

# Reuse of Research Data in Iran: Elements and Requirements

## Ameneh Soleimani

PhD Candidate in Knowledge and Information Science; University of Tehran; Tehran, Iran Email: amenehsoleimani94@gmail.com

## Fatima Fahimnia\*

PhD in Information Science and Knowledge Studies; Associate Professor; Department of Information Science and Knowledge Studies; University of Tehran; Tehran, Iran Email: fahimnia@ut.ac.ir

## Nader Naghshineh

PhD in Information Sciences and Knowledge Studies; Assistant Professor of Information Sciences and Knowledge Studies; University of Tehran; Tehran, Iran Email: nnaghshineh@gmail.com

## Adel Soleimani Nezhad

PhD in Knowledge and Information Sciences; Assistant Professor; Department of Knowledge and Information Science; Shahid Bahonar University of Kerman; Kerman, Iran; Email: adelss2004@yahoo.com

Received: 16, Jan. 2018 Accepted: 13, May 2018

**Abstract:** This paper attempts to identify the elements and requirements for reuse of research data among researchers in Iran.

The present study is qualitative research. This research was carried out in two stages. At first, Initial elements were extracted by reviewing and analyzing the content of the texts, but due to the fact that the extracted factors were not native, we used semi-structured interview. The statistical population in the first stage includes articles about reuse in PubMed, Scopus, Science direct and ProQuest databases. The statistical population of the second phase was researchers who were familiar with the reuse of research data. The interviews began using targeted sampling and continued using snowball strategy and sampling based on the introduction of the contributors. The data gathering tool was a semi-structured interview and data were analyzed by MaxQDA software.

As a result of analyzing the content of the literature and interviews, the main elements of reuse of research data were identified in three main dimensions: human factors, organizational factors and infrastructures. The components of each dimension were identified and, at the end, suggestions were made to reuse existing data.

**Keywords:** Data Reuse, Research Data, Research Data Reuse, Repositories

\* Corresponding Author

Iranian Journal of  
**Information  
Processing and  
Management**

Iranian Research Institute  
for Information Science and Technology  
(IranDoc)

ISSN 2251-8223

eISSN 2251-8231

Indexed by SCOPUS, ISC, & LISTA

Vol. 35 | No. 3 | pp. 663-692

Spring 2020

<https://doi.org/10.35050/JIPM010.2020.035>



# استفاده مجدد از داده‌های پژوهشی

## در ایران: شناسایی عناصر و الزامات

آمنه سلیمانی ده دیوان

دانشجوی دکتری علم اطلاعات و دانش‌شناسی؛  
دانشگاه تهران؛ تهران، ایران؛  
amenehsoleimani94@gmail.com

فاطمه فهیم‌نیا

دکتری علم اطلاعات و دانش‌شناسی؛ دانشیار؛  
گروه علم اطلاعات و دانش‌شناسی؛ دانشگاه تهران؛  
تهران، ایران؛  
fahimnia@ut.ac.ir

نادر نقشبند

دکتری علم اطلاعات و دانش‌شناسی؛ استادیار؛  
گروه علم اطلاعات و دانش‌شناسی؛ دانشگاه تهران؛  
تهران، ایران naghshineh@gmail.com

عادل سلیمانی نژاد

دکتری علم اطلاعات و دانش‌شناسی؛ استادیار؛  
گروه علم اطلاعات و دانش‌شناسی؛ کرمان، ایران؛  
adels2004@yahoo.com



مقاله برای اصلاح به مدت ۲۵ روز نزد پدیدآوران بوده است.

پدیش: ۱۳۹۸/۰۳/۲۰

دریافت: ۱۳۹۷/۰۸/۱۸

نشریه علمی | رتبه بین‌المللی  
پژوهشگاه علوم و فناوری اطلاعات ایران  
(ایرانداک)

شاپا (چاپی) ۲۲۳-۲۵۱

شاپا (الکترونیکی) ۸۳۱-۲۲۵

نمایه در SCOPUS، ISC، LISTA، و

jipm.irandoc.ac.ir

دوره ۳۵ | شماره ۳ | صص ۶۶۳-۶۹۲

بهار ۱۳۹۹

<https://doi.org/10.35050/JIPM010.2020.035>



**چکیده:** محققان در این پژوهش به دنبال شناسایی عناصر و الزامات مورد نیاز برای استفاده مجدد از داده‌های پژوهشی در بین پژوهشگران ایرانی هستند. پژوهش حاضر از نوع کیفی است و در دو مرحله انجام گرفت. ابتدا با مرور و تحلیل محتوای متون، عناصر اولیه استخراج شد، اما با توجه به بومی نبودن عوامل مستخرج، به وسیله مصاحبه نیمه‌ساختار یافته از نظر خبرگان استفاده شد. جامعه آماری در مرحله اول، شامل پژوهش‌های انجام‌شده در حوزه استفاده مجدد در پایگاه‌های اطلاعاتی معتبر است. جامعه آماری مرحله دوم پژوهشگرانی بودند که با بحث استفاده مجدد از داده‌های پژوهشی آشنایی داشتند. مصاحبه‌ها با استفاده از نمونه‌گیری هدفمند شروع شد، سپس با استفاده از استراتژی گلوله‌برفی و نمونه‌گیری مبتنی بر معرفی مشارکت‌کننده ادامه پیدا کرد. ابزار گردآوری داده‌ها در مرحله اول، متون و منابع و در مرحله دوم، مصاحبه نیمه‌ساختار یافته است، و در نهایت، داده‌ها با نرم‌افزار «مکس کیودا» تجزیه و تحلیل شد. کدها و مقوله‌هایی در نتیجه تحلیل محتوای متون و مصاحبه با خبرگان شناسایی و در نهایت، در ذیل سه بُعد اصلی عوامل انسانی، عوامل سازمانی، و زیرساخت‌ها دسته‌بندی شدند. سپس، مؤلفه‌های هر بُعد مشخص و در

پایان، پیشنهادهای برای به‌کارگیری استفاده مجدد از داده‌های پژوهشی موجود در ایران مطرح شد.

کلیدواژه‌ها: استفاده مجدد، داده‌های پژوهشی، مدیریت داده‌های پژوهشی، مخازن داده

## ۱. مقدمه

در دهه گذشته، تولید داده‌های دیجیتال در بسیاری از زمینه‌ها افزایش یافته است. برخی آمارها نشان می‌دهد که هر روز در جهان حدود ۲/۵ کوین تریلیون<sup>۱</sup> (معادل ۱۰ به توان ۱۸) بایت داده تولید می‌شود (Dobre and Xhafa 2014) و تا سال ۲۰۲۰ بیش از ۴۰ زتابایت (یا ۴۰ تریلیون گیگابایت) داده تولید و مصرف خواهد شد (Gantz and Reinsel 2012). این حجم عظیم داده ما را با پدیده‌ای تحت عنوان سیل عظیم داده<sup>۲</sup> مواجه می‌سازد و حجم داده‌ها و نقش آن‌ها در پژوهش باعث ایجاد الگویی جدید از علم تحت عنوان پارادایم چهارم<sup>۳</sup> شده است که توسط «جیم گری»<sup>۴</sup> در دهه ۱۹۹۰ مطرح شد. هدف از این پارادایم، جهانی است که در آن همه متون پژوهشی، آثار علمی و داده‌های مربوط به آن‌ها به‌صورت برخط در دسترس هستند. در این پارادایم، تنها بر پژوهش‌هایی تأکید می‌شود که مبتنی بر داده‌های به‌اشتراک گذاشته شده هستند و پژوهشگران، دیگر به دنبال تولید و جمع‌آوری داده نیستند و از داده‌های تولیدشده توسط سایرین استفاده می‌کنند (وزیری، نقشینه و نوروزی ۱۳۹۶). به این ترتیب، دسترس پذیری داده‌ها به‌طور روزافزون مورد توجه قرار گرفت. در تلاش برای گسترش دسترسی به داده‌های پژوهشی و مطابق با دستورالعمل‌های جدید دولتی در این باره، برخی از ناشران، مجلات، مؤسسات دانشگاهی و پژوهشی و حامیان مالی پژوهش‌ها شروع به تأسیس و یا گسترش مخازن داده<sup>۵</sup> و دعوت به اشتراک‌گذاری<sup>۶</sup> داده‌ها نمودند (Curty 2016). با ایجاد مخازن داده، روزبه‌روز داده‌های بیشتری در سرورهای این مخازن انباشته شد و بنابراین، توجه‌ها به سمت پیدا کردن راه‌هایی برای حفظ ارزش و به حداکثر رساندن استفاده از داده‌ها در داخل و حتی بین رشته‌ها معطوف گردید (Faniel

1. quintillion

2. data deluge

3. fourth paradigm

4. Jim Gray

5. data repositories

6. data sharing

and Zimmerman 2011) و بحث استفاده مجدد<sup>۱</sup> از داده‌های پژوهشی به صورت جدی مطرح شد. استفاده مجدد از داده‌های پژوهشی به معنای استفاده از یک مجموعه داده، توسط شخص دیگری غیر از گردآورندگان اصلی داده‌ها برای اهداف جدید است. به عبارت دیگر، استفاده مجدد از داده‌های پژوهشی یعنی محقق از داده‌هایی که شخصاً جمع‌آوری نکند یا از ترکیب این داده‌ها با هم، برای پاسخ دادن به یک سؤال جدید استفاده کند (Zimmerman 2008; Pasquetto, Randles and Borgman 2017). بین سال‌های ۲۰۰۷ تا ۲۰۱۰ دانشمندان ژنومیک ۱۱۵۰ مقاله منتشر کردند که داده‌های آن‌ها را خودشان جمع‌آوری نکرده بودند (Fear 2013). در سال ۲۰۱۰، و بعد از سال‌ها برنامه‌ریزی، اشتراک‌گذاری و استفاده مجدد از داده‌های دیگران گروهی از محققان موفق به تشخیص زود هنگام آلزایمر شدند (Mueller et al. 2005). در سال ۲۰۱۳، یک دانشجوی اقتصاد توانست با استفاده مجدد از داده‌هایی که در تحقیقات سال ۲۰۱۰، جمع‌آوری و منتشر شده بود، خطایی را شناسایی کند که تأثیرات مهمی در سیاست‌های عمومی جامعه داشت (Herndon, Ash and Pollin 2013). اینها نمونه‌هایی از روند روبه‌رشد استفاده مجدد از داده‌های پژوهشی در دهه گذشته است. عوامل متعددی این رشد را توضیح می‌دهند؛ از جمله: جنبش داده‌های باز، سیاست‌های ناشران و حامیان مالی پژوهش‌ها در حمایت از اشتراک‌گذاری داده‌ها<sup>۲</sup>، دیدگاه پژوهشگران نسبت به مزایای استفاده از منابع به اشتراک‌گذاری شده و خدمات و زیرساخت‌هایی<sup>۳</sup> که دسترسی به مجموعه داده‌ها را بهبود می‌بخشند (Bishop and Kuula-Lummi 2017). از آنجا که ارزش واقعی داده در کاربرد و نه حجم آن است (Coffey 2016)، داده به عنوان یک دارایی محسوب شده که استفاده مجدد از آن منجر به بازگشت سرمایه و ایجاد ارزش افزوده خواهد شد (El Hani, Rivest and Maranzana 2012). به این ترتیب، استفاده مجدد از داده‌های موجود چند سالی که است

#### 1. data Re-use

۲. به طور مثال، بنگاه‌های مالی علمی مانند مؤسسه ملی سلامت، مؤسسه ملی سرطان، شورای پژوهش‌های انگلستان برای پژوهش‌هایی که با بودجه آنها و یا بودجه عمومی انجام شده باشند، درخواست اشتراک‌گذاری داده‌های پژوهشی و طرح مدیریت داده را دارند (نقل در وزیری، نقشینه و نوروزی ۱۳۹۶).

۳. مخازنی مانند مندلی دیتا، زندو، ژنبنک، درآی‌آد، دیتاشر و سرویس داده انگلستان از جمله زیرساخت‌هایی هستند که امکان دسترسی به مجموعه داده‌ها را فراهم می‌کنند.

به‌عنوان یک بحث پژوهشی مستقل، و نه نتیجه اشتراک‌گذاری داده‌ها مورد توجه قرار گرفته است. به‌رغم اقبال از بحث استفاده مجدد از داده‌ها در دنیا، در ایران این رویکرد حداقل به‌صورت رسمی انجام نمی‌شود. بر اساس آخرین اطلاعات مستخرج از پایگاه استنادی «کلاریویت آنالیتیکس»<sup>۱</sup> ایران در سال ۲۰۱۷، بالاترین نرخ رشد تولید علم در میان ۲۵ کشور برتر دنیا را داشته است (پایگاه استنادی جهان اسلام ۱۳۹۶). این حجم بالای تولیدات علمی طبیعتاً منجر به تولید داده‌های پژوهشی زیادی در کشور خواهد شد که می‌توان استفاده مجدد از آن‌ها را در دستور کار قرار داد و این مفهوم را به یک فرهنگ و رویکرد متداول در جامعه علمی کشور تبدیل کرد. بدین‌منظور در راستای پذیرش و استقرار نظامی به‌منظور استفاده مجدد از داده‌های پژوهشی در ایران، در ابتدای راه نیازمند شناسایی عناصر و الزامات این حوزه هستیم تا با شناخت ابعاد به‌بومی‌سازی آن‌ها پردازیم. در همین راستا، در این پژوهش سعی بر این است که ابعاد به‌کارگیری استفاده مجدد از داده‌های پژوهشی در ایران شناسایی شوند. پژوهش حاضر در صدد پاسخگویی به این سؤالات است: ابعاد و مؤلفه‌های استفاده مجدد از داده‌های پژوهشی بر اساس تحلیل محتوای متون کدام‌اند؟ ابعاد و مؤلفه‌های استفاده مجدد از داده‌های پژوهشی در ایران از نظر خبرگان کدام‌اند؟

## ۲. پیشینه پژوهش

مروری بر متون و مقالات نشان می‌دهد که با وجود گذشت بیش از دو دهه از مطرح شدن بحث داده‌های پژوهشی در سطح بین‌المللی و علی‌رغم اهمیت آن در دنیا، پژوهش‌های اندکی در این حوزه در ایران انجام شده است. به‌طور مثال، «وزیری، نقشینه و نوروزی» در تحقیقی با عنوان «به‌اشتراک‌گذاری داده‌های پژوهشی»، به این نتیجه رسیدند که علاوه بر سازمان‌های جهانی، ناشران نشریات نیز اشتراک‌گذاری داده‌ها را به‌عنوان یک الزام در نظر گرفته و شرط انتشار مقالات را اشتراک‌گذاری داده‌ها قرار داده‌اند<sup>۲</sup>.

1. Clarivate Analytics

۲. به‌طور مثال، مجلاتی از جمله «پی‌ال‌اَس» و «ای‌جی‌و» (AGU 2013; PLOS 2014). نشریات متعددی از جمله نیچر، ساینس، سیچ و پلاس سیاست‌های اشتراک داده‌های پژوهشی خود را تنظیم نموده‌اند که در آنها نویسندگان ملزم به بارگذاری داده‌های پژوهشی خود به همراه اصل مقاله هستند (وزیری، نقشینه و نوروزی ۱۳۹۶).

(۱۳۹۶). «وزیری» و همکاران در پژوهشی، عوامل مؤثر بر اشتراک‌گذاری داده‌ها را بررسی کردند و در پژوهشی دیگر به بررسی چالش‌های اشتراک‌گذاری داده‌ها پرداختند (۱۳۹۶). در این پژوهش، محققان نتوانستند در داخل کشور پژوهشی بیابند که به‌طور مستقیم به استفاده مجدد از داده‌ها پرداخته باشد، اما در خارج از کشور در زمینه استفاده مجدد از داده‌های پژوهشی، مطالعاتی به‌قرار زیر انجام شده است. «پاسکوئو، راندل و بورگمن» در پژوهشی با عنوان «استفاده مجدد از داده‌های علمی» به مسائلی از جمله به‌کارگیری استفاده مجدد در بازتولید<sup>۱</sup> پژوهش‌ها، تقابل ادغام با استفاده مجدد و انگیزه‌ها در استفاده مجدد اشاره می‌کنند (Pasquetto, Randles and Borgman 2017). «یون و کیم» در پژوهشی به بررسی باورها، گرایش‌ها و هنجارها در زمینه استفاده مجدد از داده‌ها در بین محققان علوم اجتماعی پرداختند (Yoon and Kim 2017). «بیشاپ و کوآلا-لامی» به بازبینی استفاده مجدد از داده‌های کیفی پرداختند. آن‌ها با بیان چند مطالعه موردی در زمینه استفاده مجدد، وسعت و تنوع این متدولوژی در حال رشد را نشان دادند (Bishop and Kuula-Luumi 2017). «ثانوس» در پژوهشی مفاهیم اساسی و موانع استفاده مجدد از داده‌ها را از بُعد فنی بررسی می‌کند (Thanos 2017). «کارتی» و همکاران در تحقیقی نگرش‌ها و هنجارهای مؤثر بر استفاده مجدد را بررسی و تأکید کردند که با نشان دادن ارزش استفاده مجدد باعث تشویق این رفتار خواهیم شد (Curty et al. 2017). «کارتی» در پژوهشی به بررسی عوامل مؤثر بر استفاده مجدد از داده‌ها در علوم اجتماعی پرداخته است (Curty 2015). «یون» به مطالعه استفاده مجدد از داده‌ها و قضاوت‌های اعتماد کاربران پرداخت. وی معتقد است که به علت فقدان استانداردی برای اطمینان از کیفیت داده‌ها و وجود آسیب‌های بالقوه از داده‌های با کیفیت پایین، ارزیابی داده‌ها برای استفاده مجدد به مسئله‌ای مهم بدل شده است (Yoon 2015). «فدرر» و همکاران اشتراک‌گذاری و استفاده مجدد از داده‌های پژوهشی بیومدیکال را بررسی کرده و بیان کردند که بیشتر پاسخ‌دهندگان به استفاده مجدد اهمیت می‌دهند، اما در این زمینه تخصص کمی دارند (Federer et al. 2015).

با مرور پیشینه‌ها مشخص می‌شود که پژوهش‌های انجام‌شده کمتر به

1. reproducibility

بررسی همه ابعاد استفاده مجدد از داده‌ها پرداخته‌اند. به‌طور مثال، اکثر مطالعات به بررسی نگرش محققان نسبت به استفاده مجدد (Federer et al. 2015; Yoon and Kim 2017; Curty et al. 2017; Bishop and Kuula-Luumi 2017)، اندازه‌گیری میزان استفاده مجدد (2017; Ball and Duke 2015; Park and Wolfram 2017; Fear 2013)، شیوه‌های استفاده مجدد (2017; Wallis, Rolando, Wallis, Rolando & Borgman 2013; Faniel et al. 2013;)، مسائل اخلاقی در زمینه استفاده مجدد (Grinyer 2009; Bishop 2014)، بررسی‌های تطبیقی (Sands et al. 2013; Faniel et al. 2012)، اعتماد به داده‌ها (Yoon 2015; Yakel et al. 2013)، مدل‌سازی (German 2015) و بررسی زیرساخت‌ها (Murillo 2010) پرداخته‌اند و کمتر پژوهشی به شناسایی تمامی ابعاد در این حوزه پرداخته است. مزیت این پژوهش جامعیت آن است که سعی کرده تمامی جنبه‌های استفاده مجدد را بسنجد. علاوه بر این، تاکنون در این حوزه در ایران پژوهشی انجام نشده است.

### ۳. روش پژوهش

۳-۱. نوع پژوهش: پژوهش پیش‌رو از نظر دسته‌بندی، کیفی، از نظر استراتژی از نوع پیمایشی و با استفاده از شیوه‌های مطالعه مستندات و مدارک و مصاحبه صورت گرفته است. این پژوهش در دو مرحله انجام شد. در مرحله اول، ابتدا جست‌وجوی نظام‌مند مقالات در پایگاه‌های اطلاعاتی «اسکوپوس»<sup>۱</sup> و «ساینس دایرکت»<sup>۲</sup>، «پاب‌مد»<sup>۳</sup>، «پروکوئست»<sup>۴</sup> و «امرالِد»<sup>۵</sup> با کلیدواژه‌های «data reus\*» و «data reutilization» به همراه کلیدواژه «research» در فیلد عنوان و چکیده انجام شد. شایان ذکر است که دلیل انتخاب این پایگاه‌های اطلاعاتی قدمت، اعتبار و حجم منابع اطلاعاتی نمایه‌شده در آن‌ها بود. جست‌وجو بر اساس نوع مدرک، به «مقاله» و زبان، به «انگلیسی» محدود شد. اما محدودیت زمانی برای جست‌وجو در نظر گرفته نشد (جدول ۱).

1. Scopus  
4. ProQuest

2. Science Direct  
5. Emerald

3. PubMed

### جدول ۱. راهبرد جست‌وجو و تعداد نتایج در پایگاه‌های اطلاعاتی

پایگاه	کلیدواژه	فرمول جستجو	نوع مدرک	محدودیت	نتیجه
ساینس دایرکت	"data reus" and research	"data reutilization"	مقاله	زبان	۸۳
اسکوپوس	"data reutilization"	[Title/Abstract] and research	کنفرانسی و علمی-پژوهشی	انگلیسی	۱۲۰
پاب مد	and research	[Title/Abstract]			۲۰۰
پروکوئست		[Title/ Abstract]			۱۳۳
امرالذ					۵۵
جمع					۵۹۱

در جست‌وجو از پایگاه‌های مذکور ۵۹۱ مقاله بازیابی شد. از آنجا که برخی از این پایگاه‌ها همپوشانی دارند، پس از انتقال نتایج به دست آمده به نرم‌افزار «اندنوت»<sup>۱</sup>، به حذف موارد تکراری پرداخته شد. قابل ذکر است که به دلیل این که تحلیل محتوای متون مستلزم دسترسی به متن کامل مقاله است، بنابراین، مقالاتی که به صورت تمام‌متن در دسترس نبودند از جریان تحلیل محتوا حذف شدند. پس از حذف موارد تکراری و مقالاتی که به صورت تمام‌متن در دسترس نبودند، در این مرحله مقالاتی که عنوان و چکیده آن‌ها با موضوع اصلی پژوهش مرتبط نبود نیز حذف شدند. سپس، کیفیت مقالات باقی‌مانده ارزیابی شد. جهت کنترل کیفیت مطالعات، ابزارهایی طراحی شده است. در این پژوهش از چارچوب ارزیابی کیفی که توسط پژوهشگران «مرکز ملی تحقیقات اجتماعی»<sup>۲</sup> در انگلستان توسعه یافته است (Spencer, Price & Walsh 2014) استفاده شد. سؤالات این ابزار ارزیابی با مفاهیم اصلی پژوهش‌های کیفی مرتبط است. این ابزار شامل ۱۷ سؤال بر مبنای ۴ اصل است (MacEachen et al. 2006). هنگام استفاده از این ابزار، هر مقاله توسط دو نفر از پژوهشگران بررسی و در یکی از کیفیت‌های پایین، متوسط، بالا و یا بسیار بالا دسته‌بندی شد (جدول شماره ۲، حاوی چهار سطح ارزیابی است). در مواردی که اختلاف نظر وجود داشت، مرورگران به بحث و بررسی پرداختند تا به اجماع برسند. اگر اجماع به دست نیامد، با نفر سوم مشورت شد و در پایان این

1. EndNote

2. National Centre for Social Research



مرحله، تنها مقالاتی که کیفیت بالا یا بسیار بالا داشتند وارد فرایند تحلیل محتوا شده و بقیه مطالعات حذف شدند. با توجه به معیارهای ورود به مطالعه، حذف مقالات تکراری، حذف مقالاتی که به صورت تمام متن در دسترس نبودند و بررسی کیفیت، در نهایت، تعداد ۴۳ مقاله انتخاب، با روش تحلیل محتوا بررسی و ابعاد استفاده مجدد از داده‌ها در تولیدات علمی سایر کشورها شناسایی شدند؛ اما به دلیل بومی نبودن ابعاد مستخرج از تحلیل محتوا، در مرحله دوم پژوهش به وسیله مصاحبه نیمه‌ساختاریافته با خبرگان این عناصر تعدیل و تکمیل شدند.

## جدول ۲. چهار سطح ارزیابی کیفیت مطالعات (راهنمای ارزیابی کیفیت مقالات) (Spencer, Price & Walsh 2014)

درجه کیفی توضیحات	
پایین	داده‌ها بیش از حد یکنواخت <sup>۱</sup> هستند، به دلیل تجزیه و تحلیل ناکافی یا استراتژی نمونه‌گیری، داده‌ها حقیقی نیستند و به نظر می‌رسد که نویسندگان مجموعه‌ای از ایده‌های خود را قرار داده‌اند.
متوسط	ماهیتاً توصیفی است، اما در توصیف زمینه و جزئیات ضعیف است و این منجر به ظاهر سطحی کار می‌شود.
بالا	توصیفی است، اما شامل سطح کافی از تجزیه و تحلیل است، همراه با در نظر گرفتن زمینه، ارائه یک تصویر پویاتر از شرکت کنندگان در مطالعه و محیط پیچیده‌ای که در آن کار می‌کنند.
بسیار بالا	شامل یک تمرکز نظری با توجه به در نظر گرفتن فرایندهای داخلی درگیر در ایجاد وضعیتی که توصیف شده بود (به‌عنوان مثال، پیوندهایی با ساختارهای ماکرو) و با مقادیر اکتشافی است که می‌تواند به سایر عرصه‌های تحقیق منتقل شود.

۲-۳. جامعه آماری: جامعه آماری در مرحله تحلیل محتوا، پژوهش‌های انجام شده در حوزه استفاده مجدد از داده‌ها در پایگاه‌های اطلاعاتی «اسکوپوس» و «ساینس دایرکت»، «پاب‌مد»، «پرو کوئست» و «امرالند» هستند. در مرحله دوم پژوهش، جامعه آماری، پژوهشگرانی منتخب هستند که قصد استفاده مجدد از داده‌های پژوهشی را دارند یا پژوهشگرانی که تجربه استفاده مجدد از داده‌ها را دارند (البته، این تجربه می‌تواند موفق یا ناموفق بوده باشد). برای انتخاب این پژوهشگران، ابتدا لیستی از سامانه‌هایی که امکان استفاده مجدد از داده‌ها را در سطح بین‌المللی فراهم کرده‌اند،

1. invariable

از جمله «مندلی دیتا»<sup>۱</sup>، «زندو»<sup>۲</sup>، «ژن‌بنک»<sup>۳</sup>، «دریاد»<sup>۴</sup> و «سرویس داده انگلستان»<sup>۵</sup> تهیه شد و در مرحله بعد چندتن از پژوهشگران ایرانی که در این سامانه‌ها فعال بودند انتخاب و مصاحبه با آن‌ها شروع شد. سپس، با استفاده از استراتژی گلوله‌برفی و نمونه‌گیری مبتنی بر معرفی مشارکت‌کننده<sup>۶</sup>، مصاحبه‌ها ادامه پیدا کرد. مصاحبه‌ها ضبط و بلافاصله به صورت کلمه‌به‌کلمه مکتوب شد و به نرم‌افزار «مکس کیودا»<sup>۷</sup> منتقل و کدگذاری آن انجام شد. به این ترتیب، بعد از مصاحبه نوزدهم محقق احساس کرد که به نقطه اشباع نظری و حد کفایت در گردآوری داده‌ها رسیده است.

۳-۳. ابزار گردآوری: ابزار مرحله اول پژوهش مقالاتی هستند که در پایگاه‌های مذکور نمایه شده‌اند. در مرحله دوم، ابزار گردآوری، مصاحبه نیمه‌ساختاریافته است. در جریان مصاحبه از پرسش‌های باز استفاده شد تا مصاحبه‌شوندگان نظرات واقعی خود را بیان کرده و محدود به سؤالات مصاحبه‌کننده نشوند و در عین حال، به منظور هدایت مسیر مصاحبه، پرسشنامه‌ای با سؤالات باز بر اساس عناصر مستخرج از مرحله اول، به عنوان راهنمای مصاحبه طراحی شد.

۴-۳. روایی و پایایی: در پژوهش‌های کیفی به سه نوع روایی اشاره شده است: توصیفی، تفسیری، و نظری (رنجبری، نقشینه و اسمعیلی ۱۳۹۷). روایی توصیفی به صحت داده‌ای موضوع گزارش شده توسط محققان اشاره دارد. برای حصول روایی توصیفی، مصاحبه‌ها همان روز پیاده‌سازی و تحلیل شدند. راه اطمینان از روایی تفسیری، دریافت بازخورد از مشارکت‌کنندگان است. ما در این پژوهش پس از کدگذاری، محتوای کدگذاری شده مصاحبه را برای مصاحبه‌شونده ارسال و بازخورد دریافت کردیم. برای حصول روایی نظری از تجربه پنج نفر از متخصصان علم اطلاعات استفاده کردیم. برای سنجش پایایی پژوهش نیز از معیار اطمینان‌پذیری<sup>۸</sup> که از معیارهای مطرح شده توسط «گوبا»<sup>۹</sup> و «لینکلن»<sup>۱۰</sup> است، استفاده شد؛ به این صورت که دو نفر برای انجام مصاحبه به صورت جداگانه اما موازی انتخاب و

1. Mendeley Data

2. Zendo

3. Genbank

4. Drayad

5. UK Data Service

6. Respondent-driven sampling

7. MaxQDA

8. dependability

9. Guba

10. Lincoln

یافته‌های آن‌ها با هم مقایسه شد.

#### ۴. یافته‌ها

##### ۴-۱. تحلیل محتوای متون

برای پاسخ به سؤال اول پژوهش از روش تحلیل محتوای متون استفاده شد؛ به این صورت که بعد از جست‌وجوی نظام‌مند در پایگاه‌های اطلاعاتی معتبر و بررسی مقالات بازبایی شده در نهایت، ۴۳ مقاله تحلیل محتوا و تمامی عوامل مستخرج از آن‌ها به‌عنوان یک کد در نظر گرفته شدند. پس از بررسی کدها با توجه به مفهوم آن‌ها ارتباطات مشخص شد. سپس، کدهای با مفاهیم مشابه در کنار یکدیگر قرار گرفته و مقوله‌ها به‌دست آمد. این شیوه تبدیل مفاهیم به مقوله‌ها ادامه پیدا کرد تا ابعاد اصلی اولیه شناسایی شدند. به این ترتیب، با تحلیل محتوای متون ۱۳۵ کد در ۱۹ مقوله و ۳ بُعد اصلی شامل عوامل انسانی (با ۵ مقوله)، عوامل سازمانی (با ۷ مقوله) و زیرساخت‌ها (با ۷ مقوله)، شناسایی شده و کدها، مقوله‌ها و ابعاد به همراه منابع پشتیبانی‌کننده و فراوانی هر کدام در قالب جدولی ارائه شد. اما به دلیل حجم زیاد و جلوگیری از تکرار (زیرا کدها، مقوله‌ها و ابعاد نهایی به‌طور مشروح در جدول شماره ۴، همین مقاله آمده‌اند)، برای نشان دادن شیوه کار و یافته‌های بخش تحلیل محتوای متون، تنها دو مؤلفه از بُعد زیرساخت در جدول شماره ۳، ارائه شده است.

### جدول ۳. بخشی از کدها، مقوله‌ها و ابعاد حاصل از مرحله فراترکیب

ابعاد	مقوله‌ها	کدها	منابع	فراوانی
زیرساخت	سخت‌افزار	لزوم طراحی مخازن داده	Kim and Nah 2018; Sposito 2017; Yoon and Kim 2017; Carlson and Anderson 2007; Faniel and Zimmerman 2011; Duke and Porter 2013; Curty and Qin 2014; Chao 2011; Mooney and Newton 2012	۹
		لزوم طراحی زیرساخت برای پشتیبانی از جمع‌آوری، مدیریت و کشف داده	Sands et al. 2012; Wilson et al. 2011; Wallis, Rolando, Borgman 2013; Sposito 2017; Rolland and Lee 2013; Thanos 2017; Murillo 2014;	۷
		امکانات ذخیره‌سازی در مقیاس بالا	Wilson et al. 2011;	۱
		کامپیوترهای با ظرفیت بالا برای دسترسی و ذخیره داده‌ها	Wilson et al. 2011;	۱
		توسعه ابزار مدیریت داده	Yoon and Kim 2017;	۱
		نگهداری از زیرساخت‌ها	Rolland and Lee 2013;	۱
ابرداده		اختصاص شناساگر منحصر به فرد	Schöpfel et al. 2015; Curty et al. 2017; Holub et al. 2018; Bishop 2014; Thanos 2017; Chao 2011; Mooney and Newton 2012;	۷
		تعریف استانداردهای ابرداده‌ای برای الصاق به داده‌ها	Tenopir, et al. 2015; Wallis, Rolando, Borgman 2013; Commission notice 2014; McAllister 2018; Sposito 2017; Murillo 2014; Yoon and Kim 2017; Thanos 2017; Duke and Porter 2013; Chao 2011; Mooney and Newton 2012;	۱۱
		نیاز به توصیف و ساختار یافتگی داده‌ها برای کشف آن‌ها	Murillo 2014; Darby, et al. 2012; Schöpfel et al. 2015;	۳

ابعاد	مقوله‌ها	کدها	منابع	فراوانی
زیرساخت ابر داده	مستندسازی داده‌ها		Pasquetto, Randles and Borgman, 2017; Mannheimer, Stermann, and Borda 2016; Faniel and Zimmerman 2011; Vision and Piwovar 2013; Commission notice 2014; Yardley, et al. 2014; Yoon 2016; Rolland and Lee 2013; Carlson and Anderson 2007; Thanos 2017; Curty and Qin 2014; Moore 2006; Mooney and Newton 2012; Faniel and Jacobsen 2010;	۱۶

#### ۴-۲. مصاحبه

برای پاسخ به سؤال دوم پژوهش در این مرحله، از آنجا که ابعاد مستخرج از مرحله تحلیل محتوای متون بومی نیستند و به‌منظور تجدید نظر و اصلاح و تعدیل این عناصر و ابعاد، از نظرات ۱۹ تن از خبرگان استفاده شد که البته برای رسیدن به پاسخ‌های آگاهانه به‌طور خاص این مصاحبه‌ها با پژوهشگرانی که با مقوله استفاده مجدد آشنایی و یا تجربه آن را داشتند، انجام شد. مصاحبه‌ها با نرم‌افزار «مکس کیودا» کدگذاری شد. کدگذاری داده‌های حاصل از مصاحبه مشکل‌تر از کدگذاری داده‌های حاصل از پیمایش است و طبیعتاً مراحل بیشتر و طولانی‌تری دارد. یکی از معیارهای دسته‌بندی در این کدگذاری، یافتن شباهت‌ها و تفاوت‌ها و پیدا کردن الگوهایی در داده‌های به‌دست آمده است (استراوس و کوربین ۱۹۹۰، ۲۱-۲۲). تحلیل داده‌های حاصل از مصاحبه منجر به شناسایی ۳۴۸ کد در ۳ بُعد شد. با این که عوامل مستخرج از تحلیل محتوا مبنای هدایت مصاحبه‌ها بود، اما کدهای حاصل از مصاحبه با کدهای حاصل از تحلیل محتوا در برخی مقوله‌ها و کدها یکسان نبودند. به‌طور مثال، مصاحبه‌شوندگان به فراداده کمتر پرداختند، اما بُعد سازمانی در مصاحبه‌ها بیشتر مورد توجه بود.

#### ۵. بحث

در نتیجه ادغام نتایج حاصل از دو مرحله تحلیل محتوای متون و مصاحبه با خبرگان، در نهایت، ابعاد ۳ گانه و مقوله‌های ۲۱ گانه استفاده مجدد از داده‌های پژوهشی در ایران حاصل شد. جدول شماره ۴، ابعاد و مقوله‌های زیرمجموعه هر یک از آن‌ها را نشان می‌دهد. این سه بعد عبارت‌اند از: زیرساخت‌های فنی و محتوایی، منابع انسانی، و عوامل سازمانی.

#### جدول ۴. ابعاد و مؤلفه‌های استفاده مجدد حاصل از مرور متون و مصاحبه<sup>۱</sup>

بعد زیرساخت‌ها	
مدیریت داده‌ها (۱۲)	
ارزیابی و مدیریت داده‌ها قبل از استفاده مجدد	مشخص کردن متغیرها در فایل مجموعه داده
نظارت بر روش شناسی پژوهش‌ها به منظور پرهیز از داده‌سازی	مشخص کردن معیارها در یک فایل الصاقی
مشخص کردن قالب مشخص برای داده‌ها در هر رشته	اصل استقلال مقاله اصلی
ارائه اطلاعات زمینه‌ای (باقی) در مورد داده‌ها	مشخص بودن منشأ داده
مشخص بودن واحد اندازه‌گیری	لینک داده‌ها به پایان‌نامه‌ها و مقالات چاپ شده در مجلات
طبقه‌بندی داده‌ها قبل از استفاده مجدد	لزوم تگ گذاری داده‌ها قبل از طبقه‌بندی
فراداده‌ها (۴)	
توجه به مستندسازی و ساختاریافتگی داده‌ها	اختصاص شناساگر منحصربه‌فرد به هر مجموعه داده
تعریف استانداردهای ابرداده‌ای برای الصاق به داده‌ها	چگونگی ارتباط (الصاق) ابرداده با داده
زیرساخت‌های ارتباطی و مخابراتی (۶)	
نیاز به سخت‌افزارهای شبکه	نیاز به خطوط وایرلس
نیاز به زیرساخت‌های ارتباطی و مخابراتی	گسترش شبکه با پهنای باند بالا
استفاده از رایانش ابری	سازگاری زیرساخت‌های ارتباطی با زیرساخت‌های نرم‌افزاری و سخت‌افزاری
زیرساخت‌های سخت‌افزاری (۱۰)	
نیاز به رسانه‌های ذخیره‌سازی داده	نیاز به سیستم پردازشگر یا سی‌پی‌یو قوی چند هسته‌ای
کامپیوترهای با ظرفیت بالا برای ذخیره، مدیریت و دسترسی به داده‌ها	نیاز به حافظه داخلی یا حافظه نهان
فراهم‌نمودن سرور به عنوان یکی از زیرساخت‌های اصلی	نیاز به دیسک‌های سخت قوی
لزوم توجه به نگهداری از زیرساخت‌ها	نیاز به رید کردن <sup>۱</sup> چندین هارد یا نوار مغناطیسی
نیاز مخزن به پشتیبانی کافی و لحظه‌ای	توجه به نوع داده در تهیه زیرساخت‌ها

1. raid

### زیروساخت‌های نرم‌افزاری (۱۵)

ایجاد یک پایگاه دادهٔ مجتمع یا چند پایگاه جداگانه برای انواع داده	نیاز به نرم‌افزارهای خاص با توجه به نوع داده
طراحی یک رابط کاربر کاربرپسند	نیاز به نرم‌افزار مشابهت‌جو
طراحی رابط کاربر گرافیکی	نیاز به نرم‌افزارهای بیوانفورماتیک <sup>۱</sup>
لزوم طراحی موتور کاوش	نیاز به نرم‌افزارهای آماری
طراحی یک پلت‌فرم با قابلیت جست‌وجوی یکپارچه	لزوم طراحی پایگاهی برای داده‌های کیفی
طراحی فهرستی از داده‌ها	طراحی پایگاهی برای پشتیبانی از داده‌های چندرسانه‌ای
طراحی هستی‌شناسی (آنتولوژی) دامنهٔ خاص	لزوم طراحی سیستم‌های ورودی و خروجی

توجه به خدمات نوین در دنیای دیجیتال (از جمله آراس‌اس<sup>۲</sup>، آلرت<sup>۳</sup> و غیره)

### اقدامات امنیتی (۵)

حفظ امنیت داده‌ها	نیاز به زیرساخت‌های سایبری
توجه به سرقت و هک داده‌ها	محافظت فیزیکی از داده‌ها در مخازن
لزوم توجه به ابزارهای پشتیبان‌گیری	

### مسائل اقتصادی (۵)

نیاز به منابع مالی برای شروع کار	تأمین قسمتی از بودجه با ارزش افزوده حاصل از استفادهٔ مجدد از داده‌ها
نیاز به منابع مالی برای نگهداری و تعمیر	لزوم داشتن شفافیت در مسائل اقتصادی
دریافت هزینه در قبال خدمات مخزن	

### بعد منابع انسانی

### نیروی انسانی (۱۰)

لزوم به کارگیری نیروی انسانی خاصی جهت سنجش کیفیت داده	لزوم به کارگیری متخصصان حقوقی
لزوم به کارگیری نیروی انسانی خاص جهت نگهداری و مدیریت داده	لزوم به کارگیری متخصصان علم اطلاعات
نیاز به کارگیری پژوهشگران خیلی قوی	لزوم به کارگیری متخصصان علم‌سنجی
لزوم به کارگیری متخصصان داده و داده‌کاوی	کاهش نقش نیروی انسانی به منظور کاهش خطا در مخازن

1. bioinformatic

2. Rss

3. alert

لزوم به کارگیری متخصصان سخت‌افزار و نرم‌افزار | لزوم به کارگیری متخصصان موضوعی (با توجه به حوزه موضوعی مخزن)

#### مسائل اجتماعی (۵)

ترمیم سرمایه اجتماعی پایین و عدم اعتماد مردم به دولت | ترویج مشارکت عمومی  
ترمیم عدم اعتماد مردم به پژوهشگران | لزوم افزایش اعتماد اجتماعی  
توجه به منافع (سود) عمومی

#### اخلاق حرفه‌ای (۳)

بالا بردن روحیه مشارکت و اعتماد | احساس مسئولیت  
نهادینه کردن بحث عدالت و انصاف

#### نگرش و انگیزش (۶)

تشویق پژوهشگران در قبال سپردن داده‌ها | تغییر سیستم ارتقای اعضای هیئت علمی در دانشگاه‌ها  
نیاز به تغییر نگرش محققان نسبت به استفاده مجدد | طراحی سیستم‌های پاداش  
نیاز به اعتمادسازی بین پژوهشگران | استفاده از مکانیسم‌های تشویقی برای محقق صاحب داده و استفاده‌کننده

#### اقدامات فرهنگی (۶)

فرهنگ‌سازی برای کار گروهی | نیاز به تبلیغات از طریق رسانه‌های علمی  
عدم داده‌سازی در کارهای گروهی | طراحی یک کاتالوگ برای معرفی خدمات ارائه شده  
پایین بودن میزان خطا در پژوهش‌های گروهی | نیاز به فرهنگ‌سازی برای استفاده مجدد

#### آموزش (۶)

لزوم تدوین برنامه‌های آموزش رسمی | توجه به سوادهای مورد نیاز (سواد اطلاعاتی، دیجیتالی، رسانه‌ای، و ...)  
آشنایی نیروی انسانی با اطلاعات مرتبط با نوع داده | برگزاری دوره‌های مشاوره برای کاربران  
نیاز به برگزاری آموزش‌های ضمن خدمت برای کارکنان | نیاز به آموزش و توانمندسازی پژوهشگران

#### بعد سازمانی

#### مسائل بوروکراتیک (۸)

توجه به رهبری سازمان | خارج شدن تولید علم و داده از حیطه دولتی (خصوصی‌سازی)  
طراحی استراتژی برای سازمان | نیاز به ساختار دقیق، منظم و مشخص برای سازمان



لړوم توجه به تقسیم کار

توجه به ساختار سلسله‌مراثی سازمان

تناسب تعداد نیروی انسانی با حجم کار

نیاز به برنامه‌ریزی مدون

### اعتبار سازمان (۳)

مدیریت بهتر داده‌ها با سازمان تخصصی‌تر

اعتماد محققان بر اساس اعتبار سازمان

کنترل بهتر کیفیت داده با سازمان تخصصی‌تر

### الزامات سازمانی (۶)

تأسیس مؤسساتی که داده تولید کنند (کارشان صرفاً تولید داده باشد)

نیاز به منابع مالی برای سازمان

الزام دانشگاه‌ها به طراحی مخزن داده

جذب پژوهشگر توسط سازمان

نیاز به قدرت اجرایی سازمان

توجه به مأموریت‌ها و امکانات سازمان

### سازمان متولی (۱۲)

سازمانی می‌تواند متولی باشد که دسترسی به داده داشته باشد.

مشارکت شرکت‌های خصوصی به‌عنوان سازمان متولی

نیاز به طراحی مخزن در سطح کشور نه مؤسسات

شبکه‌های تحقیقاتی به‌عنوان سازمان متولی

همکاری و مشارکت چند سازمان به‌عنوان سازمان متولی

سازمان پژوهش‌های علمی و صنعتی به‌عنوان سازمان متولی

وزارت ارتباطات و فناوری اطلاعات به‌عنوان سازمان متولی

انجمن‌های تخصصی به‌عنوان سازمان متولی

وزارت بهداشت به‌عنوان سازمان متولی

ایرانداک به‌عنوان سازمان متولی

وزارت علوم به‌عنوان سازمان متولی

کتابخانه ملی به‌عنوان سازمان متولی

### ارتباطات ملی و بین‌المللی (۷)

مبادله بین‌المللی داده (کنش‌پذیری یا همکاری با سایر کشورها)

الگوبرداری و استفاده از قابلیت‌های شبکه‌های علمی در برقراری ارتباطات

برقراری امکان ارتباط و مکاتبه بین صاحبان داده با مدیران سیستم

جلب پشتیبانی مؤسسات، دانشگاه‌ها، جوامع علمی و پژوهشگران

همکاری مجلات با مخازن

کمک گرفتن از مسئولان امور پژوهشی دانشگاه‌ها و مؤسسات پژوهشی

لړوم برقراری ارتباط بین محقق صاحب داده با استفاده‌کننده مجدد

### مسائل اخلاقی و حقوقی (۱۲)

توجه به مسائل اخلاقی و لړوم تدوین راهنمای اخلاقی

دادن تعهد توسط استفاده‌کنندگان مجدد

رعایت کپی‌رایت و مالکیت معنوی داده‌ها	مشخص کردن سهم و تعهد هر دو محقق
مشخص کردن نحوه استناد صحیح به مجموعه داده‌های مخزن	کسب اجازه از محقق اول
تدوین دستنامه مخصوص شیوه استناددهی	استناددهی به محقق اول
رعایت حریم خصوصی و محرمانه ماندن اطلاعات	پرداخت هزینه به محقق صاحب داده
تضمین جلوگیری از سوء استفاده از داده	دادن امتیازات خاص به محقق اول

#### خط‌مشی‌های اجرایی (۱۵)

الگو گرفتن از قابلیت‌های شبکه‌های اجتماعی	تسهیل چاپ مقاله برای تشویق محقق به سپردن داده به مخازن
ایجاد امکان آشنایی و همکاری افراد با زمینه پژوهشی مشابه	پرهیز از دوباره کاری و تکرار مخازن موجود خارج از کشور
مشخص کردن و به کارگیری راه‌های دریافت داده از محققان	الزام در دریافت داده به همراه نسخه پایان‌نامه یا مقاله
امکان ثبت درخواست داده در سامانه	ایجاد امکان اطلاع‌رسانی به اعضا در صورت انتشار داده جدید
امکان ارسال درخواست داده از سامانه به محقق صاحب داده	نیاز به سیستم ارزیابی و نظارت
معرفی بخشی از داده‌ها به عنوان نمونه در ابتدا	تعامل با کاربر و شناسایی نیازهای کاربر (گرفتن بازخورد از ذی‌نفعان)
طراحی یک سیستم مشاوره و پرسش و پاسخ	توجه به بحث آمارگیری در مواقع نیاز
بهره‌گیری از شاخص‌های کتاب‌شناختی برای اندازه‌گیری میزان استفاده مجدد از داده‌ها	

#### کیفیت داده (۶)

لزوم سنجش کیفیت داده‌ها	لزوم وجود فرایندی برای داوری داده‌ها
لزوم وجود معیارهایی خاص برای سنجش کیفیت داده‌ها	طراحی سیستمی شفاف برای اندازه‌گیری و ارزش‌دهی به داده‌ها
بررسی و توجه به اندازه و حجم داده‌ها	عدم دستکاری داده در مخازن

#### قوانین و استانداردها (۶)

نیاز به تدوین چارچوب قانونی خاص (با قوانین شفاف)	نیاز به استانداردهای بازیابی و ذخیره‌سازی داده
عقد قرارداد بین استفاده‌کننده مجدد و مخزن	وضع قانونی به منظور استعلام از مخزن قبل از چاپ مقالات
نیاز به تدوین استانداردهای داده (در کل چرخه حیات داده)	نظارت بر اجرای صحیح قوانین

سیاست‌های دسترسی به داده‌ها (۸)

دسترسی کنسر سیومی	لزوم قرار دادن محدودیت در دسترسی به داده‌ها
مشخص و شفاف بودن شرایط دسترسی به داده‌ها	لزوم ایجاد حساب کاربری برای هر کاربر
تنظیم موافقت‌نامه (رضایت‌نامه) برای کسب اجازه در استفاده مجدد	قرار دادن سطوح دسترسی به خدمات مخزن
فراهم آوردن امکان استفاده مجدد از داده بعد از یک بازه زمانی مشخص	دسترسی آزاد و آسان

۵-۱. بُعد زیرساخت‌ها

دسترسی و استفاده مجدد از داده‌های موجود در اولین گام نیازمند زیرساخت‌هاست و بدون وجود سطح معتدلی از آن‌ها استفاده مجدد از داده‌ها امکان‌پذیر نخواهد بود. «والیس، رولاندو و بورگمن» عنوان می‌کنند که علی‌رغم اقبال کنونی به بحث‌های مدیریت و استفاده مجدد از داده‌ها، زیرساخت‌ها و انگیزه‌های لازم برای این امور وجود ندارد (Wallis, Rolando, Borgman 2013). «افخمی» نیز معتقد است که استفاده بهینه و حداکثری از منابع دسترسی آزاد و همگانی پژوهشگران بدون وجود زیرساخت‌های قوی که بتواند در درازمدت نیازهای مختلف پژوهشگران و تغییرات مداوم در فناوری اطلاعات را پوشش دهد، به دست نخواهد آمد (۱۳۸۹). بُعد زیرساخت شامل مؤلفه‌های مدیریت داده و ابرداده، زیرساخت‌های سخت‌افزاری، نرم‌افزاری، ارتباطی و مخابراتی، مسائل اقتصادی و امنیتی است. در زمینه زیرساخت‌های سخت‌افزاری بیشتر از همه بر نیاز به کامپیوترهای با ظرفیت‌های بالا برای ذخیره و دسترسی به داده‌ها و امکانات ذخیره‌سازی در مقیاس بالا (Wilson et al. 2011) تأکید شده است و مصاحبه‌شوندگان نیز مواردی مانند سیستم پردازشگر قوی چند هسته‌ای، حافظه داخلی یا حافظه نهان<sup>۱</sup>، حافظه‌های با حجم بسیار بالا و دیسک‌های سخت با ظرفیت بالا را عنوان کرده‌اند. در بحث زیرساخت‌های نرم‌افزاری نیز، مواردی مانند نرم‌افزار رایگان و آسان برای انتقال داده‌ها (Faniel and Zimmerman 2011) لزوم طراحی پایگاه داده متناسب با انواع داده‌ها، رابط کاربر گرافیکی، نرم‌افزارهای آماری، بیوانفورماتیک و مشابهت‌جو عنوان

1. Cache memory

2. hard disk

شدند. مؤلفه بعدی این بُعد زیرساخت‌های ارتباطی است. گسترش شبکه با پهنای باند بالا از جمله این زیرساخت‌ها محسوب می‌شود (Wilson et al. 2011). برخی مصاحبه‌شوندگان معتقدند که مشکل اصلی ما در کشور بحث زیرساخت‌های ارتباطی است: «شرکت‌هایی در کشور هستند که نرم‌افزارهای بسیار جامع‌تر از این هم طراحی کردن. به نظر من مشکل نرم‌افزاری و سخت‌افزاری در کشور نداریم. مشکل اصلی ما در زیرساخت‌های ارتباطیه». استفاده بهینه از داده‌های پژوهشی تنها در صورتی امکان‌پذیر است که هزینه‌های مدیریت داده‌ها به‌عنوان یک هزینه ثابت و جزء لاینفک فرایند پژوهش در نظر گرفته شود (افخمی ۱۳۸۹). مسائل اقتصادی از دیگر مواردی بود که در بحث زیرساخت‌ها مورد توجه قرار گرفت. نیاز به سرمایه‌گذاری کافی و لزوم تأمین منابع مالی (Pasquetto, Randles, and Borgman 2017; Darby et al. 2012; Commission notice 2014)، تأمین هزینه‌های تعمیر و نگهداری زیرساخت‌ها و شفافیت در مسائل مالی (Commission notice 2014) در مقوله مسائل اقتصادی مطرح شدند. البته، به این نکته نیز توجه شد که خود بحث استفاده مجدد از داده‌های دیگران برای مقاصد جدید می‌تواند باعث بازگشت سرمایه و جبران هزینه‌ها شود. همان‌گونه که یکی از مصاحبه‌شوندگان معتقد است «زیرساخت‌های اقتصادی و تأمین بودجه برای ایجاد پایگاه داده می‌تواند از ارزش افزوده حاصل از همین استفاده مجدد از داده‌ها تأمین شود». وجود پنج کد در تحلیل محتوای متون نشان‌دهنده این است که بحث‌های امنیتی جایگاه ویژه‌ای در این حوزه دارند. مصاحبه‌ها نیز به این مطلب صحنه گذاشتند: «به نظر من بیشتر، حفاظت از داده‌ها مهمه. من به‌عنوان یک محقق توقع دارم از جایی داده‌ها را بگیرم که امنیت داشته باشه». باید ابزار و امکانات حفاظت از جامعیت و امنیت داده‌ها در برابر تهدیداتی چون از بین رفتن، گم شدن، تغییرات ناخواسته و دسترسی غیرمجاز به‌صورت عمدی و یا غیرعمدی پیش‌بینی شود. از سوی دیگر، پایگاه داده و ابزار ذخیره‌سازی آن‌ها باید در مقابل خطرات محیطی چون گرما، گردوخاک، شوک‌های الکتریکی، ربایش و تخلیه‌های الکتریکی مصون باشد. علاوه بر حفاظت فیزیکی داده‌ها، تمهیداتی نیز درباره مسئله کنترل و جلوگیری از افشای آماری که منجر به شناسایی افراد و اطلاعات مربوط به آن‌ها می‌شود، باید اندیشیده شود (افخمی ۱۳۸۹). داده‌های علمی در صورتی که به‌درستی سازماندهی و

مدیریت شوند، می‌تواند دیدگاه‌های جدید و نوآوری‌هایی را در پژوهش‌ها ایجاد کند. بنابراین، باید قبل از هر اقدامی برای استفاده مجدد از داده‌ها، مدیریتی صحیح بر روی آن‌ها اعمال و استانداردهای ابرداده‌ای به این داده‌ها الصاق شود. ابرداده به‌عنوان اطلاعاتی درباره محتوای، کیفیت، شرایط، علّت و دیگر خصوصیات داده تعریف می‌شود و تهیه آن نیازمند رعایت الزامات و استانداردهای ویژه‌ای است (فرجی سبکبار و همکاران ۱۳۹۲).

#### ۲-۵. بعد منابع انسانی

بعد عوامل انسانی شامل محققانی است که داده‌های خود را برای استفاده مجدد به مخازن می‌سپارند؛ محققانی که از داده‌های مخازن استفاده می‌کنند و همچنین کسانی که در روند آماده‌سازی داده‌ها همکاری می‌کنند. این بعد شامل مؤلفه‌های نیروی انسانی، آموزش، نگرش و انگیزش، اقدامات فرهنگی و مسائل اجتماعی است. در بحث نیروی انسانی، نویسندگان به لزوم به کارگیری افراد با تخصص‌های مرتبط اشاره کرده‌اند (Darby et al. 2012) مصاحبه‌شوندگان نیز بر این مسئله تأکید دارند: «من فکر می‌کنم برای نیروی انسانی به کمک متخصصان فناوری اطلاعات، مسئولین پژوهشی دانشگاه‌ها، سردبیران و داوران مجلات علمی پژوهشی، متخصصان علم اطلاعات و علم‌سنجی نیاز». بعد از به کارگیری و مدیریت کارشناسان متخصص، لزوم آموزش رسمی و توانمندسازی آنان مطرح می‌شود (Tenopir et al. 2015; Bishop 2012; Darby et al. 2012; Wilson et al. 2011; Wilson et al. 2014) البته بحث آموزش پژوهشگران نیز یک ضرورت محسوب می‌شود (Prince et al. 2017; Wallis, Rolando, and Borgman 2013). آموزش‌های ضروری از نظر پژوهشگران عبارت‌اند از: سازماندهی فایل‌ها، پیوند یادداشت‌ها به محتوا، پشتیبان‌گیری، آگاهی از این که چه ابزارهای نرم‌افزاری در دسترس هستند و کدام‌یک از نرم‌افزارها برای مقابله با چالش‌های تحقیق خاص مناسب است، و ساختاربندی داده‌ها به‌ویژه در پایگاه‌های داده رابطه‌ای (Wilson et al. 2011). دیدگاه و نگرش محققان در خصوص استفاده مجدد از داده‌های موجود، مؤلفه بعدی این بعد است. باید نظام تشویق و ترغیب در جهت دسترسی آزادانه و عمومی به داده‌ها را پیش‌بینی نمود؛ هرچند قوانین و مقررات در تسهیل دسترسی به داده‌ها ضروری‌اند، ولی در عین حال هنجارها و مشوق‌ها برای پژوهشگران مفید

خواهد بود. از جمله مسائل فرهنگی در این زمینه تغییر نگاه پژوهشگران است. شاید بتوان گفت فاکتورهای فرهنگی و اجتماعی از تأثیرگذارترین فاکتورها در استفاده مجدد از داده‌ها هستند (Curty et al. 2017; Faniel and Zimmerman 2011; Wallis, Rolando and Borgman 2013; Darby et al. 2012). علاوه بر مسائل فرهنگی، به مسائل اجتماعی همچون لزوم اعتماد اجتماعی و توجه به سرمایه اجتماعی نیز باید توجه شود. به‌طور مثال، یکی از مصاحبه‌شوندگان معتقد است که «یک بیماری تاریخی در ایران داریم به نام عدم اعتماد یا سرمایه اجتماعی پایین. عدم اعتماد بین مردم و دولت. مردم به پژوهشگران هم اعتماد کمی دارند». به اعتقاد ایشان ما برای رسیدن به داده‌های صحیح و قوی، نیاز به ترمیم این وضعیت جامعه داریم. «ترمیم این اعتماد هم جزو زیرساخته و جزو کارهایی است که در کوتاه‌مدت نمی‌توانیم بهش برسیم، یک بیماریه که در طولانی مدت به وجود آمده. تا اون عدم اعتماد رو درست نکنیم داده‌ها مون قوی نیست. هر چقدر هم که پژوهشگر قوی عمل کنه من مطمئنم که جواب‌هایی که از جامعه آماری دریافت می‌کنه، به خاطر عدم اعتماد، واقعی نیستند پس من ترجیح می‌دم خودم داده جمع کنم». اخلاق حرفه‌ای نیز از مقوله‌هایی است که در بعد منابع انسانی مورد توجه بود. در این بعد بر بالا بردن مشارکت و کار گروهی تأکید شد. به‌طور مثال، مصاحبه‌شونده‌ای معتقد است که: «اگر کارهای پژوهشی به‌صورت گروهی انجام بشه میزان خطا کمتر میشه. همه که با هم باشیم میزان خطا میاد پایین و داده‌سازی هم نمیشه».

### ۳-۵. بعد سازمانی

از آنجا که سازمان‌ها و ناشران به‌شدت می‌توانند رفتار پژوهشی محققان را تحت تأثیر قرار دهند، عوامل سازمانی از جمله الزامات و اعتبار سازمان، سیاست‌ها و خط‌مشی‌های سازمان، مسائل اخلاقی و حقوقی، کیفیت داده، و سیاست‌های دسترسی به داده از عوامل تأثیرگذار بر استفاده مجدد هستند. زمان، هزینه، دانش و مهارت لازم برای آماده‌سازی داده‌ها، به‌منظور استفاده مجدد توسط دیگران، خارج از توان پژوهشگر است و بایستی یک سازمان، متولی این امر شود که در مورد سازمان متولی بین مصاحبه‌شوندگان اختلاف نظر بود؛ وزارت‌خانه‌ها (وزارت علوم، وزارت بهداشت و وزارت ارتباطات)، «پژوهشگاه علوم و فناوری اطلاعات ایران»،

دانشگاه‌ها، سازمان‌های خصوصی و «کتابخانه ملی» از جمله گزینه‌هایی بودند که به‌عنوان سازمان متولی مطرح شدند. اما اکثر مصاحبه‌شوندگان هم‌عقیده بودند که سازمان متولی باید ساختار دقیق و منظم، امکانات فنی و اقتصادی، و قدرت اجرایی بالایی داشته باشد. فارغ از این که این سازمان دولتی یا خصوصی باشد، نیاز به یک سری سیاست‌ها، خط‌مشی‌ها و قوانین دارد. به جرأت می‌توان ادعا کرد که یکی از چالش‌های اساسی در بحث استفاده مجدد از داده‌ها، بحث‌های مربوط به مسائل اخلاقی و قانونی آن است. عدم اطمینان قانونی می‌تواند مانعی اساسی بر سر راه استفاده مجدد از داده‌ها باشد و در مقابل، ارائه چارچوب قانونی برای حمایت از پژوهشگران به آنان انگیزه خواهد داد. حفظ حریم خصوصی، محرمانگی، حق مؤلف<sup>۱</sup>، حق استفاده و مالکیت آن‌ها، کسب رضایت آگاهانه، مسائل مربوط به امنیت ملی، قوانین ملی و توافقات بین‌المللی، رایج‌ترین مسائل حقوقی مربوط به استفاده مجدد از داده‌ها هستند (Schöpfel et al. 2015; Darby et al. 2012; Fecher, Friesike and Hebing 2015; Houlb et al. 2018; Commission notice 2014). تحریف یا دستکاری داده‌ها، سوءاستفاده تجاری و رقابتی از آن‌ها، و سوءتفسیر کاربران از دیگر مسائلی است که ضرورت توجه به الزامات حقوقی و سیاسی استفاده مجدد از داده‌ها را بیشتر می‌کند (Fecher, Friesike, and Hebing 2015). تدوین یک چارچوب بر اساس قوانین ملی و توافق‌های بین‌المللی، پایبندی به مسائل اخلاقی، کسب رضایت آگاهانه از محققان و پرهیز از هرگونه ضرر و زیان به دیگران از مواردی است که مورد توجه مصاحبه‌شوندگان بود. حتی مصاحبه‌شونده‌ای به لزوم تدوین یک راهنمای اخلاقی مستقل برای استفاده مجدد اشاره کرد. مطالعات نشان می‌دهند که مسئله استناددهی به مجموعه داده‌ها، همچنین نیاز به تدوین استانداردهای شفاف و دقیق استناددهی برای هماهنگی در استناددهی، از الزامات روند استفاده از داده‌های دیگران است (Curty and Qin 2014; Pasquetto, Randles, and Borgman 2017; Bishop and Bishop 2013; Piowar and Vision 2013; Wallis, Rolando, and Borgman 2013; Kuula-Lummi 2017). «مونی و نیوتن» معتقدند که تصمیم نویسنده برای استناددهی به داده‌ها تحت تأثیر استانداردهایی قرار می‌گیرد که توسط دستنامه‌های استناد، سیاست‌های مجله، و

ارائه‌دهندگان داده‌ها تدوین شده‌اند (Mooney and Newton 2012). بنابراین، سه عنصر سبک، استانداردهای استناد و رفتار نویسنده باید در بحث استناددهی مورد توجه قرار گیرند. برخی پژوهشگران شرط استفاده مجدد از داده‌ها را سنجش کیفیت آن‌ها ذکر کرده‌اند (Darby et al. 2012; Curty et al. 2017; Commission notice 2014; Holub et al.). ارزش (Vision & Piwowar 2013; Faniel and Zimmerman 2011; Schöpfel et al. 2015; 2018). ارزش داده‌ها به کیفیت آن‌ها بستگی دارد. کیفیت بدین معناست که شرایط مبهم مرتبط با تولید داده‌ها (مانند تکنیک‌های جمع‌آوری، آرشیو، مقیاس‌های اندازه‌گیری و توزین آن‌ها) توضیح داده شود. استانداردهای افزایش کیفیت باید با مشاوره با پژوهشگران هر حوزه علمی تدوین شود (افخمی ۱۳۸۹). مؤلفه بعدی این بُعد بحث قابلیت و کنترل دسترسی به داده‌هاست. نویسندگان زیادی طرفدار دسترسی آزاد به داده‌ها هستند (Bishop and Kuula-Lummi 2017; Mannheimer, Sterman and Borda 2016; Tenopir et al. 2015; Schöpfel and Južnič 2015; Bishop 2014; Faniel et al. 2013; 2018). این در حالی است که دسترسی آسان و بلندمدت به داده‌ها (Commission notice 2014) و سهولت کشف داده‌ها (Darby et al. 2012; Curty 2016; Wilson et al. 2011; Vision & Piwowar 2013;) منجر به افزایش تمایل پژوهشگران برای استفاده مجدد از داده‌ها خواهد شد. در داده‌های پژوهشی که با بودجه عمومی تولید شده‌اند، اصل بر این است که دسترسی حداکثری و ارزان باشد. البته، محدودیت‌های قانونی و اخلاقی در دسترسی در چارچوب سیاست‌ها وجود دارد. از جمله این محدودیت‌ها دلایل امنیتی، حفاظت از حقوق کاربر، و یا حفظ اسرار افراد است (افخمی ۱۳۸۹).

## ۶. نتیجه‌گیری

پیشرفت‌های سریع در فناوری‌های اطلاعاتی و ارتباطی همه جنبه‌های زندگی بشر را تحت تأثیر قرار داده است. یکی از مباحثی که چند سالی است تحت تأثیر این فناوری‌ها مورد توجه قرار گرفته، بحث تولید روزافزون داده‌های پژوهشی است. دسترسی گسترده به داده‌های پژوهشی و استفاده مجدد از آن‌ها باعث بازگشت سرمایه و حفظ ارزش داده‌ها خواهد شد. در همین راستا، این پژوهش به وسیله تحلیل محتوای متون و مصاحبه با خبرگان به شناسایی عناصر لازم به منظور استفاده مجدد از داده‌های پژوهشی در ایران پرداخت.



عناصر شناسایی شده در ۳ بعد زیرساخت‌ها، منابع انسانی، و عوامل سازمانی تقسیم‌بندی شدند. زیرساخت‌های سخت‌افزاری، نرم‌افزاری، ارتباطی، مسائل اقتصادی، مسائل امنیتی و فراداده از جمله زیرساخت‌ها هستند. مؤلفه‌های بُعد عوامل انسانی شامل نیروی انسانی، آموزش، اقدامات فرهنگی، نگرش و انگیزش، اخلاق حرفه‌ای و مسائل اجتماعی است. توجه به مسائل اخلاقی و حقوقی، کیفیت داده، سیاست‌های دسترسی، مالکیت معنوی و استناددهی، الزام و اعتبار سازمانی، خط‌مشی‌های اجرایی و مسائل بوروکراتیک از مؤلفه‌های اصلی بُعد سازمانی است. با توجه به نقش داده‌ها در روند پژوهش‌ها و توسعه مفهوم استفاده مجدد از داده‌های موجود در دنیا، این مفهوم بایستی به فرهنگ و رویکرد معمول و متداول در جامعه علمی کشور بدل گردد. طبیعی است که اجرای این رویکرد با چالش‌ها و مشکلات فنی و غیرفنی زیادی مواجه شود. این چالش‌ها تا حدودی با حمایت تمامی نهادها و سازمان‌های درگیر در امر پژوهش و یک برنامه‌ریزی دقیق و مدون تعدیل خواهند شد. پیشنهادهایی که حاصل از مرور متون و مصاحبه‌اند و می‌توانند زمینه استفاده مجدد از داده‌های موجود را فراهم کنند، عبارت‌اند از: طراحی یک مخزن ملی باز به صورت یکپارچه و جامع به منظور میزبانی داده‌های پژوهشی در کشور؛ انتخاب یک سازمان با قدرت اجرایی بالا و ساختاری دقیق به عنوان سازمان متولی؛ تدوین سیاست‌هایی شفاف؛ تدارک زیرساخت‌های لازم فنی و غیرفنی؛ الزام مؤسسات پژوهشی کشور به ارائه داده‌های پژوهشی همراه با طرح‌های پژوهشی، پایان‌نامه‌ها و مقالات؛ در نظر گرفتن سازوکارهایی برای سنجش کیفیت داده‌ها؛ تعیین سطح دسترسی به داده‌ها برای افراد عضو و غیرعضو با مجوزهای خاص؛ اتخاذ تدابیر امنیتی؛ تدوین استانداردهای ابرداده‌ای خاص؛ برنامه‌ریزی برای تغییرات فرهنگی و اجتماعی؛ تشویق پژوهشگران و ایجاد انگیزه در آن‌ها به منظور استفاده مجدد از داده‌های موجود؛ برگزاری دوره‌های آموزشی رسمی و خاص برای آموزش استفاده مجدد.

### فهرست منابع

- استراوس، آنسلم، و جولیت کوربین. ۱۹۹۰. اصول روش تحقیق کیفی. ترجمه بیوک محمدی. ۱۳۸۷. تهران: پژوهشگاه علوم انسانی و مطالعات فرهنگی.
- افخمی، رضا. ۱۳۸۹. مدیریت چرخه تولید تا مصرف داده‌های پژوهشی: چالش‌ها و فرصت‌ها. همایش ملی مدیریت پژوهش و فناوری، دانشگاه امام صادق (ع). تهران.

پایگاه استنادی جهان اسلام. ۱۳۹۷. جایگاه ایران در تولید علم جهان ارتقا یافت، از اینترنت:

<https://isc.gov.ir/fa/news/995> (دسترسی در ۱۳۹۸/۰۱/۳۱)

رنجبری، فاطمه، نادر نقشینه، و محمدرضا اسمعیلی. ۱۳۹۷. بازمفهوم‌سازی سواد اطلاعاتی برای تیم‌های طراحی نرم‌افزار کتابخانه. *فصلنامه پردازش و مدیریت اطلاعات* ۳۳ (۴): ۱۶۵۹-۱۶۸۴.

فرجی سبکبار، حسنعلی، علی منصوریان، منصور رضاعلی، و محمدرضا آقاآزاداحمدچالی. ۱۳۹۲. ارائه مدل مفهومی برای مدیریت زیرساخت داده مکانی بخشی شهرداری. *پژوهش‌های جغرافیایی انسانی* ۴۵ (۱): ۲۳-۴۴.

فقیهی، ابوالحسن، و محسن عزیزاده. ۱۳۸۴. روایی در تحقیق. *مدیریت فرهنگ سازمانی* ۲ (۳): ۵-۱۹.

وزیری، اسماعیل، نادر نقشینه، و عبدالرضا نوروزی چالکی. ۱۳۹۶. اشتراک‌گذاری داده‌های پژوهشی رویکردی جهانی و ملی. *فصلنامه پژوهشنامه پردازش و مدیریت اطلاعات* ۳۳ (۳): ۱۰۱۳-۱۰۴۰.

\_\_\_\_\_، میترا دیلمقانی، و شهرام توفیقی. ۱۳۹۶. عوامل مؤثر بر اشتراک‌گذاری داده‌های پژوهشی. *مطالعات ملی کتابداری* ۲۸ (۳): ۱۲۳-۱۳۹.

\_\_\_\_\_ . ۱۳۹۶. چالش‌های اشتراک‌گذاری داده‌های پژوهشی. *فصلنامه مدیریت اطلاعات سلامت* ۷۰ (۲): ۷۹-۹۳.

## References

- AGU2013 .. AGU Publications Data Policy. <http://publications.agu.org/author-resource-enter/publication-policies/data-policy/>, 2014-04-14. (accessed Aug. 01, 2018).
- Ball, A., and M. Duke. 2015. How to track the impact of research data with metrics (DCC How-to Guides)? Edinburgh, Scotland. Retrieved from <http://www.dcc.ac.uk/resources/how-guides/track-data-impact-metrics#sthash.sTZOPKLq.dpuf> (accessed July 23, 2018).
- Bishop, L. 2014. Re-using qualitative data: A little evidence, on-going issues and modest reflections. *Studia Sociologica* 214 (3): 167-176.
- Bishop, L., and A. Kuula-Luumi. 2017. *Revisiting qualitative data reuse: A decade on*. Sage Open 7:(1) 1-15.
- Carlson, S., and B. Anderson. 2007. What Are Data? The Many Kinds of Data and Their Implications for Data Re-Use. *Journal of Computer-Mediated Communication* 12 (2): 635-651.
- Chao, T. C. 2011. Disciplinary reach: Investigating the impact of dataset reuse in the earth sciences. *Proceedings of the American Society for Information Science and Technology* 48 (1): 1-8.
- Coffey, lissa. 2016. Smart Data - What is it and How it is Different from Big Data? MIT Initiative on the digital economy. Ver1. <http://digitalcommunity.mit.edu/docs/DOC-1232> (accessed Nov. 10, 2017).
- Commission notice — Guidelines on recommended standard licences, datasets and charging for the reuse of document. *Official Journal of the European Union*, 2401-10 ..
- Curry, R. G. 2015. Beyond "data thrifting": An investigation of factors influencing research data reuse in the social sciences. Doctoral dissertation. Syracuse University.
- Curry, R. G. 2016. Factors influencing research datafire--use in Social Sciences: An exploratory study. *International Journal of Digital Curation* 11 (1): 96-117.
- \_\_\_\_\_, and J. Qin. 2014. *Towards a model for research data reuse behavior*. Proceedings of the Association for Information Science and Technology 51 (1): 1-4.

- Curty, R. G. 2016. "Factors influencing research datafire--use in Social Sciences: An exploratory study." *International Journal of Digital Curation* 11 (1):96-117.
- Curty, R. G., K. Crowston, A. Specht, B. W. Grant, and E. D. Dalton. 2017. Attitudes and norms affecting scientists' data reuse. *PLoS one* 12 (12): e0189288. <https://doi.org/10.1371/journal.pone.0189288> (accessed May 16, 2018)
- Darby, R., S. Lambert, B. Matthews, M. Wilson, K. Gitmans, S. Dallmeier-Tiessen, S. Mele, and J. Suhonen. 2012. Enabling scientific data sharing and re-use. IEEE 8th International Conference on E-Science, Chicago, IL, USA, 2012.
- Dobre, C. Sc Xhafa. F. 2014. Intelligent sendees for big data science. *Future Generation Computer Systems* 37: 267–281.
- Duke, C. S., and J. H. Porter. 2013. The ethics of data sharing and reuse in biology. *BioScience* 63 (6): 483-489.
- El Hani, M. A., L. Rivest, and R. Maranzana. 2012. *Product data reuse in product development*. In IFIP International Conference on Product Lifecycle Management (pp. 243-256). Springer, Berlin, Heidelberg.
- Faniel, I. M. and T. E. Jacobsen. 2010. Reusing scientific data: How earthquake engineering researchers assess the reusability of colleagues' data. *Computer Supported Cooperative Work (CSCW)* 19 (3-4): 355-375.
- Faniel, I. M., and A. Zimmerman. 2011. Beyond the data deluge: a research agenda for large-scale data sharing and reuse. *International Journal of Digital Curation* 6 (1): 58-69.
- Faniel, I. M., J. Barrera-Gomez, A. Kriesberg, and E. Yakel. 2013. A comparative study of data reuse among quantitative social scientists and archaeologists. *iConference 2013 Proceedings*. Illinois (797-800).
- Fear, K. M. 2013. Measuring and Anticipating the Impact of Data Reuse. PhD Dissertation. University of Michigan.
- Fecher, B., S. Friesike, and M. Hebing. 2015. What drives academic data sharing? *PLoS one* 10 (2). <https://doi.org/10.1371/journal.pone.0118053> (accessed Nov. 10, 2017)
- Federer, L. M., Y. L. Lu, D. J. Joubert, J. Welsh, and B. Brandys. 2015. Biomedical data sharing and reuse. *PLoS one* 10 (6): e0129506. <https://doi.org/10.1371/journal.pone.0129506> (accessed May 11, 2018)
- Gantz, J., & D. Reinsel. 2012. The Digital Universe in 2020: Big data, bigger digital shadows, and biggest growth in the Far East. IDC – EMC Corporation. Online Available at <http://www.emc.com/collateral/analyst-reports/idc-the-digital-universe-in2020.pdf> (Accessed January 16, 2016)
- German, L. 2015. *Academic research data re-usage in a digital age: modelling best practice*. Doctoral dissertation. University of Southampton.
- Grinyer, A. 2009. The ethics of the secondary analysis and further use of qualitative data. *Social Research Update* 56 (4): 1-4.
- Herndon, T., M. Ash, and R. Pollin. 2013. Does High Public Debt Consistently Stifle Economic Growth? A Critique of Reinhart and Rogoff. Retrieved from <http://www.peri.umass.edu/236/hash/31e2ff374b6377b2ddec04deaa6388b1/publication/566/> (accessed Jun. 21, 2018)
- Holub, P., F. Kohlmayer, F. Prasser, M. T. Mayrhofer, I. Schlünder, G. M. Martin, ... & D. Strapagiel. 2018. Enhancing reuse of data and biological material in medical research: From FAIR to FAIR-health. *Biopreservation and biobanking* 16 (2): 97-105.
- Kim, Y., and S. Nah. 2018. Internet researchers' data sharing behaviors: An integration of data reuse experience, attitudinal beliefs, social norms, and resource factors. *Online Information Review* 42 (1): 124-142.

- Kim, K. K., D. K. Browe, H. C. Logan, R. Holm, L. Hack, & L. Ohno-Machado. 2014. Data governance requirements for distributed clinical research networks: triangulating perspectives of diverse stakeholders. *Journal of the American Medical Informatics Association* 21 (4): 714-719.
- McAllister, J. W. 2018. Scientists' Reuse of Old Empirical Data: Epistemological Aspects. *Philosophy of Science* 85 (5): 755-766.
- MacEachen, E., J. Clarke, R. L. Franche, and E. Irvin. 2006. Systematic review of the qualitative literature on return to work after injury. *Scandinavian Journal of Work Environment and Health* 32 (4): 257-269.
- Mannheimer, S., L. B. Sterman, and S. Borda. 2016. Discovery and Reuse of Open Datasets: An Exploratory Study. *Journal of eScience Librarianship* 5 (1): 5.
- Mooney, H. and M. P. Newton. 2012. The anatomy of a data citation: Discovery, reuse, and credit. *Journal of Librarianship & Scholarly Communication* 1 (1): 1-6.
- Moore, N. 2006. The contexts of context: Broadening perspectives in the (re)use of qualitative data. *Methodological Innovations Online* 1 (2): 21-32.
- Mueller, S. G., M. W. Weiner, L. J. Thal, R. C. Petersen, C. R. Jack, W. Jagust, ... & L. Beckett. 2005. Ways toward an early diagnosis in Alzheimer's disease: The Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimer's & Dementia* 1 (1): 55-66.
- Murillo, A. P. 2010. An investigation of selected cyberinfrastructure and interoperability elements: Data Sharing and reuse in the sciences. *Bulletin of IEEE Technical Committee on Digital Libraries* 10 (1): 35-42.
- Murillo, A. P. 2014. Examining data sharing and data reuse in the DataONE environment. *Proceedings of the American Society for Information Science and Technology* 51 (1): 1-5.
- Park, H., and D. Wolfram. 2017. An examination of research data sharing and re-use: implications for data citation practice. *Scientometrics* 111 (1): 443-461.
- Pasquetto, I., B. Randles, and C. Borgman. 2017. On the reuse of scientific data. *Data Science Journal* 16 (8). doi:10.5334/dsj-2017-008.
- PLOS. 2014. PLOS Editorial and Publishing Policies. <http://www.plosone.org/static/policies.action>, 2014-04-10. (accessed June 21, 2018)
- Prince, K., M. Jones, A. Blackwell, A. Simpson, S. Meakins, & A. Vuylsteke. 2018. Barriers to the secondary use of data in critical care. *Journal of the Intensive Care Society* 19 (2): 127-131.
- Rolland, B., & C. P. Lee. 2013. Beyond trust and reliability: Reusing data in collaborative cancer epidemiology research. In CSCW '13 Proceedings of the 2013 Conference on Computer Supported Cooperative Work (pp. 435-444). New York: ACM
- Sands, A., C. L. Borgman, L. Wynholds, & S. Traweek. 2012. Follow the data: How astronomers use and reuse data. *Proceedings of the American Society for Information Science and Technology* 49 (1): 1-3.
- Schöpfel, J., P. Južnič, H. Prost, C. Malleret, A. Češarek, & T. Koler-Povh. 2015. Dissertations and data. In GL17 International Conference on Grey Literature, Amsterdam.
- Spencer L, J. Ritchie, J. Lewis, and L. Dillon. 2003. *Quality in qualitative evaluation: a framework for assessing research evidence*. London: Government Chief Social Researcher's Office. 2 (i-98).
- Sposito, F. A. 2017. What do Data Curators Care About? Data Quality, User Trust, and the Data Reuse Plan. *Paper Presented at IFLA WLIC 2017*. Available online: <http://library.ifla.org/1797/> (accessed June 29, 2018)
- Tenopir, C., E. D. Dalton, S. Allard, M. Frame, I. Pjesivac, B. Birch, ... & K. Dorsett. 2015. Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *PLoS one*, 10 (8). <https://doi.org/10.1371/journal.pone.0134826>

- Tenopir C, S. Allard, K. Douglass, A. U. Aydinoglu, L. Wu, E. Read, ... . 2011. Data sharing by scientists: Practices and perceptions. *PLoS one* 6 (6). <https://doi.org/10.1371/journal.pone.0021101> (accessed Aug. 18, 2018)
- Thanos, C. 2017. Research Data Reusability: Conceptual Foundations, Barriers and Enabling Technologies, Publications, MDPI. *Open Access Journal* 5 (1): 1-19.
- Vision T. J., & H. A. Piwovar. 2013. Data reuse and scholarly reward: understanding practice and building infrastructure. *PeerJ PrePrints* 1:e14v1 <https://doi.org/10.7287/peerj.preprints.14v1> (accessed June 21, 2018)
- Wallis J. C., E. Rolando, & C. L. Borgman. 2013. If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology. *PLoS ONE* 8 (7): e67332. <https://doi.org/10.1371/journal.pone.0067332> (accessed Aug. 6, 2018)
- Ward, C., J. Warren, and R. Hanisch. 2014. Making materials science data more valuable research products. *Integrating Materials and Manufacturing Innovation* 3 (1): 22.
- Wilson, J. A., L. Martinez-Urbe, M. A. Fraser, and P. Jeffreys. 2011. An Institutional Approach to Developing Research Data Management Infrastructure. *International Journal Digital Curation* 6 (2): 274-287.
- Yakel, E., I. M. Faniel, A. Kriesberg, and A. Yoon. 2013. Trust in digital repositories. *International Journal of Digital Curation* 8 (1): 143-156.
- Yardley, S. J., K. M. Watts, J. Pearson, & J. C. Richardson. 2014. Ethical issues in the reuse of qualitative data: Perspectives from literature, practice, and participants. *Qualitative Health Research* 24 (1): 102-113.
- Yoon, A. 2015. Data reuse and users' trust judgments: Toward trusted data curation. Doctoral dissertation, The University of North Carolina at Chapel Hill.
- Yoon, A. 2016. Red flags in data: Learning from failed data reuse experiences. In Proceedings of the 79th ASISandT Annual Meeting: Creating Knowledge, Enhancing Lives through Information and Technology (p. 126). American Society for Information Science.
- \_\_\_\_\_, and Y. Kim. 2017. Social scientists' data reuse behaviors: Exploring the roles of attitudinal beliefs, attitudes, norms. *Library and Information Science Research* 39 (3): 224-233.
- Zimmerman, A. S. 2008. New knowledge from old data: The role of standards in the sharing and reuse of ecological data. *Science, Technology, and Human Values* 33 (5): 631-652.

### آمنه سلیمانی

متولد ۱۳۵۹، دانشجوی دکتری علم اطلاعات و دانش‌شناسی دانشگاه تهران است.

نظام‌های ذخیره و بازیابی اطلاعات، وب معنایی، وب نامرئی، مدیریت داده‌های پژوهشی، اشتراک‌گذاری داده و استفاده مجدد از داده‌های پژوهشی از علائق پژوهشی وی است.



#### فاطمه فهیم‌نیا

متولد ۱۳۴۵، دارای مدرک دکتری علم اطلاعات و دانش‌شناسی دانشگاه تهران است. ایشان هم‌اکنون دانشیار دانشکده مدیریت دانشگاه تهران است. مطالعات اقتصادی، مدیریت اطلاعات، آموزش علم اطلاعات، فراهم‌آوری و مطالعات کودکان از جمله علایق پژوهشی وی است.



#### نادر نقشینه

متولد سال ۱۳۴۰، دارای مدرک دکتری علم اطلاعات و دانش‌شناسی دانشگاه تهران است. ایشان هم‌اکنون استادیار گروه علم اطلاعات و دانش‌شناسی دانشگاه تهران است. مسائل مربوط به زندگی دیجیتال، حفاظت، رخنه دیجیتال و Digital emergence از علایق پژوهشی ایشان است.



#### عادل سلیمانی‌نژاد

دارای مدرک دکتری علم اطلاعات و دانش‌شناسی است. ایشان هم‌اکنون استادیار بخش علم اطلاعات و دانش‌شناسی دانشگاه شهید باهنر کرمان است. داده‌کاوی، متن‌کاوی، نظام‌های ذخیره و بازیابی اطلاعات، علم‌سنجی و مدیریت داده‌های پژوهشی از جمله علایق پژوهشی ایشان است.

