



معرفی دیتاورس

مآئده کباری، Maede.kobari@gmail.com
بانظارت: مبارکه آجدانی، mobarakehajdani@gmail.com

مقدمه

پروژه دیتاورس^۱ یک برنامه وب منبع آزاد است که به اشتراک گذاری، حفظ، استناد، کاوش و تحلیل داده‌های پژوهشی می‌پردازد. این پروژه امکان دسترسی به داده‌ها را برای دیگران فراهم می‌کند و به پژوهشگران این امکان را می‌دهد که کارهای دیگران را به راحتی بازتولید کنند. پژوهشگران، نشریات، نویسندگان داده، ناشران، توزیع کنندگان داده و مؤسسه‌های وابسته همگی اعتبار علمی و دیده شدن در وب را دریافت می‌کنند. این پلتفرم توسط پروژه دیتاورس توسعه یافته و بیش از یک دهه است که برای ساخت مخازن داده در سراسر جهان استفاده می‌شود (Brooke, 2020). دیتاورس بر پایه اصول علم آزاد^۲ بنا شده و هدف آن افزایش دسترسی پذیری، قابلیت استفاده مجدد و اعتبار داده‌های پژوهشی است (King, 2007). این سیستم توسط مؤسسه علوم اجتماعی کمی دانشگاه هاروارد پشتیبانی می‌شود و از یک رویکرد توسعه ترکیبی بهره می‌برد که شامل تیم اصلی توسعه و جامعه جهانی مشارکت کنندگان است (Brooke, 2020). دیتاورس به‌ویژه برای داده‌های کوچک و متوسط^۴ مفید است و به افزایش شفافیت^۵ داده‌ها کمک می‌کند (Conzett, 2020). پروژه دیتاورس در طول زمان به طور قابل توجهی رشد کرده و اکنون یک پروژه همکاری بین‌المللی بزرگ است.

اهداف

- اهداف استراتژیک پروژه دیتاورس راهنمای نقشه راه انتشار همکاری‌های دیتاورس با جامعه و خدماتی است که ارائه داده می‌شود. در حال حاضر، اهداف پروژه عبارت‌اند از:
- رشد جامعه دیتاورس
 - توانمندسازی جامعه منبع باز برای کاوش و پیاده‌سازی برنامه‌ها، ابزارها و خدمات جدید دیتاورس
 - توسعه قابلیت مدیریت داده‌های حساس و داده‌های کلان
 - گسترش ویژگی‌های داده و فراداده برای رشته‌های موجود و جدید

^۱ Dataverse Project

^۲ Open Science

^۳ Institute for Quantitative Social Science at Harvard University

^۴ long tail of research

^۵ Findable, Accessible, Interoperable, Reusable (FAIRness)



• گسترش ویژگی‌های آرشیوی و حفاظتی

• افزایش تعامل پذیری از طریق پیاده‌سازی استانداردها

• افزایش مشارکت‌ها از جامعه توسعه منبع‌باز

• بهبود تجربه کاربری دیتاورس

همکاری

مؤسسه علوم اجتماعی کمی^۶ با کتابخانه دانشگاه هاروارد و سازمان فناوری اطلاعات دانشگاه هاروارد همکاری می‌کند تا نصب واسپارگاه دیتاورس هاروارد را به طور عمومی در دسترس پژوهشگران و جمع‌آوردگان داده از سراسر رشته‌ها قرار دهد. آی کیو اس اس رهبری توسعه نرم‌افزار منبع آزاد پروژه دیتاورس را بر عهده دارد و با برنامه کمک به داده‌های آزاد در هاروارد، همکاری با کتابخانه هاروارد، دفتر ارتباطات علمی و آی کیو اس اس پشتیبانی کاربری ارائه می‌دهد. خدمات فناوری کتابخانه در اچ یو آی تی^۷، پشتیبانی از میزبانی و پشتیبان‌گیری واسپارگاه دیتاورس هاروارد را فراهم می‌کند. پروژه دیتاورس همچنین با کنسرسیوم جامعه جهانی دیتاورس همکاری می‌کند تا نیازهای جامعه را پشتیبانی کند.

ویژگی‌های کلیدی

۱. ذخیره‌سازی و حفظ داده‌ها

دیتاورس امکان ذخیره‌سازی داده‌های خام، مشتق‌شده و فراداده‌ها را فراهم می‌کند و از حفظ بلندمدت داده‌ها با استفاده از استانداردهایی مانند دی او آی^۸ پشتیبانی می‌کند (Sweeney & Crosas, 2013) (King, 2007). این پلتفرم برای داده‌های بزرگ^۹ و داده‌های حساس^{۱۰} گسترش یافته است (Brooke, 2020). همچنین، قابلیت ادغام با سیستم‌های حفظ ماندگار^{۱۱} و سی کی اس اس^{۱۲} برای ذخیره‌سازی دیجیتال بلندمدت وجود دارد، در مورد داده‌های زیستی، دیتاورس برای حفظ ساختار فایل سیستم و دسترسی مستقیم به محاسبات^{۱۲} گسترش یافته است (McKinney et al, 2016).

آز مایستگاه علم آزاد

⁶ IQSS

⁷ HUIT

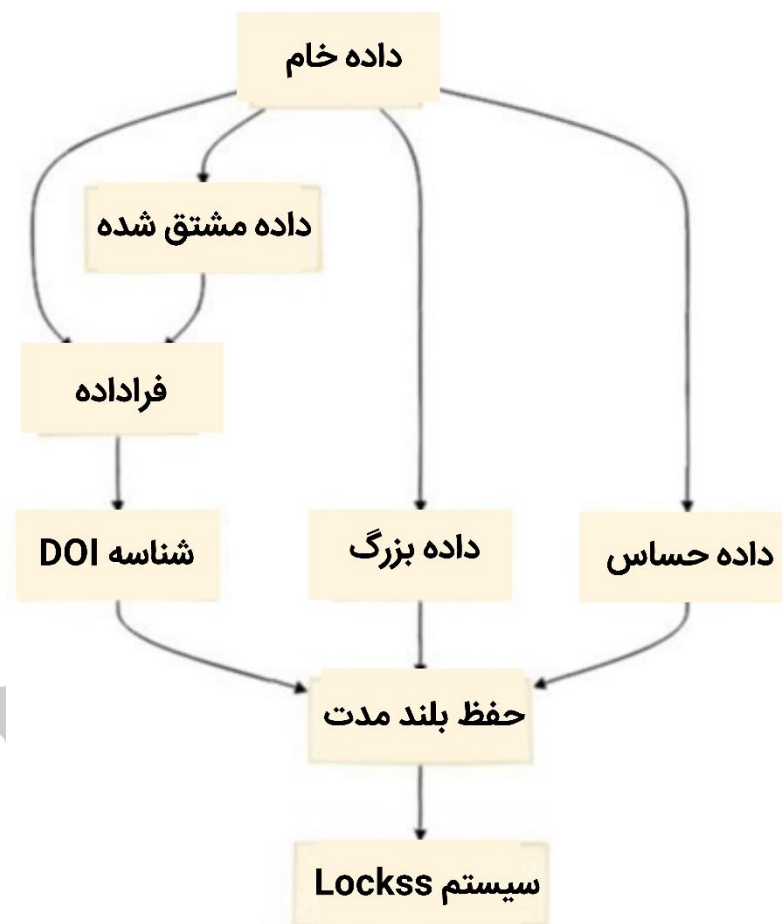
⁸ DOI (Digital Object Identifier)

⁹ large data support

¹⁰ sensitive data support

¹¹ Lots of Copies Keep Stuff Safe (LOCKSS)

¹² in-place computation



نمودار ۱- اجزای کلیدی (SCINITO AI, 2025)

۲. پشتیبانی از اصول فیر^{۱۳}

دیتاورس به طور خاص برای افزایش شفافیت طراحی شده و ویژگی‌هایی مانند جستجوی ساخت یافته، دسترسی عمومی یا خصوصی، و ادغام با فرمت‌های استاندارد برای داده‌های تحلیلی ارائه می‌دهد (Conzett, 2020). کاربران می‌توانند داده‌ها را با مجوزهای عمومی منتشر کنند و فراداده‌ها را با اسکیمای^{۱۴}هایی مانند دیتا سایت^{۱۵} سازگار سازند این پلتفرم برای داده‌های ناهمگن مانند داده‌های کووید ۱۹ (شامل موارد روزانه، رسانه‌های اجتماعی، تحرک جمعیتی و غیره) مناسب است و امکان ادغام بر پایه جی‌آی‌اس^{۱۶} را فراهم می‌کند (McKinney et al., 2016, Hu et al, 2020).

¹³ FAIR Compliance and Accessibility

¹⁴ schema

¹⁵ DataCite

¹⁶ GIS

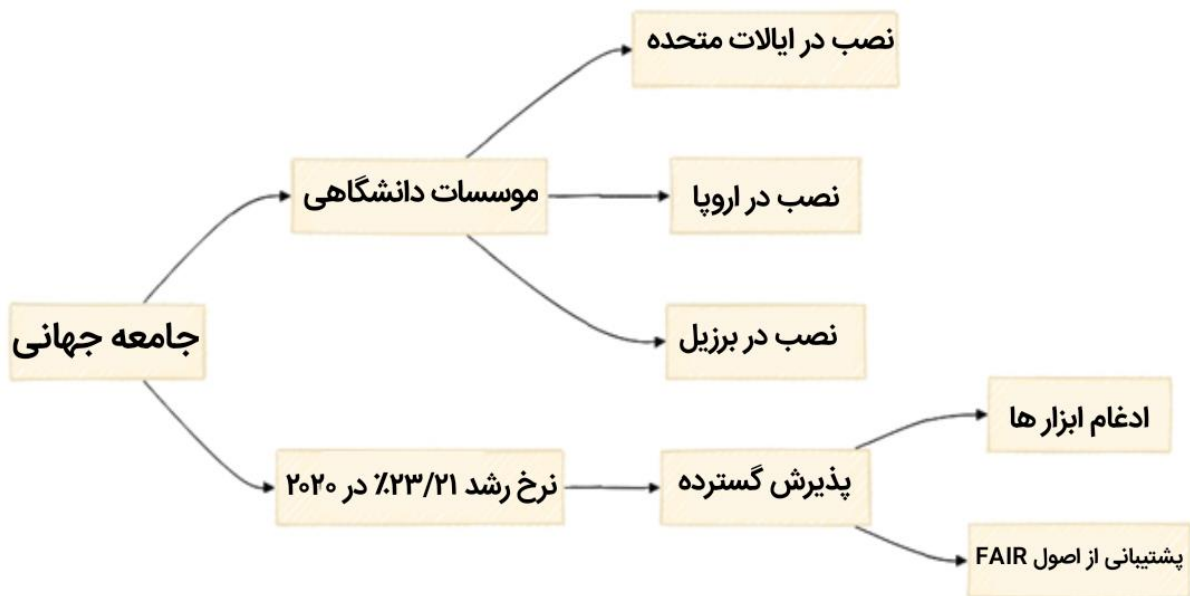


۳. مدیریت فراداده و استانداردسازی

دیتاورس از چارچوب فراداده غنی پشتیبانی می‌کند، با بیش از ۷۵ المان فراداده در ۱۳ دسته‌بندی (مانند فراداده، شناسه‌ها، احراز هویت، دسترسی به داده، حفظ و پایداری است. در مقایسه با پلتفرم‌های دیگر دیتاورس ویژگی‌های بهتری در مدیریت فراداده و ادغام با استانداردهای جهانی ارائه می‌دهد (Mahato & Gajbe, 2018).

۴. حمایت از جامعه و گسترش جهانی

دیتاورس توسط جامعه جهانی توسعه می‌یابد و بیش از ۸۰٪ نصب‌های آن در پنج سال اخیر (۲۰۱۶-۲۰۲۰) رخ داده است (Chen et al, 2022). تا سال ۲۰۲۰، ایالات متحده با ۱۳ نصب بیشترین تعداد را دارد و اروپا با ۲۵ نصب پیشتاز است (Jamwal & Kaur, 2021). کتابخانه‌های دانشگاهی نقش کلیدی در پذیرش دیتاورس دارند و ۱۸ دانشگاه از ۲۹ دانشگاه برتر جهانی خدمات داده تحقیقاتی بر پایه آن ارائه می‌دهند (Chen et al, 2022). در برزیل، دیتاورس برای ذخیره اسناد تحقیقاتی و مجلات الکترونیکی استفاده می‌شود (Unknown Author, 2017). رشد جهانی آن ۲۳.۲۱٪ در سال ۲۰۲۰ بوده است (Jamwal & Kaur, 2021).



نمودار ۲- گسترش جهانی دیتاورس (SCINITO AI, 2025)

این نمودار گسترش جهانی دیتاورس و نقش جامعه، مؤسسات دانشگاهی، نصب‌ها در ایالات متحده، اروپا و برزیل، و نرخ رشد را به تصویر می‌کشد.



۵. ویژگی‌ها و کاربردهای دیگر

- انتشار و اشتراک‌گذاری: امکان انتشار داده‌ها با دی.او.آی پایدار و ارجاع‌دهی برای افزایش اعتبار علمی (Sweeney & Crosas, 2013) (King, 2007).
- پشتیبانی از داده‌های خاص: مانند داده‌های جغرافیایی و داده‌های کووید ۱۹ برای تحلیل الگوهای انتقال (Brooke, 2020) (Hu et al., 2020) (Hu et al., 2020).
- مقایسه با رقبا: در مقایسه با سی کی ای ان^{۱۷}، دیتاورس در قابلیت‌های حفظ، ادغام و دسترس‌پذیری برتر است، هرچند سی کی ای ان در سادگی و نصب قوی‌تر عمل می‌کند (Mahato & Gajbe, 2018).
- پایداری و اعتماد: دیتاورس به‌عنوان واسپارگاه قابل اعتماد عمل می‌کند و با واسپارگاه آر. دی. ای^{۱۸} سازگار است (Conzett, 2020) (Kim, 2018).

برای واسپاری داده‌ها به واسپارگاه دیتاورس هاروارد باید:

۱. به آدرس <http://dataverse.harvard.edu> بروید.
۲. گزینه "ثبت‌نام" را انتخاب کنید تا یک حساب کاربری ایجاد کنید این کار کمتر از یک دقیقه طول می‌کشد.
۳. گزینه "اضافه کردن داده" را انتخاب کنید.
۴. هر نوع داده کمی یا کیفی در هر فرمت را بارگذاری کنید.
۵. زمانی که آماده‌اید داده‌های خود را با دیگران به اشتراک بگذارید، روی "انتشار" کلیک کنید و یک ارجاع علمی رسمی برای داده‌های خود دریافت خواهید کرد.

قوانین و شرایط کلی استفاده از وب‌گاه دیتاورس دانشگاه هاروارد

در استفاده از خدمات، شما تأیید می‌کنید که:

۱. قوانین جامعه وب‌گاه را مطالعه و درک کرده‌اید.
 ۲. سیاست حفظ حریم خصوصی وب‌گاه را مطالعه و درک کرده‌اید.
- شما موظف به رعایت تمامی قوانین و مقررات محلی، ایالتی، ملی و بین‌المللی در استفاده از خدمات هستید و استفاده شما از خدمات نباید:
- حقوق مالکیت معنوی دیگران، از جمله حق ثبت اختراع، علامت تجاری، حق نشر و حق حریم خصوصی را نقض کند.

• محتوا نباید غیرقانونی، تهدیدآمیز، آزاردهنده، افتراآمیز، فریبنده، کلاهبردانه یا به‌نوعی توهین‌آمیز باشد.

• شامل هرگونه هرزنامه باشد.

• موجب ایجاد یا استفاده از ویروس‌های نرم‌افزاری یا هر کد، فایل یا برنامه‌ای شود که به‌منظور مختل کردن

¹⁷ Comprehensive Knowledge Archive Network (CKAN)

¹⁸ RDA Repository



وزارت علوم، تحقیقات و فناوری

پژوهشگاه علوم و فناوری اطلاعات ایران
(ایرانداک)

عملکرد صحیح نرم‌افزار یا سخت‌افزار طراحی شده باشد یا به اطلاعات سیستم‌های دانشگاه هاروارد یا اشخاص ثالث آسیب برساند یا دسترسی غیرمجاز به آن‌ها پیدا کند.

به طور خلاصه، کاربران باید از قوانین و سیاست‌های سایت پیروی کنند و از هر گونه فعالیت غیرقانونی یا مخرب خودداری کنند.

محدودیت‌ها

۱. مطابقت با قوانین: هنگام ارسال محتوا به وب‌گاه، باید اطمینان حاصل کنید که محتوا با شرایط استفاده مطابقت دارد. در صورت عدم تطابق، دیتابیس هاروارد حق دارد به‌تنهایی اقدام به حذف محتوا کند.
۲. مسئولیت قانونی: شما به‌عنوان کاربر مسئولیت قانونی و مالی تمامی خسارات ناشی از نقض این توافق‌نامه را بر عهده دارید.
۳. تضمین‌ها و نمایندگی‌ها: با ارسال محتوا، شما تضمین می‌کنید که:
 - محتوا حقوق مالکیت معنوی دیگران را نقض نمی‌کند.
 - محتوا قوانین را نقض نمی‌کند.
 - در صورت بروز هر گونه مشکل پس از ارسال، به‌سرعت دیتابیس هاروارد و مدیران مربوطه را مطلع خواهید کرد.
 - محتوا شامل ویروس‌های نرم‌افزاری یا کدهای مخرب نیست.
 - تمامی مجوزهای لازم برای انتشار محتوا را دریافت کرده‌اید.
 - محتوا باید فاقد هر گونه اطلاعات شناسایی باشد تا شناسایی مجدد افراد ممکن نباشد.
۴. استثنای برای اطلاعات شناسایی: اطلاعات شناسایی تنها در شرایط خاصی مجاز است، از جمله:
 - اطلاعات قبلاً به عموم منتشر شده باشد.
 - اطلاعات مربوط به شخصیت‌های عمومی باشد.
 - زمان کافی از جمع‌آوری اطلاعات گذشته باشد.
 - تمامی افراد شناسایی شده رضایت صریح برای انتشار داده‌ها را داده باشند.
 - تمامی افراد شناسایی شده فوت کرده باشند و هیچ قانونی مانع انتشار داده‌ها نشود.

مجوزها و اجازه‌ها

- شما به دیتابیس هاروارد تمامی مجوزهای لازم برای آرشیو، حفظ و دسترسی به محتوای ارسال‌شده را اعطا می‌کنید. این شامل اجازه برای:
- انتشار مجدد محتوای شما در فرمت‌های مختلف.



• تبلیغ و معرفی محتوا در هر نوع اطلاع‌رسانی برای دیتاورس‌ها روارد و دیتا پس^{۱۹} است.

۱. عدم مسئولیت دیتاورس‌ها روارد:

- دیتاورس‌ها روارد هیچ‌گونه مسئولیتی در قبال قانونی‌بودن، قابلیت اجرایی، دقت یا مطلوبیت توافق‌نامه استفاده از داده‌ها ندارد.

- این وب‌گاه طرف قرارداد توافق‌نامه استفاده از داده‌ها نیست و نمی‌تواند به‌خاطر شرایط موجود در آن پاسخگو باشد.

- دیتاورس‌ها روارد هیچ الزامی برای کمک یا حمایت از هر یک از طرفین توافق‌نامه در اجرای شرایط آن ندارد.
۲. مسئولیت کاربران:

- کاربران مسئول تعیین، حفظ و اجرای شرایط مجوزی هستند که برای دسترسی و استفاده از محتوای خود می‌خواهند استفاده کنند.

- دیتاورس‌ها روارد مسئولیتی در قبال نادرستی‌ها، شرایط غیرقابل اجرا یا مسئولیت‌هایی که ممکن است از انتخاب گزینه‌های موجود در این توافق‌نامه ناشی شود، ندارد و همچنین مسئولیتی در قبال بررسی یا اجرای انطباق با هر یک از شرایطی که کاربر ممکن است انتخاب کند، ندارد.

۳. دانلود ارسال‌های کاربران:

- وب‌گاه تلاش می‌کند تا دسترسی باز به داده‌ها را برای کاربران فراهم کند تا بتوانند آن‌ها را دانلود کنند، مشروط به محدودیت‌های تعیین شده توسط واسطاران و هرگونه محدودیت قانونی قابل اعمال.

منابع

1. Brooke, D. (2020). Community built infrastructure: The Dataverse project. N/A. <https://doi.org/10.5194/equosphere-egu2020-12006>
2. Conzett, P. (2020). DataverseNO: A national, generic repository and its contribution to the increased FAIRness of data from the long tail of research. *Ravnetrykk*. <https://doi.org/10.7557/15.5514>
3. King, G. (2007). An introduction to the Dataverse network as an infrastructure for data sharing. *Sociological Methods & Research*. <https://doi.org/10.1177/0049124107306660>
4. Mahato, S. S., & Gajbe, S. B. (2018). A comparative study of open source data repository software: Dataverse and CKAN. *Library Herald*. <https://doi.org/10.5958/0976-2469.2018.00005.2>
5. Sweeney, L., & Crosas, M. (2013). An open science platform for the next generation of data. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2302084>
6. Chen, H.-L., Chiu, T.-H., & Cline, E. (2022). Academic libraries and research data management: A case study of Dataverse global adoption. *Information Discovery and Delivery*. <https://doi.org/10.1108/idd-04-2022-0028>
7. Kim, S. (2018). Functional requirements for research data repositories. *International Journal of Knowledge Content Development & Technology*. (DOI not available)
8. McKinney, B., Meyer, P., Crosas, M., & Sliz, P. (2016). Extension of research data repository system to support direct compute access to biomedical datasets: Enhancing

¹⁹ Data Preservation Alliance for the Social Sciences (Data-PASS)



وزارت علوم، تحقیقات و فناوری

پژوهشگاه علوم و فناوری اطلاعات ایران
(ایرانداک)

- Dataverse to support large datasets. *Annals of the New York Academy of Sciences*. <https://doi.org/10.1111/nyas.13272>
9. Jamwal, V., & Kaur, S. (2021). Global presence of open-source research data management platform for libraries: The Dataverse project. *Library Hi Tech News*. <https://doi.org/10.1108/lhtn-10-2021-0066>
 10. Giess, T., Itzigehl, S., Range, J., Schömig, R., Bruckner, J. R., & Pleiss, J. (2023). FAIR and scalable management of small-angle X-ray scattering data. *Journal of Applied Crystallography*. <https://doi.org/10.1107/s1600576723001577>
 11. Hu, T., Guan, W. W., Zhu, X., Shao, Y., Liu, L., Du, J., Liu, H., Zhou, H., Wang, J., She, B., Zhang, L., Li, Z., Wang, P., Tang, Y., Hou, R., Li, Y., Sha, D., Yang, Y., Lewis, B., Kakkar, D., & Bao, S. (2020). Building an open resources repository for Covid-19 research. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3587704>
 12. Scientific data management on a Dataverse network at IBICT. (2017). N/A. <https://doi.org/10.14324/000.learn.16>

