‌های پژوهشی و ناآشنایی برخی پژوهشگران با آن، کتابدار داده باید علاوه بر آشنایی با چگونگی تدوین طرح مدیریت داده، بتواند ضرورت مدیریت داده‌ها را به پژوهشگران و موسسات پژوهشی شرح دهد و از فواید تدوین طرح مدیریت داده دفاع کند (۳۶). شاید بتوان گفت فرایند طرح مدیریت داده، قسمت کلیدی مدیریت داده‌های پژوهشی است؛ زیرا با انجام این فرایند نقش زیربنایی برای اجرای سایر فرایندهای مدیریت داده‌های پژوهشی دارد.

«سیاست‌ها و انطباق با داده‌ها» دومین فرایند مرحله‌ی برنامه‌ریزی و طراحی است. آشنایی با سیاست‌ها و خط‌مشی‌های مربوط به اشتراک‌گذاری داده‌ها و آگاهی از الزامات حامیان مالی پژوهش‌ها، ناشران و مجلات برای پژوهشگران (۳۲ و ۲۴)؛ و توانایی مشارکت در سیاست‌گذاری داده‌ها (۲۷ و ۲۳)؛ از دیگر توانمندی‌ها و قابلیت‌های مورد انتظار کتابدار داده بوده است. عصاره و همکاران در مطالعه‌ی خود به سه سطح سیاست‌گذاری کلان، میانه و خرد



در مرحله‌ی گردآوری و ایجاد، مجموعه داده‌ها گردآوری و سپس سازماندهی و یکپارچه‌سازی می‌شوند. وقتی نوبت به مدیریت داده‌ها و اسناد در یک پروژه می‌رسد، باید فرایندهایی را اجرا کرد که متناسب با ویژگی‌های آن پروژه باشد. مستندسازی مناسب داده‌ها اطلاعات لازم برای درک و تفسیر کامل داده‌ها را فراهم می‌کند؛ همچنین فرایند گردآوری فراداده‌ها کشف، استفاده‌ی مجدد، تکرارپذیری، حفظ و آرشیو داده‌ها را تسهیل می‌کند. این مرحله و مرحله‌ی «تجزیه و تحلیل و همکاری»، ارتباط تنگاتنگی با یکدیگر دارند؛ به طوری که در برخی موارد این دو مرحله ادغام و از آن با عنوان «گردآوری و تجزیه - تحلیل» یاد شده است (۷). اولین فرایند این مرحله «ابزارها و نرم‌افزارهای مبتنی بر همکاری» است. مطابق یافته‌های پژوهش، آشنایی با ابزارها، روش‌ها، نرم‌افزارها و پلتفرم‌های مدیریت داده‌های پژوهشی (۳۰ و ۲۷ و ۲۴ و ۱۸)؛ و آشنایی با ابزارهای تجزیه و تحلیل داده (۳۰ و ۲۲ و ۲۰ و ۱۸)؛ از توانمندی‌ها و قابلیت‌های مورد انتظار کتابدار داده بوده است. در برخی از مطالعات مروری به ضرورت یادگیری مهارت‌های برنامه‌نویسی اشاره شده است (۴۵ و ۴۴). Pinto و Semeler در پژوهش‌های خود اشاره می‌کنند که مرتبط‌ترین دانش فناوری برای کتابداران داده شامل زبان‌های برنامه‌نویسی، طراحی و ساختار پایگاه داده، رابط برنامه‌نویسی کاربردی برای بازیابی داده‌ها، طراحی کاربرمحور، ابزارهای پردازش زبان طبیعی، و اینترنت اشیا است؛ اما یک کتابدار داده نیازی به برنامه‌نویس شدن ندارد؛ بلکه باید علاقمند به یادگیری زبان‌ها و منطق برنامه‌نویسی باشد؛ و در مورد عملکرد طیف وسیعی از نرم‌افزارهای علمی که برای تبدیل داده‌ها استفاده می‌شود، آگاهی داشته باشد (۱۵ و ۱۴). بنابراین به نظر می‌رسد نیازی نباشد که کتابدار داده، برنامه‌نویس حرفه‌ای باشد و آشنایی کلی با زبان‌های برنامه‌نویسی هم کفایت می‌کند. Creamer و همکاران در کتاب خود به ضرورت آشنایی با ابزارهای دیگری مانند data interview و data curation profiles پرداختند. data interview ابزاری است که در اختیار کتابدار قرار می‌گیرد تا به پژوهشگران کمک کند تا اهداف خود را برای حفظ و اشتراک‌گذاری داده‌های خود پس از پروژه کشف کنند؛ و کتابداران از ابزار data curation profiles برای مصاحبه با پژوهشگران و بررسی چرخه‌ی حیات یک پروژه‌ی تحقیقاتی و الزامات و مسایل مربوط به داده‌های حاصل از پژوهش استفاده می‌کنند (۳۷). آشنایی با ابزارهای مدیریت کتاب‌شناختی نیز از دیگر قابلیت‌هایی است که باید کتابدار داده از آن برخوردار باشد (۳۵). به طور کلی، با توجه به این که پژوهش یک کار تیمی است،

برای مدیریت داده‌ها نیاز به ابزارها و نرم‌افزارهایی است که هم امکان دسترسی و استفاده از فایل‌ها، داده‌ها و منابع به صورت مشترک را فراهم کند و هم برای پژوهشگران، امکان مدیریت داده‌ها و منابع‌شان را در طول انجام پژوهش میسر کند. کتابداران داده با آموختن مهارت کار با این ابزارها و نرم‌افزارها می‌توانند در راستای این دو مهم به پژوهشگران یاری برسانند.

دومین فرایند مرحله گردآوری و ایجاد «نوت‌بوک‌های آزمایشگاهی الکترونیکی» است. نوت‌بوک آزمایشگاهی ابزاری برای ثبت سوابق است که توسط محققان برای مستندسازی آزمایش‌ها، مشاهدات و نتایج حاصل از آن‌ها استفاده می‌شود. این ابزار می‌تواند کاغذی یا الکترونیکی باشد؛ اما نوع الکترونیکی آن می‌تواند ویژگی‌های پیشرفته‌ای را برای مدیریت داده‌ها فراهم کند. آشنایی با نوت‌بوک‌های آزمایشگاهی الکترونیکی از دیگر توانمندی‌ها و قابلیت‌های مورد انتظار برای کتابدار داده است (۲۵ و ۱۸).

سومین فرایند مرحله گردآوری و ایجاد «مستندسازی و فراداده» است. از دیگر توانمندی‌ها و قابلیت‌های اشاره شده در متون برای کتابدار داده می‌توان آشنایی با چگونگی مستندسازی داده‌ها و فراداده‌ها (۳۲ و ۳۰ و ۲۶ و ۲۴ و ۲۱ و ۱۹)؛ توانایی مصورسازی داده‌های پژوهشی (۲۶ و ۲۲ و ۲۱)؛ توصیف داده‌ها جهت بازیابی آسان‌تر در آینده (۲۷ و ۲۶ و ۲۱)؛ و مکان‌یابی مجموعه داده‌ها (۲۹ و ۲۶ و ۱۸) را نام برد. چند مطالعه که به صورت مروری به موضوع مدیریت داده‌های پژوهشی پرداخته‌اند، به ضرورت آشنایی کتابداران با استانداردها و طرح‌های فراداده برای نقش‌آفرینی در این حوزه اشاره نمودند (۴۵ و ۴۴ و ۳۹). Cendon و Almeida در مقاله مروری خود به رده‌بندی خدمات مدیریت داده‌های پژوهشی پرداختند. آن‌ها علاوه بر یافته‌های پژوهش حاضر، به ارائه خدمات دیکشنری داده نیز اشاره نمودند و آن را در زیررده‌ی مستندسازی داده و فراداده قرار دادند (۳۵). این فرایند بیشترین میزان توجه را در مطالعات مورد بررسی به خود اختصاص داده بود. بنابراین می‌توان گفت کسب توانمندی‌ها و قابلیت‌های یافت شده مربوط به این فرایند برای کتابدار داده ضرورت بیشتری نسبت به سایر مراحل و فرایندهای چرخه‌ی حیات داده‌های زیست‌پزشکی هاروارد دارد.

آخرین فرایند مرحله گردآوری و ایجاد «تکرارپذیری» است. ارائه مشاوره‌ی روش‌شناختی برای انجام پژوهش‌های دقیق و تکرارپذیر (۱۹)، از دیگر توانمندی‌ها و قابلیت‌های مورد انتظار برای کتابدار داده است. با توجه به اهمیت تکرارپذیری در امر پژوهش، کتابدار داده باید در این زمینه مهارت‌های خود

است. در مورد این فرایند، در پژوهش‌های مورد بررسی، داده‌ای یافت نشد. اما می‌توان آشنایی با اصول مدیریت داده‌های تصویری پژوهش‌ها، آشنایی با ابزارها و نرم‌افزارهای ضبط و مدیریت تصاویر، و آشنایی با شیوه‌های گردآوری تصاویر (اسکن داخلی یا عکاسی، دیجیتالی کردن، یا خرید از منابع خارجی) را برای نقش‌آفرینی کتابدار داده در نظر گرفت.

«کنترل نسخه» آخرین فرایند مرحله تجزیه و تحلیل و همکاری است. در این فرایند، آشنایی با پروتکل‌های نام‌گذاری فایل‌ها (۲۸)؛ آشنایی با نرم‌افزار کنترل نسخه (۲۶ و ۲۴)؛ توانایی نسخه‌سازی فایل داده‌ها (۳۰ و ۲۸)؛ توانمندی‌ها و قابلیت‌های مورد اشاره در متون مورد بررسی برای کتابدار داده بوده است. به نظر می‌رسد که کتابداران داده با اتکا به توانمندی‌ها و قابلیت‌هایشان در کنترل نسخه، بتوانند پژوهشگران را در راستای مدیریت پرونده‌های متعدد پژوهشی یاری برسانند.

• مرحله چهارم: ذخیره‌سازی و مدیریت

هر مرحله از چرخه‌ی حیات داده‌های زیست‌پزشکی هارواد حول محور مدیریت ذخیره‌سازی داده‌ها می‌چرخد. برای حفظ امنیت داده‌ها و مطابق با پروتکل‌های ایمنی، حفاظت و نگهداری مناسب از ذخیره‌سازی در طول چرخه‌ی حیات بسیار مهم است. در چرخه‌ی حیات داده‌ی زیست‌پزشکی هاروارد این مرحله و مرحله بعد یعنی «ارزیابی و آرشیو» با هم پیوند داده شده‌اند و مفاهیم و فرایندهای این دو مرحله، ذیل یک مفهوم واحد با عنوان «ذخیره‌سازی و ارزیابی» شرح داده شده‌است (۷). از این رو بحث در مورد یافته‌ها و توضیحات بیشتر ذیل مرحله پنجم ارائه می‌شود.

• مرحله پنجم: ارزیابی و آرشیو

این مرحله، متمرکز بر شناسایی سوابق پژوهشی ضروری و ارزیابی آن‌ها به منظور حفظ و نگهداری است. به طور کلی برای هر پروژه پژوهشی، نیاز به مدیریت مناسب داده‌ها و امنیت آن‌ها وجود دارد. اولین فرایند، مرحله‌ی ذخیره‌سازی و ارزیابی، «امنیت داده» است. از جمله توانمندی‌ها و قابلیت‌هایی که کتابداران داده برای ایفای نقش در این فرایند باید داشته باشند: توانایی برنامه‌ریزی برای حفظ امنیت و محرمانگی داده‌ها (۳۰) و توانایی مدیریت داده‌های حساس (۲۶ و ۲۱) است. عروجی و همکاران در پژوهشی به این نتیجه دست یافتند که حریم خصوصی و محرمانگی داده‌ها باید جزئی از برنامه درسی رشته‌ی علم اطلاعات و دانش‌شناسی باشد (۴۵). آن‌ها در پژوهش دیگری به این مهم اشاره

را افزایش دهد تا هم بتواند به پژوهشگران خدمات مشاوره‌ای ارائه دهد هم به‌عنوان عضوی از تیم پژوهش در راستای رعایت اصل تکرارپذیری پژوهش متمرکس باشد.

• مرحله سوم: تجزیه و تحلیل و همکاری

این مرحله و مرحله‌ی «گردآوری و ایجاد»، ارتباط تنگاتنگی با یکدیگر دارند (۷). تجزیه و تحلیل داده‌ها فرایند پاکسازی، تبدیل و مدل‌سازی داده است که با هدف کشف اطلاعات مفید، اطلاع‌رسانی در خصوص نتایج، و پشتیبانی از تصمیم‌سازی‌ها صورت می‌گیرد. اولین فرایند مرحله‌ی تجزیه و تحلیل و همکاری «مجموعه داده‌های آماده‌ی تجزیه و تحلیل» است. در این فرایند، توانایی گردآوری، بازبینی و آماده‌سازی مجموعه داده‌های آماده‌ی تجزیه و تحلیل (توانایی ایجاد، قالب‌بندی، اعتباریابی، استانداردسازی، پاکسازی و مستندسازی مجموعه داده‌های آماده‌ی تجزیه و تحلیل) (۲۸ و ۲۷ و ۲۵ و ۱۹ و ۱۳)؛ و آشنایی با ابزارهای تجزیه و تحلیل داده (۳۰ و ۲۲ و ۲۰ و ۱۸)، از جمله توانمندی‌ها و قابلیت‌های مورد انتظار کتابدار داده بوده است. Semeler و همکاران در مطالعه‌ی خود استدلال کردند که علم داده می‌تواند به‌عنوان یک حوزه‌ی مطالعه‌ی جدید برای کتابداران داده در نظر گرفته شود که به دنبال مشارکت در مسایل مربوط به مدیریت و تجزیه و تحلیل داده‌ها هستند؛ و این وظیفه کتابداران است که در پردازش داده‌ها مهارت کسب کنند تا وظایف یک متخصص داده را انجام دهند (۱۵). اصنافی و عبدی نیز در یک مطالعه مروری به این نتیجه دست یافتند که نقش اصلی یک کتابدار داده، جستجوی تخصص فنی برای گردآوری، تجزیه و تحلیل و تفسیر داده‌هاست (۴۴). همچنین چندین مطالعه‌ی مرور نظام‌مند دیگر که به بررسی متون مرتبط با مدیریت داده‌های پژوهشی پرداخته بودند به ضرورت آشنایی کتابدار داده با چگونگی تجزیه و تحلیل داده و ابزارهای آن اشاره نمودند (۴۱ و ۳۹ و ۳۴). می‌توان این‌گونه استدلال کرد که از آنجایی که داده‌ها اغلب دچار به هم‌ریختگی و آشفتگی هستند، هدف از مدیریت داده‌ها این است که آن‌ها قبل از تجزیه و تحلیل، پاکسازی شوند. ایجاد، قالب‌بندی، اعتباریابی، استانداردسازی، پاکسازی داده، و مستندسازی مفاهیمی هستند که بیانگر فرایندهای ضروری برای ایجاد «مجموعه داده‌های آماده‌ی تجزیه و تحلیل» هستند. همچنین آشنایی با ابزارهای تجزیه و تحلیل، و پردازش داده‌ها می‌تواند به پیشبرد این فرایند کمک کند.

«مدیریت تصاویر» دومین فرایند مرحله‌ی تجزیه و تحلیل و همکاری

کرده‌اند که امنیت داده‌ها از مهم‌ترین موارد در حوزه‌ی سواد مدیریت داده‌های پژوهشی محسوب می‌شود (۳۹). نتایج مطالعه‌ی Semeler و Pinto نیز مهر تأییدی بر این توانمندی و قابلیت کتابدار داده گذاشته است (۱۴).

دومین فرایند مرحله‌ی ذخیره‌سازی و ارزیابی، «نگهداری داده» است. آگاهی از چگونگی ذخیره‌سازی و ارزیابی فراداده‌ها (۳۲ و ۲۶)؛ توانایی برنامه‌ریزی برای ذخیره‌سازی و نگهداری امن داده‌ها (۳۰ و ۲۸)، از جمله توانمندی‌ها و قابلیت‌های مورد انتظار برای کتابدار داده بوده است. مجیدی و همکاران نیز در پژوهش خود به این نتیجه رسیدند که کارکردهای حفاظت و آرشیو در مدیریت داده‌های پژوهشی در محیط‌های علمی و دانشگاهی ضرورت دارد (۴۳). در فرایند «نگهداری داده‌ها» باید مشخص شود که داده‌ها کجا و برای چه مدت ذخیره شوند. پس برای انجام صحیح آن آگاهی از چگونگی ذخیره‌سازی و ارزیابی فراداده‌ها، توانایی حفظ و نگهداری از خروجی‌های پژوهش‌ها، توانایی برنامه‌ریزی برای ذخیره‌سازی و نگهداری امن داده‌ها، ایجاد آرشیو برای نگهداری و حفاظت داده‌ها و فراداده‌ها، و آشنایی با الزامات نگهداری داده‌ها ضروری می‌باشد.

سومین فرایند مرحله‌ی ذخیره‌سازی و ارزیابی، «آرشیو و مدیریت رکورد» است. بر اساس یافته‌های این پژوهش، توانایی ارزیابی داده‌ها برای حفظ و نگهداری (۳۲ و ۲۷)؛ آشنایی با فایل README (۲۱)؛ و رایج مشاوره در زمینه ارزیابی و انتخاب داده‌ها برای حفظ و نگهداری (۲۷)؛ از دیگر توانمندی‌ها و قابلیت‌های کتابدار داده است. شیخ و همکاران در پژوهش خود به نقش کتابدار داده در حفظ داده‌ها یا فراهم کردن بستری برای تبادل داده با دیگران دست یافتند که نتیجه‌ای مشابه با پژوهش حاضر داشت (۳۴). فرایند «آرشیو و مدیریت رکورد» جهت مشخص شدن این که «چه چیزی باید حفظ شود» انجام می‌گیرد. بنابراین آشنایی با فایل README، رایج مشاوره در زمینه ارزیابی و انتخاب داده‌ها برای حفظ و نگهداری، ارزیابی و انتخاب داده‌ها برای واسپارگاه‌ها، و ارزیابی داده‌ها برای حفظ و نگهداری از جمله فعالیت‌هایی است که کتابدار داده باید در انجام آن‌ها توانایی لازم را داشته باشد.

آخرین فرایند مرحله ذخیره‌سازی و ارزیابی «انهدام داده» است. کتابدار داده باید توانایی ارزیابی داده‌ها برای انهدام یا حذف دائمی داده‌ها (۳۲ و ۲۷) را داشته باشد. مسئولیت پژوهشگر پس از پایان پژوهش در قبال داده‌هایش، و دانستن این که کدام داده‌ها و سوابق برای انهدام و تخریب بی‌خطر هستند، ممکن است پیچیده باشد. بنابراین کتابدار داده باید در این زمینه برای کمک به پژوهشگران

توانایی لازم را داشته باشد.

• مرحله ششم: اشتراک‌گذاری و توزیع

داده‌ها با توجه به الزامات به اشتراک‌گذاری داده در واسپارگاه‌ها، به اشتراک گذاشته می‌شوند. اشتراک‌گذاری داده‌ها برای ترجمان سریع نتایج تحقیقات به دانش، محصولات و رویه‌ها در راستای ارتقای سلامت انسان ضروری است (۷). این مرحله و مرحله «انتشار و استفاده مجدد» بسیار در هم تنیده هستند و نمی‌توان مرزی میان این مراحل قایل شد. از این رو بحث در مورد یافته‌ها و توضیحات بیشتر ذیل مرحله‌ی بعد ارایه شده است.

• مرحله هفتم: انتشار و استفاده مجدد

در این مرحله تلاش می‌شود تا اطمینان حاصل شود که تلاش‌های داده‌ای محققان پیشین، به‌طور گسترده توسط سایر محققان بهره‌برداری شود. همان‌گونه که در راهنمای چرخه‌ی حیات داده‌ی زیست‌پزشکی هاروارد مشهود است، مفاهیم و فرایندهای این مرحله و مرحله قبل، ذیل یک مفهوم واحد با عنوان «اشتراک‌گذاری و انتشار» شرح داده شده است (۷). اولین فرایند مرحله اشتراک‌گذاری و انتشار «اشتراک‌گذاری داده» است. در این فرایند، آشنایی با سیاست‌ها و قوانین اشتراک‌گذاری و استفاده‌ی مجدد از داده‌ها (۲۷ و ۲۴ و ۲۳ و ۲۱ و ۱۸)؛ توانایی متقاعد کردن پژوهشگران برای اشتراک‌گذاری داده‌ها (۲۸ و ۱۸) و توانایی ایجاد دسترسی و استفاده‌ی مجدد از داده‌ها (۲۶ و ۲۳)، از جمله توانمندی‌ها و قابلیت‌های مورد انتظار برای کتابدار داده بوده است. عصاره و همکاران نیز در پژوهش خود این‌گونه نتیجه گرفتند که در عصری که با پارادایم جدیدی از پژوهش روبروست، کتابخانه‌ها در تلاش هستند خدمات پژوهشی جدیدی را برای حمایت از پژوهشگران رایج دهند و نگرانی‌های پژوهشگران را در ارتباط با اشتراک‌گذاری داده‌ها درک کنند (۳۸). عروجی و همکاران، و شیخ و همکاران نیز به نتیجه‌ای مشابه دست یافتند (۳۹ و ۳۴). بنابراین کتابدار داده از سوی باید با سیاست‌ها و قوانین اشتراک‌گذاری و استفاده‌ی مجدد از داده‌ها آشنایی داشته باشد و از سوی دیگر با دغدغه‌ها و نگرانی‌های پژوهشگران در این زمینه آشنا باشد تا بتواند با رایج مشاوره‌ای هدفمند در این زمینه نقش‌آفرینی کند.

فرایند دوم مرحله‌ی اشتراک‌گذاری و انتشار، «دسترسی آزاد» است. در این فرایند، آگاهی از سیاست‌های دسترسی آزاد و اشتراک‌های دسترسی آزاد (۲۷ و ۲۳)، و آشنایی با Sherpa Romeo منبع آنلاین خط مشی دسترسی

ششمین فرایند مرحله‌ی اشتراک‌گذاری و انتشار، «پیش‌چاپ و انتشار» است. در مورد این فرایند، توانمندی و قابلیت برای کتابدار داده در متون مورد بررسی یافت نشد. واسپارگاه‌های پیش‌چاپ، فرصتی برای به اشتراک گذاشتن دست‌نوشته‌ها و مقالات کاری قبل از انتشار مجله فراهم می‌کنند. سرورهای اصلی پیش‌چاپ شامل یک تالار گفتگوی خاص بازخورد است که به پژوهشگران اجازه می‌دهد تا نظرات، نقد و بررسی و ارزیابی‌های شفاف روی نسخه‌های خطی پیش‌چاپ را به صورت آنلاین ارائه کنند (۴۷). بنابراین به نظر می‌رسد که آشنایی کتابداران با واسپارگاه‌ها و قوانین حاکم بر آن‌ها و فرصت‌هایی که ارائه می‌دهند می‌تواند یاری بخش پژوهشگران در انتشار نتایج پژوهش‌هایشان در این مخازن ارزشمند باشد.

آخرین فرایند مرحله‌ی اشتراک‌گذاری و انتشار، «واسپارگاه داده» است. در این فرایند، آشنایی با واسپارگاه‌های داده‌های پژوهشی (۳۰ و ۲۷ و ۲۵ و ۲۴ و ۱۸)، آشنایی با چگونگی انتشار مجموعه داده در واسپارگاه (۲۴ و ۲۱ و ۱۸)؛ دیگر توانمندی‌ها و قابلیت‌های مورد انتظار کتابدار داده در متون مورد بررسی بوده است. عروجی و همکاران، و مجیدی و همکاران نیز در پژوهش‌هایشان به اهمیت آگاهی کتابدار داده از چگونگی راه‌اندازی واسپارگاه‌ها اشاره کرده‌اند (۴۵ و ۴۳). شیخ و همکاران نیز عنوان کرده‌اند که کتابداران باید بتوانند پژوهشگران را در راستای حفظ داده‌ها از طریق توسعه واسپارگاه‌های داده راهنمایی کنند (۳۴). همچنین Semeler و Pinto نیز در مقاله‌ی خود توصیه می‌کنند که کتابداران داده با واسپارگاه‌های داده‌های پژوهشی کار کنند تا از پژوهشگران از آغازین مراحل تحقیقات علمی حمایت کرده، به فرایند مستندسازی کمک کنند و اطمینان حاصل کنند که داده‌های پژوهش برای مدت طولانی حفظ و قابل استفاده مجدد باشند (۱۴). بنابراین کتابدار داده جهت ایفای نقش در این فرایند ضروری است که ضمن آشنایی با واسپارگاه‌های ملی و بین‌المللی داده‌ها، با چگونگی راه‌اندازی و مدیریت واسپارگاه داده و نیز چگونگی انتشار داده در آن آشنایی داشته باشد تا بتواند نقش مؤثری در این فرایند از مدیریت داده‌های پژوهشی ایفا کند.

• ویژگی‌های عمومی

توانایی برقراری ارتباط کلامی (۲۰ و ۱۳ و ۱۲)؛ مهارت‌های آموزشی و اطلاع‌رسانی (۳۱ و ۲۹ و ۲۸ و ۱۸ و ۱۲) و توانایی ایجاد آگاهی در بین محققان در مورد مدیریت داده‌های پژوهشی (۳۱ و ۲۸ و ۱۸ و ۱۲)؛ از جمله توانمندی‌ها و قابلیت‌های مورد انتظار کتابدار داده است. Semeler و همکاران استدلال

آزاد (۳۰ و ۲۷)؛ توانمندی‌ها و قابلیت‌های مورد انتظار کتابدار داده بوده است. Almeida و Cendon، و نیز Creamer و همکاران استدلال کردند که در این فرایند، کتابدار داده باید تلاش کند تا از سیاست‌های دسترسی آزاد و اشتراک داده‌های دسترسی آزاد آگاهی یابد (۳۷ و ۳۵). بنابراین آگاهی کتابداران داده در این زمینه، همچنین آشنایی با منابع و ابزارهای این حوزه می‌تواند در زمینه‌ی پیشبرد مدیریت داده‌های پژوهشی مفید باشد.

سومین فرایند مرحله‌ی اشتراک‌گذاری و انتشار، «قرارداد استفاده‌ی داده» است. قرارداد استفاده‌ی داده‌ها یک قرارداد الزام آور بین سازمان‌هاست که بر انتقال و استفاده از داده‌ها حاکم است. شرایط و ضوابط آن بسته به قوانین و مقررات حاکم بر نوع خاص داده و همچنین سیاست‌ها و یا الزامات ارائه‌دهنده متفاوت است. ارائه مشاوره به پژوهشگران در مورد استانداردها و واسپارگاه فراداده‌ها (۲۷)؛ از جمله توانمندی‌ها و قابلیت‌های مورد انتظار کتابدار داده بوده است. Creamer و همکاران نیز این‌گونه استدلال کرده‌اند که کتابدار داده باید از استانداردها و قوانین واسپارگاه‌ها آگاهی داشته باشد تا بتواند در این زمینه به پژوهشگران مشاوره دهد (۳۷). بنابراین داشتن توانمندی و قابلیت کتابدار در این فرایند در پیشبرد سایر فرایندها اثرگذار است.

چهارمین فرایند مرحله‌ی اشتراک‌گذاری و انتشار، «مالکیت معنوی» است. آشنایی با ملاحظات قانونی و اخلاقی و اصول استناددهی به داده‌ها برای استفاده‌ی مجدد (۳۰ و ۲۹ و ۲۵ و ۲۴ و ۱۸)؛ آگاهی از سیاست‌های اشتراک‌گذاری و استفاده‌ی مجدد از داده‌ها (۲۶ و ۲۱) از دیگر توانمندی‌ها و قابلیت‌های مورد انتظار کتابدار داده است. عروجی و همکاران نیز در پژوهش خود به اهمیت آگاهی از حقوق مالکیت اخلاقی مربوط به اشتراک‌گذاری داده‌ها در سواد مدیریت داده‌های پژوهشی اشاره کرده‌اند (۳۹). بنابراین آشنایی کتابدار داده با ملاحظات قانونی و اخلاقی و اصول استناددهی به داده‌ها برای استفاده‌ی مجدد، همچنین آگاهی از سیاست‌های اشتراک‌گذاری و استفاده‌ی مجدد از داده‌ها می‌تواند در این فرایند، کارآمد باشد.

پنجمین فرایند مرحله‌ی اشتراک‌گذاری و انتشار، «برون‌داده‌های علمی» است. ایجاد شناسه دیجیتال شی (DOI) (۲۶ و ۲۳) از دیگر توانمندی‌ها و قابلیت‌های مورد انتظار کتابدار داده است. مهم است که کتابداران بتوانند در راستای اختصاص شناسه‌های دائمی به برون‌داده‌های علمی پژوهشگران به‌منظور تسهیل دسترسی پذیرسازی مجموعه داده‌های آن‌ها یاری برسانند (۴۶).



کردند اولین مجموعه از مهارت‌های مورد نیاز کتابداران داده شامل مهارت‌های بین فردی و ویژگی‌های رفتاری، مهارت‌های ارتباطی در قالب ارتباط نوشتاری رسمی و نگارش اسناد فنی و مطالعات موردی است (۱۵). برخی پژوهشگران نیز توانایی برقراری ارتباط و داشتن مهارت‌های آموزشی را برای مدیریت داده‌های پژوهشی ضروری دانسته‌اند (۴۴ و ۳۹ و ۳۸). بنابراین می‌توان گفت که توانمندی در برقراری ارتباط کلامی با پژوهشگران و داشتن مهارت در ارائه آموزش و مشاوره به آنان از مهم‌ترین ویژگی‌های عمومی است که کتابدار داده را قادر می‌سازد تا نقش خود را در مراحل و فرایندهای مختلف مدیریت داده‌های پژوهشی به بهترین نحو ایفا کند.

نتیجه‌گیری

در این مطالعه‌ی مرور دامنه، با بررسی عمیق متون، توانمندی‌ها و قابلیت‌های مورد انتظار کتابدار داده شناسایی شد. در مجموع، ۹۲ توانمندی و قابلیت شناسایی شد که در ۲۳ فرایند از ۷ مرحله‌ی چرخه‌ی حیات داده‌های زیست‌پزشکی هاروارد دسته‌بندی شدند. بیشترین تعداد توانمندی‌ها و قابلیت‌ها در مرحله دوم یعنی «گردآوری و ایجاد» و فرایند «مستندسازی و فراداده» آن قرار گرفت. با توجه به اهمیت روزافزون مدیریت داده‌های پژوهشی در سطوح کلان و خرد، نتایج این پژوهش می‌تواند اطلاعات ارزشمندی در خصوص شایستگی‌های مورد انتظار کتابدار داده برای تصمیم‌سازان، سیاست‌گذاران و برنامه‌ریزان فراهم کند. دست‌اندرکاران بازنگاری برنامه‌های آموزشی کتابداری و اطلاع‌رسانی

پزشکی و همچنین علم اطلاعات و دانش‌شناسی می‌توانند از نتایج این پژوهش برای بررسی برنامه‌های آموزشی مقاطع کارشناسی، کارشناسی ارشد و دکتری استفاده کنند تا ظرفیت موجود این مقاطع برای تربیت دانش‌آموختگانی مجهز به توانمندی‌ها و قابلیت‌های مورد انتظار کتابدار داده را ارزیابی کنند. همچنین، متولیان آموزش ضمن خدمت در سازمان‌های ذی‌ربط می‌توانند از یافته‌های این پژوهش برای طراحی برنامه‌های بازآموزی به‌منظور توانمندسازی کتابداران شاغل در دانشگاه‌ها و مراکز تحقیقاتی برای ایفای نقش در مراحل مختلف مدیریت داده‌های پژوهشی استفاده کنند.

باتوجه به این‌که این مطالعه یک مطالعه‌ی مروری است و محدودیت‌های خاص این نوع مطالعات بر آن حاکم است، پیشنهاد می‌شود تا با انجام پژوهش‌های کیفی و بهره‌گیری از نظرات خبرگان، به طراحی مدل شایستگی برای کتابدار داده مبادرت شود.

تشکر و قدردانی

این مطالعه، حاصل طرح تحقیقاتی برگرفته از پایان‌نامه با عنوان «ارایه مدل کتابدار داده برای مدیریت داده‌های پژوهشی» است که با حمایت دانشکده پیراپزشکی دانشگاه علوم پزشکی شهیدبهشتی با کد طرح ۴۳۰۰۹۹۲۱ و مصوبه اخلاق IR.SBMU.RETECH.REC.1403.083 صورت گرفته است. نویسندگان مقاله مراتب سپاس و قدردانی خود را به سازمان و تمام همکارانی که در این مسیر راه‌گشا بوده‌اند، تقدیم می‌کنند.

References

- Masinde J, Chen J, Wambiri D & Mumo A. Research librarians' experiences of research data management activities at an academic library in a developing country. *Data and Information Management* 2021; 5(4): 412-24.
- Soleymani N, Mansouri A, Shabani A & Siadat SA. Components of research data management (RDM) in academic libraries. *Librarianship and Information Organization Studies* 2022; 33(1): 20-39 [Article in Persian].
- Tripathi M, Shukla A & Sonkar SK. Research data management practices in university libraries: A study. *DESIDOC Journal of Libraru and Information Technology* 2017; 37(6): 417-24.
- Hickson S, Poulton K, Connor M, Richardson J & Wolski M. Modifying researchers' data management practices: A behavioural framework for library practitioners. *International Federation of Library Associations and institutions Journal* 2016; 42(4): 253-65.
- Pinfield S, Cox AM & Smith J. Research data management and libraries: Relationships, activities, drivers and influences. *PLoS One* 2014; 9(12): e114734.
- Tang R & Hu Z. Providing research data management (RDM) services in libraries: Preparedness, roles, challenges, and training for RDM practice. *Data and Information Management* 2019; 3(2): 84-101.

7. Harvard Medical School. Harvard biomedical data management | Data management. Available at: <https://datamanagement.hms.harvard.edu/>. 2023.
8. Sinaeepourfard A, Masip-Bruin X, Garcia J & Marin-Tordera E. A survey on data lifecycle models: Discussions toward the 6Vs challenges. Available at: <https://www.ac.upc.edu/app/research-reports/html/RR/2015/18.pdf>. 2015.
9. Davidson J, Jones S, Molloy L & Kejsler UB. Emerging good practice in managing research data and research information within UK universities. *Procedia Computer Science* 2014; 33(1): 215-22.
10. Ritchie A & Kuusniemi ME. Developing research data management practices at a university hospital library in Australia-an international collaborative project. Seville, Spain: 15th EAHIL Conference Knowledge, Research, Innovation ... e-Health, 2016.
11. Read KB, Koos J, Miller RS, Miller CF, Phillips GA, Scheinfeld L, et al. A model for initiating research data management services at academic libraries. *Journal of the Medical Library Association* 2019; 107(3): 432-41.
12. Federer L. Defining data librarianship: A survey of competencies, skills, and training. *Journal of the Medical Library Association* 2018; 106(3): 294-303.
13. Ohaji IK, Chawner B & Yoong P. The role of a data librarian in academic and research libraries. *Information Research* 2019; 24(4): 844.
14. Semeler AR & Pinto AL. Data librarianship as a field study. *Transinformacao* 2020; 32(1): 1-12.
15. Semeler AR, Pinto AL & Rozados HBF. Data science in data librarianship: Core competencies of a data librarian. *Journal of Librarianship and Information Science* 2019; 51(3): 771-80.
16. Eric J. Working as a data librarian: A practical guide. California: Libraries Unlimited; 2018: 14-23.
17. Calarco P, Shearer K, Schmidt B & Tate D. Librarians' competencies profile for scholarly communication and open access. Available at: https://web.archive.org/web/20180413235412id_/https://www.coar-repositories.org/files/Competencies-for-ScholComm-and-OA_June-2016.pdf. 2016.
18. Silkotch Ch, Garcia-Milian R & Hersey D. Partnering with health sciences libraries to address challenges in bioimaging data management and sharing. *Histochemistry and Cell Biology* 2023; 160(3): 193-8.
19. Sinha P, Subaveerapandiyan A & Sinha MK. Research data management services in academic libraries: A comparative study of South Asia and Southeast Asia. Available at: https://www.researchgate.net/publication/370600835_Research_data_management_services_in_academic_libraries_a_comparative_study_of_South_Asia_and_Southeast_Asia. 2023.
20. Fuhr J. Developing data services skills in academic libraries. *College and Research Libries* 2022; 83(3): 474-502.
21. Xiao S, Ng TYF & Yang T. Research data stewardship at the university of Hong-Kong. *Library Management* 2022; 43(1/2): 128-47.
22. Amanullah S & Abrizah A. The landscape of research data management services in Malaysian academic libraries: Librarians' practices and roles. *The Electronic Library* 2023; 41(1): 63-86.
23. Ghahnaviyeh H, Rasouli-Azad MR, Babalhavaeji F & Hariri N. The most widely used components of research data management (case study: Librarians of university of medical sciences libraries in 7 th. zone in Iran). *Journal of Knowledge Studies* 2022; 14(55): 107-20[Article in Persian].
24. Bishop BW, Collier H, Orehek AM & Ihli M. Potential roles for science librarians in research data management: A gap analysis. *Issues in Science and Technololgy Librarianship* 2021; 98(1): 1-17.
25. Kim S & Syn SY. Practical considerations for a library's research data management services: The case of the national institutes of health library. *Journal of the Medical Library Association* 2021; 109(3): 450-8.



26. Lafferty-Hess S, Rudder J, Downey M, Ivey S, Darragh J & Kati R. Conceptualizing data curation activities within two academic libraries. *Journal of Librarianship and Scholarly Communication* 2020; 8(1): 1-17.
27. Nwabugwu MJ & Godwin LS. Research data management (RDM) services in libraries: Lessons for academic libraries in Nigeria. *Library Philosophy and Practice (e-journal)* 2020; 4238(1): 1-18.
28. Federer L, Foster ED, Glusker A, Henderson M, Read K & Zhao S. The medical library association data services competency: A framework for data science and open science skills development. *Journal of the Medical Library Association* 2020; 108(2): 304-9.
29. Hamad F, Al-Fadel M & Al-Soub A. Awareness of research data management services at academic libraries in Jordan: Roles, responsibilities and challenges. *New Review of Academic Librarianship* 2021; 27(1): 76-96.
30. Chukanova S. Research data management and digital curation as a library activity. *Folia Bibliologica* 2020; 61(1): 85-98.
31. Ogier A, Brown A, Petters J, Hilal A & Porter N. Enhancing collaboration across the research ecosystem: Using libraries as Hubs for discipline-specific data experts. *Practice and Experience in Advanced Research Computing* 2018; 60(1): 1-6.
32. Faniel IM & Connaway LS. Librarians' perspectives on the factors influencing research data management programs. *College and Research Libraries* 2018; 79(1): 100-19.
33. Boateng KA. Academic library and research data management roles the case of Norwegian libraries [Thesis]. Norway: Oslo and Akershus University College of Applied Sciences; 2015.
34. Sheikh A, Malik A & Adnan R. Evolution of research data management in academic libraries: A review of the literature. *Information Development* 2023; OnlineFirst: 1-15.
35. Almeida FG & Cendon BV. Library support for research data management: A taxonomy of services. *Qualitative and Quantitative Methods in Libraries* 2020; 9(2): 109-26.
36. Wong GKW & Chan DLH. Designing library-based research data management services from bottom-up. In book: *Future directions in digital information*. UK: Chandos Publishing; 2021: 55-68.
37. Creamer A, Kafel D & Martin ER. Research data management and the health sciences librarian. In book: *Health sciences librarianship chapter: Research data management and the health sciences librarian*. USA: Rowman and Littlefield and Medical Library Association; 2014: 252-74.
38. Osareh F, Sam S & Ghayebi A. Identifying the dimensions of the development of research data management services in scientific centers. *Library and Information Science* 2021; 24(2): 85-111 [Article in Persian].
39. Oroji M, Hariri N & Babalhavaeji F. Systematic review of researches in the field of research data of management literacy. *Librarianship and Information Organization Studies* 2023; 34(4): 133-66 [Article in Persian].
40. Koltay T. Accepted and emerging roles of academic libraries in supporting research. *The Journal of Academic Librarianship* 2019; 45(2): 75-80.
41. Ashiq M & Warraich NF. A systematized review on data librarianship literature: Current services, challenges, skills, and motivational factors. *Journal of Librarianship and Information Science* 2022; 55(3): 1-20.
42. Peters W. Library services and research data support at king's college London – profiling the research data management librarian. Available at: https://library.ifla.org/id/eprint/1729/1/S06-wayne_peters-en.pdf. 2017.
43. Majidi A, Naghshineh N, Esmaili Givi M & Hashemi M. Study of the foundation, models and issues of research data curation and management in scientific and academic environments. *Human Information Interaction* 2017; 4(2): 31-57 [Article in Persian].

44. Asnafi AR & Abdi S. The role of the data librarian in scientometrics and related fields. *Journal of Knowledge Studies* 2021; 13(51): 35-51[Article in Persian].
45. Oroji M, Alimohammadi D & Akhshik SS. Capacity assessment of curriculum information science and knowledge studies to support students for getting job opportunities in research data management. *National Library and Information Organization Studies* 2019; 30(1): 58-83[Article in Persian].
46. Harvard Medical School. Scholarly products | Data management. Available at: <https://datamanagement.hms.harvard.edu/share-publish/scholarly-products>. 2024.
47. Harvard Medical School. Preprints and publishing | Data management. Available at: <https://datamanagement.hms.harvard.edu/share-publish/preprints-publishing>. 2024.

Expected Competencies and Capabilities of Data Librarians in Research Data Management According to the Harvard Biomedical Data Life Cycle: A Scoping Review

Zohreh Ehteshami¹ (B.S.), Azam Shahbodaghi^{2*} (Ph.D.), Mohammad Javad Mansourzadeh³ (Ph.D.)

¹ Master of Sciences Student in Medical Library and Information Science, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

² Assistant Professor, Department of Medical Library and Information Science, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

³ Ph.D. in Medical Library and Information Science, Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Abstract

Received: 17 Aug. 2024

Accepted: 2 Dec. 2024

Background and Aim: An efficient data librarian equipped with the necessary competencies and capabilities is one of the most crucial elements in managing research data. The aim of this study is to identify the expected competencies and capabilities for data librarians in research data management according to the Harvard Biomedical Data Life Cycle.

Materials and Methods: This study is a scoping review, utilizing the Harvard Biomedical Data Lifecycle model to systematically present the findings. To retrieve relevant literature, a search strategy was employed using related keywords in databases such as Scopus, PubMed, Web of Science, Google Scholar and other reputable domestic databases, over the past five years. The research population comprised original research articles published in Persian and English that addressed the expected skills and capabilities for data librarians in managing research data.

Results: Out of 5064 documents found, 196 were selected for full-text review. After reviewing the full texts, 17 studies were included in the research. In total, 92 competencies and capabilities were identified across 23 processes within the 7 stages of the Harvard Biomedical Data Lifecycle: 16 in the first stage, 16 in the second stage, 7 in the third stage, 15 in the fourth and fifth stages, 12 in the sixth stage, 8 in the seventh stage, and 18 general competencies and capabilities. According to the findings, the most studies focused on the competencies and capabilities required for the second stage, "Collection and Creation," while the fewest studies addressed the seventh stage, "Publish and Reuse." No studies mentioned competencies and capabilities for the processes "Image Management" in the third stage and "Preprints and Publishing" in the seventh stage.

Conclusion: The results of this study indicate that among the various stages of the data lifecycle, the "Collection and Creation" stage received the most attention. Additionally, data librarians should possess not only specialized and professional skills but also general competencies and capabilities. It is recommended that the findings of this research be considered for designing short-term and long-term educational programs to train data librarians for research data management.

Keywords: Research Data Management, Data Librarian, Harvard Biomedical Data Lifecycle

* Corresponding Author:
Shahbodaghi A
Email:
shahbodaghi@sbmu.ac.ir