به نام خدا

وزارت علوم، تحقیقات، و فناوری

پژوهشگاه علوم و فناوری اطلاعات ایران

(ایرانداک)

خلاصهٔ گزارش طرح پژوهشی سازمانی

عنوان طرح پژوهشی

شناسایی چارچوبی برای مدیریت و به‌اشتراک‌گذاری داده‌های پژوهشی در مؤسسه‌های آموزشی و پژوهشی در ایران: مطالعه موردی پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک) و دانشگاه تربیت مدرس

تاریخ شروع طرح پژوهشی	**۱۴۰۳/۰۸/۱۴**	**تاریخ پایان طرح پژوهشی**	**۱۴۰۵/۰۷۰۷**
نام و نام خانوادگی مجری طرح پژوهشی	**رضا رجبعلی بگلو**		
نام و نام خانوادگی همکاران طرح پژوهشی	**اسماعیل وزیری و وحید مجیدی کوهبنانی**		

دربارهٔ طرح پژوهشی

داده‌های پژوهشی هسته فرایند پژوهش‌اند؛ گردآوری آن‌ها زمان و هزینه فراوان می‌طلبد و به‌اشتراک‌گذاری‌شان به شفافیت، بازتولیدپذیری و بازاستفاده از سرمایه‌گذاری عمومی کمک می‌کند. با این حال، نتایج پژوهش‌های پیشین نشان می‌دهد سیاست سازمان‌ها در این زمینه ناکافی بوده و نگرانی‌هایی مانند سوءتفسیر و سوءاستفاده از داده‌ها را برای جامعه علمی پدید آورده است. ایرانداک نیز با وجود نقش ملی در ثبت و اشاعه اطلاعات پژوهشی، برنامه‌ای منتشرشده برای مدیریت و به‌اشتراک‌گذاری داده‌های طرح‌های خود ندارد. مسئله طرح این است که استقرار مدیریت داده پژوهشی در یک مؤسسه ایرانی به چه مؤلفه‌هایی نیاز دارد و این مؤلفه‌ها چه نسبتی با یکدیگر دارند.

هدف اصلی طرح، شناسایی نیازمندی‌های اجرایی، مدیریتی، سیاستی، فرهنگی، اخلاقی و حقوقی مدیریت و به‌اشتراک‌گذاری داده‌های پژوهشی در مؤسسه‌های آموزشی و پژوهشی و صورت‌بندی آن‌ها در قالب یک چارچوب مفهومی بود. هدف‌های فرعی عبارت بودند از شناسایی مزایا، چالش‌ها و نیازمندی‌های پیاده‌سازی موفق، و شناسایی اجزایی که برای تدوین راهنما، خط‌مشی و برنامه مدیریت داده لازم است.

آیین‌نامه‌ها، شیوه‌نامه‌ها و خط‌مشی‌های ایرانداک و ۲۲۳ طرح پژوهشی ده سال گذشته آن بررسی شد؛ ۱۱۳

سند مدیریت داده از ۶۰ مؤسسه پیشرو در جهان تحلیل شد؛ و دیدگاه شش صاحب‌نظر در یک نشست مصاحبه گروهی گردآوری شد. یافته‌ها در چارچوبی مفهومی و اولیه، شامل ده مؤلفه، چهار لایه و یک حلقه عرضی سامان یافت و با سطوح بلوغ، نقشه راه استقرار و شاخص‌های ارزیابی تکمیل شد. چارچوب هنوز اعتبارسنجی نشده است. مقایسه با دانشگاه تربیت مدرس نیز به‌دلیل در دسترس نبودن اسناد قابل تحلیل آن انجام نشد.

روش پژوهش

پژوهش از نظر هدف کاربردی است و با رویکرد کیفی و راهبرد مطالعه موردی، در سه گام مکمل اجرا شد. در گام نخست، مبانی نظری و پیشینه علم آزاد و مدیریت داده پژوهشی مرور شد و سیاهه‌ای که بر پایه آن‌ها طراحی شده بود، برای بررسی آیین‌نامه‌ها، شیوه‌نامه‌ها و خط‌مشی‌های ایرانداک و ۲۲۳ طرح پژوهشی ده سال گذشته آن به کار رفت؛ هر طرح از نظر سه مؤلفه داده آزاد، متن‌باز و روش‌شناسی آزاد سنجیده شد. همچنین وب‌گاه ۱۰۰ مؤسسه نخست رتبه‌بندی شانگهای کاویده شد و از ۶۰ مؤسسه، ۱۱۳ سند مدیریت داده گردآوری و بر پایه یک چارچوب تحلیلی از نظر حضور هر مؤلفه کدگذاری شد.

در گام دوم، با تکیه بر خلأها و ابهام‌های شناسایی‌شده در اسناد، پرسش‌های مصاحبه گروهی تدوین شد. در گام سوم، یک نشست مصاحبه گروهی متمرکز با شش صاحب‌نظر برگزار شد که با نمونه‌گیری هدفمند و با سه معیار برگزیده شدند: سابقه عملی در چرخه طرح پژوهشی، پوشش هم‌زمان ابعاد موضوع، و تنوع جایگاه علمی. داده‌های مصاحبه پیاده و به روش تحلیل محتوای استقرایی ـ قیاسی در نرم‌افزار تحلیل کیفی کدگذاری شد. اعتمادپذیری با زاویه‌بندی پژوهشگر و داده، بازبینی همتراز کدها و ثبت مسیر تصمیم‌های کدگذاری پشتیبانی شد. نتایج سه منبع با هم تلفیق شد و هیچ مؤلفه‌ای تنها بر پایه یک منبع وارد چارچوب نشد.

دو ابزار هوش مصنوعی مولد فقط در نقش کمکی به کار رفتند: دیپ‌سیک برای گسترش واژگان جست‌وجو و شناسایی اولیه منابع پیشینه، و چت‌جی‌پی‌تی نسخه ۵٫۶ برای سازمان‌دهی پیکره اسناد و پیشنهاد چارچوب اولیه کدگذاری. تنها اسناد عمومی در اختیار ابزار قرار گرفت و متن مصاحبه، اطلاعات هویتی یا داده محرمانه وارد نشد. همه منابع و کدها را پژوهشگران با سند مبدأ تطبیق دادند و مسئولیت تفسیر نتایج با آنان بود.

یافته‌های طرح پژوهشی

مقررات ایرانداک بیش از آنکه چرخه کامل داده را پوشش دهند، بر گزارش طرح، ثبت اثر، دسترسی به متن و صیانت از حقوق مؤلف متمرکزند و انتشار داده، کد و مستندات روش را الزامی نکرده‌اند؛ وضعیت ایرانداک را می‌توان گذار از مدیریت برونداد به مدیریت چرخه داده دانست. از ۲۲۳ طرح بررسی‌شده، ۵۹٫۲ درصد

روش‌شناسی آزاد، ۵۳,۴ درصد داده آزاد و تنها ۲۴,۲ درصد متن‌باز داشتند؛ ۲۱,۱ درصد هر سه مؤلفه را هم‌زمان دارا بودند و ۲۷,۴ درصد هیچ‌یک را نداشتند. تحلیل اسناد ۶۰ مؤسسه پیشرو ده مؤلفه را آشکار کرد. خط‌مشی و راهبرد با ۹۱,۷ درصد پرتکرارترین مؤلفه بود و پس از آن هزینه و پایداری و آموزش (هر یک ۷۶,۷ درصد)، نگهداشت بلندمدت (۷۵ درصد) و واسپارگاه (۷۳,۳ درصد) قرار گرفتند. در برابر، برنامه مدیریت داده در ۶۶,۷ درصد، پایش انطباق در ۵۰ درصد و اصول فِیر و شناسه پایدار هر یک در ۴۰ درصد این مؤسسه‌ها دیده شد. فاصله خط‌مشی (۹۱,۷ درصد) با پایش (۵۰ درصد) نشان می‌دهد سیاست‌ها بیش از آنکه سنجیده شوند، نوشته می‌شوند. مصاحبه گروهی نشان داد دغدغه اصلی کنشگران ایرانی پیش از آنکه فنی باشد، حقوقی و انگیزشی است: ابهام مالکیت مادی و معنوی داده، نبود قرارداد روشن، ترس از سوءاستفاده یا از دست رفتن فرصت انتشار، فرهنگ مالکیت‌محور، نبود بستر مناسب و نبود مشوق ساختاری. در برابر، شفافیت و اعتمادپذیری پژوهش، پرهیز از دوباره‌کاری، دیده‌شدن آثار و نظام‌های تشویقی به‌عنوان پیش‌برنده بیان شد.

از تلفیق سه منبع، ده مؤلفه شناسایی و در چهار لایه صورت‌بندی شد: حکمرانی و جهت‌دهی، برنامه‌ریزی چرخه عمر، زیرساخت و استاندارد، و دسترسی مسئولانه و ظرفیت انسانی. مؤلفه هزینه، پایش و ارزیابی به‌شکل حلقه‌ای عرضی بازخورد را به لایه نخست بازمی‌گرداند. برای هر لایه درون‌داد، فرایند و برون‌داد تعریف شد. چهار سطح بلوغ (آغازین، در حال توسعه، تثبیت‌شده و یکپارچه)، نقشه راه سه‌مرحله‌ای استقرار همراه با مسئولیت بازیگران و شش شاخص ارزیابی (پوشش برنامه، شناسه‌دهی، شفافیت محدودیت، زمان پاسخ به درخواست دسترسی، بازاستفاده قابل اثبات و انطباق) نیز به چارچوب افزوده شد.

نتیجه‌گیری و پیشنهادها

مدیریت داده پژوهشی پدیده‌ای چندلایه است و با یک اقدام منفرد، چه راه‌اندازی واسپارگاه و چه تصویب خط‌مشی، محقق نمی‌شود. سیاست تنها هنگامی اثر دارد که در نقاط تصمیم چرخه طرح تعبیه شود و برنامه مدیریت داده آن را به عمل روزمره پژوهش برساند. گشودگی و حفاظت در برابر یکدیگر نیستند و قاعده غالب «دسترسی مسئولانه» است. در بافت ایران، مالکیت، قرارداد و اعتماد سازمانی در کنار زیرساخت و مهارت، متغیرهای تعیین‌کننده پذیرش‌اند. ایرانداک به دلیل تجربه ملی در ثبت و اشاعه اطلاعات پژوهشی می‌تواند نقش هماهنگ‌کننده استاندارد، شناسه، راهنما و خدمت مشترک را برعهده گیرد، مشروط بر آنکه چارچوب ملی امکان بومی‌سازی مؤسسه‌ای را حفظ کند. چارچوب ارائه‌شده مفهومی و اولیه است و از تلفیق ادبیات، تحلیل اسناد و یک مصاحبه گروهی برآمده؛ داوری مستقل خبرگان آن را نیازموده است. شواهد میدانی به یک پژوهشگاه و شش صاحب‌نظر محدود است، پس نتایج تعمیم تحلیلی‌اند، نه آماری. پیشنهادها: ۱) اعتبارسنجی چارچوب به روش دلفی در دو یا سه دور با پانلی از دست‌کم پانزده تا بیست خبره از میان سیاست‌گذاران، مدیران پژوهشی و پژوهشگران

رشته‌های گوناگون؛ ۲) اجرای آزمایشی یک‌ساله در چند مؤسسه با اندازه و مأموریت متفاوت؛ ۳) آغاز استقرار از «حداقل قابل اجرا»: خط‌مشی پایه، مسئول مشخص، برنامه مدیریت داده متناسب، ذخیره امن، مستندسازی، تصمیم دسترسی و سپرده در واسپارگاه معتبر؛ ۴) بازنگری قراردادهای پژوهشی و آیین‌نامه ارتقا برای روشن‌کردن مالکیت داده و به رسمیت شناختن آن به‌عنوان برونداد علمی.