

THỰC TIỄN QUẢN LÝ DỮ LIỆU NGHIÊN CỨU CỦA GIẢNG VIÊN TẠI TRƯỜNG ĐẠI HỌC KHOA HỌC XÃ HỘI VÀ NHÂN VĂN, ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH

TS Ninh Thị Kim Thoa, TS Nguyễn Thị Hương
Trường Đại học KHXH&NV, ĐHQG Tp. Hồ Chí Minh

Tóm tắt: Bài viết tìm hiểu thực trạng quản lý dữ liệu nghiên cứu của giảng viên Trường Đại học KHXH&NV-ĐHQG Tp. Hồ Chí Minh. Nghiên cứu đã tiến hành khảo sát 151 giảng viên và phỏng vấn 06 giảng viên cơ hữu đang công tác tại Trường. Kết quả nghiên cứu cho thấy sự đa dạng và đặc thù của dữ liệu nghiên cứu thuộc các lĩnh vực khoa học xã hội và nhân văn được tạo lập hoặc thu thập. Nghiên cứu cũng khám phá những mặt mạnh và một số hạn chế trong lưu trữ, bảo mật, quản lý và chia sẻ dữ liệu nghiên cứu. Kết quả nghiên cứu đã cung cấp cơ sở thực tiễn cho các bên liên quan trong việc xây dựng chính sách, cơ sở hạ tầng công nghệ và các dịch vụ phù hợp cho quản lý dữ liệu nghiên cứu của trường đại học một cách hệ thống và có chiến lược, đồng thời giúp giảng viên nâng cao kiến thức và kỹ năng liên quan.

Từ khóa: Dữ liệu nghiên cứu; quản lý dữ liệu nghiên cứu; giảng viên đại học.

RESEARCH DATA MANAGEMENT PRACTICES OF LECTURERS AT THE UNIVERSITY OF SOCIAL SCIENCES AND HUMANITIES, VIETNAM NATIONAL UNIVERSITY, HO CHI MINH CITY

Abstract: The article investigates the current state of research data management by faculty members at the University of Social Sciences and Humanities - Vietnam National University in Ho Chi Minh City. The study used a mixed method to conduct a survey of 151 faculty members and to interview 06 full-time faculty members currently working at the university. The research results reveal the diversity and specificity of research data generated in the fields of social sciences and humanities. The study also explores strengths and some limitations in the process of storage, security, management, and sharing of research data. The findings provide practical groundwork for stakeholders in developing policies, technological infrastructure, and appropriate services for the systematic and strategic management of the university's research data. Additionally, it aids faculty members in enhancing their knowledge and skills related to research data management.

Keywords: Research data; research data management; university faculty members.

GIỚI THIỆU

Quản lý dữ liệu nghiên cứu (QLDLNC) đang ngày càng đóng một vai trò quan trọng trong giáo dục đại học. Hoạt động QLDLNC sẽ mang đến sự đảm bảo tính toàn vẹn, khả năng truy cập, tuổi thọ của DLNC, giúp duy trì độ tin cậy và khả năng tái tạo của các sản phẩm trong quá trình nghiên cứu. Quản lý DLNC nhìn chung được hiểu là một quá trình, bao gồm việc thiết kế, tạo lập dữ liệu, lưu trữ, bảo mật, bảo quản, truy xuất, chia sẻ và tái sử dụng [1]. Ngoài ra, QLDLNC cũng được xem xét dưới các góc độ liên quan đến khả năng kỹ thuật, những vấn đề về đạo đức, pháp lý và khung quản trị. Hiện nay, trong môi trường giảng dạy và nghiên cứu ở bậc đại học, việc

tạo lập, thu thập và sử dụng DLNC là một phần không thể thiếu của các hoạt động học thuật. Việc QLDLNC hiệu quả đang trở thành một khía cạnh quan trọng đối với các nhà nghiên cứu và giảng viên (sau đây gọi chung là giảng viên/GV) bởi nó có thể giúp nâng cao chất lượng và tính ứng dụng của nội dung giảng dạy, tạo ra một hệ thống linh hoạt nhằm tối ưu hóa quá trình nghiên cứu cũng như chia sẻ kết quả nghiên cứu một cách hiệu quả.

Mục đích chính của bài viết này là đánh giá thực trạng hoạt động QLDLNC từ góc độ nhận thức của GV tại Trường Đại học KHXH&NV (ĐHKHXH&NV), ĐHQG Tp. Hồ Chí Minh (ĐHQG-HCM). Các mục tiêu cụ thể được xác định bao gồm:

- Xác định loại DLNC được thu thập hoặc tạo lập bởi GV;
- Xác định các cơ sở hạ tầng khác nhau được GV sử dụng để lưu trữ DLNC;
- Xác định thực trạng hoạt động bảo quản và quản lý dữ liệu;
- Xác định thực trạng hoạt động chia sẻ DLNC.

Việc đánh giá thực trạng hoạt động QLDLNC của GV sẽ góp phần xác định mức độ sẵn sàng trong việc áp dụng các phương pháp quản lý phù hợp dưới góc độ tổ chức, từ đó có những chiến lược chuẩn bị cho quá trình thiết kế và triển khai các dịch vụ hỗ trợ QLDLNC phù hợp cho các bên liên quan.

1. TỔNG QUAN TÀI LIỆU

Dữ liệu nghiên cứu đóng vai trò quan trọng trong quá trình phát triển tri thức và nâng cao chất lượng nghiên cứu. Đối với GV, QLDLNC đòi hỏi cần có sự quan tâm đặc biệt để đảm bảo tính minh bạch, cho phép bảo quản lâu dài và tái sử dụng hiệu quả. Tổng quan tài liệu cho thấy đã có nhiều nghiên cứu trên thế giới về QLDLNC với các hướng tiếp cận từ nhiều khía cạnh khác nhau, bao gồm quan điểm và nhận thức của nhà nghiên cứu, nhà khoa học, cũng như quan điểm của các thư viện đại học nơi cung cấp dịch vụ QLDLNC. Trong đó, thực tiễn QLDLNC của GV đại học là một lĩnh vực nghiên cứu phổ biến được đầu tư và khám phá trong nhiều công trình học thuật [1, 2, 4, 9]. Thực tiễn QLDLNC được đề cập bao gồm các hành động, thái độ và lợi ích của các nhà nghiên cứu liên quan đến việc lưu trữ, chia sẻ và tái sử dụng DLNC [1]. Theo đó, việc đánh giá thực tiễn QLDLNC của GV sẽ chú trọng đến một số nội dung chính như nhận thức của GV về QLDLNC, các loại và định dạng DLNC được tạo lập và sử dụng, vấn đề lưu trữ, bảo quản, quản lý và chia sẻ dữ liệu.

Trước hết, thông qua các công trình nghiên cứu, việc đánh giá nhận thức của GV về QLDLNC cho thấy sự đa dạng trong mức độ nhận thức của GV về tầm quan trọng của QLDLNC. Một số học giả trong lĩnh vực khoa học xã hội và nhân văn (KHXH&NV) hiểu rõ

tầm quan trọng của việc QLDLNC, tuy nhiên một số khác có thể thiếu kiến thức toàn diện hoặc có những GV vẫn chưa đủ hiểu biết về quy trình quản lý dữ liệu và các lợi ích mà nó mang lại. Chẳng hạn, nghiên cứu của Bunkar, A. R. và Aydinoglu, A. U. [3, 4] cho thấy rằng nhà nghiên cứu đã nhận thức được tầm quan trọng và lợi ích của QLDLNC, cụ thể trong việc chia sẻ dữ liệu, cho phép tiếp cận miễn phí DLNC để tái sử dụng. Tương tự, khảo sát của các tác giả M'kulama và Akakandelwa [11] cho thấy, các nhà nghiên cứu tại ZARI đã nhận thức được khái niệm QLDLNC nhưng còn hạn chế trong nhận thức nói chung về vấn đề truy cập mở vào DLNC. Tuy nhiên, nghiên cứu trước đó của các tác giả Chigwada, Chiparusha và Kasiroori [5] lại phản ánh rằng, nhà nghiên cứu ở Zimbabwe còn thiếu kiến thức và kỹ năng. Các quan điểm và nhận thức này đóng một vai trò quan trọng và ảnh hưởng đến việc áp dụng các phương thức thực hành QLDLNC. Đồng thời, mức độ nhận thức của GV về lợi ích QLDLNC cũng có thể là yếu tố tạo động lực giúp họ thực hành QLDLNC.

Sự đa dạng của các hoạt động QLDLNC, bao gồm thực tiễn QLDLNC của GV liên quan đến việc lưu trữ, chia sẻ và tái sử dụng DLNC đã được nhận diện trong nhiều nghiên cứu nhưng ở các mức độ toàn diện khác nhau. Trong khi một số nghiên cứu nhận diện toàn bộ các giai đoạn khác nhau của hoạt động QLDLNC [17, 18], thì đa số các nghiên cứu khác chú trọng đến việc tìm hiểu một hoặc một số hoạt động cụ thể, chú trọng thực hành lưu trữ, bảo quản và chia sẻ dữ liệu. Chẳng hạn, Majid, Foo và Zhang [9] nghiên cứu thực hành QLDLNC của các nhà nghiên cứu tại Đại học Công nghệ Nanyang Singapore dưới các góc độ như: loại DLNC, lưu trữ và chia sẻ dữ liệu. Nhìn chung, các nghiên cứu nhận diện rằng, nhà nghiên cứu thường lưu trữ dữ liệu của họ trong các thiết bị cá nhân và ổ cứng ngoài, trong khi đó chia sẻ dữ liệu vẫn được coi là một vấn đề phức tạp, nhất là việc chia sẻ dữ liệu thô. Việc chia sẻ dữ liệu còn hạn chế [8], đặc biệt với các nhà nghiên cứu thuộc các lĩnh vực KHXH&NV [9, 10, 13, 16].

Một số nghiên cứu khác cũng chỉ ra rằng, nhà nghiên cứu thường tự quản lý dữ liệu với những nguyên tắc cá nhân mà không dựa vào các tiêu chuẩn chung hoặc hệ thống tổ chức chặt chẽ [12, 16].

Hình thức chia sẻ dữ liệu thường thông qua các ấn phẩm [7, 14]; hoặc chỉ được chia sẻ với một nhóm đối tượng hạn chế là các thành viên và những người có liên hệ chặt chẽ với các nhà nghiên cứu [15, 18] do có những lo ngại về vấn đề an toàn bảo mật dữ liệu. Nhìn chung, thực tiễn QLĐLNC được xem như đang trong giai đoạn trưởng thành ở các nước phát triển, trong khi ở các nước đang phát triển, hoạt động này chưa được đầu tư nhiều, các GV còn thiếu kỹ năng trong QLĐLNC [2].

Bên cạnh đó, QLĐLNC được coi là một quá trình phức tạp có thể thay đổi đáng kể tùy theo lĩnh vực chuyên ngành khoa học và phương pháp nghiên cứu [1, 20]. Chẳng hạn, các nghiên cứu về KHXX&NV thường liên quan đến dữ liệu định tính, tài liệu lưu trữ và các số liệu thống kê [1, 20] do bản chất của dữ liệu trong các lĩnh vực này có thể dựa trên việc thu thập từ hành vi, nhận thức, sở thích, thái độ và quan điểm của con người [9]. Do vậy, chất lượng dữ liệu của ngành KHXX&NV thường kém hơn so với các ngành khác, đồng thời cũng gặp nhiều rủi ro và rào cản hơn trong việc chia sẻ và tái sử dụng dữ liệu [19]. Điều này đòi hỏi cần có sự phân tích, đánh giá toàn diện về các loại và định dạng dữ liệu mà GV trong lĩnh vực KHXX&NV thường xuyên tạo ra, có thể bao gồm văn bản, hình ảnh, dữ liệu thống kê, và các định dạng khác nhằm giúp hiểu rõ về tính chất đặc biệt của từng loại dữ liệu, giúp xác định yêu cầu lưu trữ và quản lý hiệu quả, để phát triển các giải pháp QLĐLNC chuyên ngành phù hợp với các phương pháp nghiên cứu và cấu trúc dữ liệu khác.

2. PHƯƠNG PHÁP NGHIÊN CỨU

Trong nghiên cứu này, phương pháp nghiên cứu hỗn hợp, bao gồm phương pháp điều tra bằng bảng hỏi và phương pháp phỏng vấn đã được sử dụng để thu thập và phân tích cơ sở thực tiễn của đề tài. Phương

pháp điều tra được sử dụng trong nghiên cứu này thông qua bảng hỏi trực tuyến (với công cụ Google Form), với tổng số mẫu tham gia khảo sát là 151 GV. Bảng hỏi được thiết kế bao gồm các câu hỏi đóng và câu hỏi mở. Trong đó, các câu hỏi về thực tiễn hoạt động lưu trữ, quản lý và chia sẻ dữ liệu là các câu hỏi có đáp án với thang đo 3 mức độ, trong đó 1 = “Không bao giờ”, 2 = “Thỉnh thoảng” và 3 = “Thường xuyên”. Mục đích, yêu cầu và cách trả lời nghiên cứu cũng được thông tin rõ ràng đến người tham gia khảo sát trong nội dung giới thiệu đề tài nghiên cứu. Dữ liệu thu được từ bảng hỏi đã được nhập vào phần mềm SPSS, làm sạch trước khi tiến hành phân tích. Nghiên cứu áp dụng phương pháp phân tích thống kê mô tả để phân tích các chỉ số thống kê như tần số, tỷ lệ phần trăm (%), giá trị trung bình (\bar{X}) và độ lệch chuẩn (σ).

Phương pháp phỏng vấn sau đó được tiến hành với 06 GV có trình độ thạc sĩ và tiến sĩ. Dữ liệu định tính thu được từ các cuộc phỏng vấn được phân tích nội dung theo quy trình bao gồm các bước như: (1) tháo băng ghi âm các cuộc phỏng vấn, mã hoá GV tham gia phỏng vấn từ GV1 đến GV6 để đảm bảo quy tắc ẩn danh; (2) đọc lướt dữ liệu phỏng vấn để tiến hành hiệu chỉnh dữ liệu trước khi tiến hành mã hoá chính thức; (3) nhận diện những hạng mục chủ đề từ dữ liệu phù hợp với các nội dung nghiên cứu; và (4) tổng hợp các chủ đề từ dữ liệu theo các nhóm đã xác định. Kết quả phân tích dữ liệu phỏng vấn được sử dụng để bổ sung hoặc điều chỉnh và làm rõ thêm kết quả phân tích dữ liệu từ bảng hỏi.

Về đặc điểm mẫu cho khảo sát bảng hỏi, trong tổng số 151 GV tham gia khảo sát, có 60 GV nam, chiếm 39.7% và 91 GV nữ, chiếm 60.3%. Các GV tham gia trong nghiên cứu này thuộc các nhóm lứa tuổi khác nhau, trong đó GV dưới 30 tuổi là 6.0%, GV từ 30 đến 39 tuổi là 46.4%, GV từ 40 đến dưới 49 tuổi chiếm 37.1%, và GV trên 50 tuổi chiếm 10.6%. Trình độ học vấn của GV cũng khác nhau với 8% số GV tham gia khảo sát có trình độ cử nhân, 57.6% GV có trình độ thạc sĩ, và 37.1% GV có trình độ tiến sĩ.

Về kinh nghiệm nghiên cứu khoa học của GV, có 145/151 GV trả lời câu hỏi về việc đã thực hiện đề tài các cấp, chiếm tỷ lệ 96%. Trong số này, có 44.1% GV đã tham gia đề tài cấp trường, 20.0% GV đã tham gia đề tài cấp ĐHQG, 2.1% GV đã tham gia đề tài cấp tỉnh/thành phố, 4.8% GV đã và đang tham gia đề tài cấp quốc gia, 0.7% GV đã tham gia đề tài mang tính hợp tác quốc tế, 25.5% GV đã đang làm đề tài nghiên cứu của cấp học thạc sĩ hoặc tiến sĩ, 2.8% GV đang làm các đề tài khác.

Bảng 1. Loại đề tài nghiên cứu giảng viên đã tham gia

Loại đề tài nghiên cứu	Tần suất	Tỷ lệ %	Tỷ lệ % hợp lệ	Tỷ lệ % tích lũy
Đề tài cấp Trường	64	42.4	44.1	44.1
Đề tài cấp Đại học Quốc Gia	29	19.2	20.0	64.1
Đề tài cấp thành phố	3	2.0	2.1	66.2
Đề tài cấp quốc gia	7	4.6	4.8	71.0
Đề tài hợp tác quốc tế	1	.7	.7	71.7
Đề tài theo cấp học (thạc sĩ, tiến sĩ)	37	24.5	25.5	97.2
Ý kiến khác	4	2.6	2.8	100.0
Tổng	145	96.0	100.0	
Giá trị khuyết	6	4.0		
Tổng	151	100.0		

3. KẾT QUẢ NGHIÊN CỨU

3.1. Về các loại dữ liệu nghiên cứu được thu thập hoặc tạo lập

Kết quả phân tích dữ liệu được trình bày trong Bảng 2 cho thấy, văn bản (với các định dạng như TXT, DOC, PDF, RTF, HTML, XML) là loại dữ liệu phổ biến nhất được GV tạo lập hoặc thu thập với 147 lượt chọn, chiếm tỷ lệ 97.4 % trong tổng số mẫu nghiên cứu. Tiếp đến là các dữ liệu đa phương tiện với 96 lượt

trả lời, chiếm 63.6% tổng số GV tham gia khảo sát, và dữ liệu của thống kê với 73 lượt trả lời, chiếm tỷ lệ 48.3% số mẫu nghiên cứu. Tỷ lệ thấp nhất liên quan đến các dữ liệu tạo lập từ các ngôn ngữ thiết kế cơ sở dữ liệu và các mã lập trình. Một số ý kiến khác nêu lên loại dữ liệu thu thập từ phỏng vấn theo phương pháp lý thuyết cơ sở (grounded theory).

Bảng 2. Loại dữ liệu nghiên cứu được giảng viên thu thập và tạo lập

Loại dữ liệu	Trả lời		Tỷ lệ trên tổng số mẫu
	Số lượt chọn	Tỷ lệ	
Văn bản (TXT, DOC, PDF, RTF, HTML, XML)	147	40.6%	97.4%
Đa phương tiện (hình ảnh, âm thanh, video)	96	26.5%	63.6%
Thống kê dữ liệu (SAS, SPSS, CSV, MAT, XLS)	73	20.2%	48.3%
Hệ thống thông tin địa lý (ArcGIS, MapInfo, QGIS, Bản đồ thế giới)	22	6.1%	14.6%
Cơ sở dữ liệu (MS Access, Oracle, SQL, MySQL, MariaDB)	11	3.0%	7.3%
Mã lập trình (mã nguồn) Java, C, Perl, Python, Ruby, PHP, JSON, Dplyr, D3, HTML)	10	2.8%	6.6%
Ý kiến khác	3	0.8%	2.0%
Tổng	362	100.0%	239.7%

3.2. Về cơ sở hạ tầng công nghệ lưu trữ dữ liệu

Trong nghiên cứu này, cơ sở hạ tầng công nghệ để lưu trữ dữ liệu được hiểu bao gồm cả các thiết bị phần cứng hoặc các dịch vụ phần mềm và đám mây giúp lưu trữ dữ liệu. Phân tích dữ liệu được trình bày trong Bảng 3 cho thấy GV thường xuyên sử dụng các thiết bị cá nhân để lưu trữ dữ liệu với $\bar{X}=2.99$ (ĐLC = .115), tiếp theo là các dịch vụ lưu trữ trực tuyến như Google Drive hay iCloud với $\bar{X}=2.64$ (ĐLC = .496) và lưu trữ dữ liệu vào các thiết bị ngoại vi như ổ cứng di động hoặc USB với $\bar{X}=2.37$ (ĐLC = .668). Ít thường

xuyên nhất là việc lưu trữ dữ liệu trên các kho lưu trữ mở cho các lĩnh vực khoa học xã hội, tâm lý, kinh tế với $\bar{X}=1.26$ ($\text{ĐLC} = .526$).

Bảng 3. Thống kê mức độ thường xuyên sử dụng các thiết bị hoặc dịch vụ lưu trữ dữ liệu của giảng viên

Thiết bị/ dịch vụ lưu trữ	Số lượt (N)	Trung bình	Độ lệch chuẩn
Máy tính cá nhân (PC hoặc Laptop)	149	2.99	.115
Các dịch vụ lưu trữ trực tuyến (ví dụ: Google Drive, Dropbox, iCloud...)	140	2.64	.496
Thiết bị lưu trữ ngoài (CD /DVD, Ổ cứng di động hoặc USB Flash)	142	2.37	.668
Hệ thống máy chủ của đơn vị/cơ quan	131	1.80	.684
Kho lưu trữ mở cho các lĩnh vực khoa học xã hội, tâm lý, kinh tế... (SocArXiv, psyArXiv, RePEc)	125	1.26	.526
Valid N (listwise)	122		

Kết quả phỏng vấn cho các thông tin tương đồng. Chẳng hạn, có 03 GV cho biết họ thường xuyên quản lý và lưu trữ dữ liệu trên máy tính cá nhân hoặc các ổ cứng di động vì tính thuận tiện và có thể sử dụng trong nhiều tình huống như không có mạng internet vẫn có thể lưu trữ và chia sẻ dữ liệu; hoặc 01 GV khác thường sử dụng Google Drive vì đây là “*công cụ rất tiện ích, linh hoạt và đặc biệt được miễn phí*”.

3.3. Về bảo mật và quản lý dữ liệu

An toàn và bảo mật dữ liệu đóng vai trò quan trọng trong đảm bảo quyền kiểm soát dữ liệu trong QL DLNC. Theo kết quả khảo sát được trình bày trong Bảng 4, đa số GV tham gia khảo sát không đặt chế độ bảo mật dữ liệu, chiếm tỷ lệ 61.6%. Chỉ có 34.4% GV đặt mật khẩu và một tỷ lệ rất nhỏ (3.3%) mã hóa dữ liệu. Trong khi đó, kết quả phỏng vấn cho thấy, có 5 GV có đặt mật khẩu cho thiết bị lưu

trữ và chỉ có 1 GV trả lời không bảo mật dữ liệu:

“Tôi không đặt mật khẩu cho bất cứ thiết bị lưu trữ nào, chỉ lưu trữ trên máy tính và chỉ cần giữ máy tính của mình cẩn thận thì sẽ ổn” (GV1).

Trong khi đó, có GV ý thức cần đặt mật khẩu mạnh bằng cách “*lấy các thuật ngữ chuyên ngành ví dụ như một câu thơ chữ Hán mà mình tâm thêm hoa thị đầu dòng và những con số...*”, có GV khác cho rằng “*chỉ đặt mật khẩu đơn giản cho ổ cứng chứ chưa có giải pháp bảo mật chuyên sâu*”.

Bảng 4. Cách thức giảng viên bảo mật dữ liệu

Cách thức bảo mật dữ liệu	Tần suất	Tỷ lệ %	Tỷ lệ % hợp lệ	Tỷ lệ % tích lũy
Đặt mật khẩu	52	34.4	34.4	34.4
Dữ liệu được mã hóa	5	3.3	3.3	37.7
Dữ liệu không được bảo mật	93	61.6	61.6	99.3
Ý kiến khác	1	.7	.7	100.0
Tổng	151	100.0	100.0	

Dữ liệu được trình bày trong Bảng 5 cũng cho thấy, GV chọn cách thức tự lưu trữ và quản lý dữ liệu trên máy tính cá nhân và các thiết bị lưu trữ ngoài chiếm tỷ lệ 85% tổng số lượt lựa chọn. Kết quả phỏng vấn cho thấy sự tương đồng với kết quả khảo sát, tất cả 6 GV cho biết họ tự quản lý và lưu trữ dữ liệu trên máy tính cá nhân, ổ cứng di động và sử dụng miễn phí hoặc mua các gói dung lượng lưu trữ trên Google Drive do chi phí thấp, có tính chủ động và thuận tiện. Kết quả phỏng vấn cho thấy sự nhận diện tương đồng khi tất cả GV chưa sử dụng sách dịch vụ lưu trữ của bên ngoài, chẳng hạn:

“Tôi tự lưu trữ dữ liệu và chưa sử dụng bất kỳ dịch vụ nào của thư viện hay doanh nghiệp vì chưa được tiếp cận, hiểu biết tường tận về các dịch vụ và có chút băn khoăn, lo ngại về phí dịch vụ QL DLNC”.

Bảng 5. Cách thức giảng viên quản lý dữ liệu

Cách thức giảng viên quản lý dữ liệu	Trả lời		Tỷ lệ trên tổng số mẫu
	Số lượt chọn (N)	Tỷ lệ	
Tự lưu trữ và quản lý	147	85.0%	97.4%
Dịch vụ lưu trữ của dự án	6	3.5%	4.0%
Đăng ký dịch vụ lưu trữ nước ngoài	9	5.2%	6.0%
Ý kiến khác	11	6.4%	7.3%
Tổng số	173	100.0%	114.6%

3.4. Về chia sẻ dữ liệu nghiên cứu

Kết quả khảo sát được trình bày trong Bảng 6 cho thấy, GV thường xuyên sử dụng email để chia sẻ thông tin và dữ liệu với $\bar{X}=2.89$ ($\text{ĐLC} = .333$) tương đương 89.9% số GV tham gia khảo sát. Tiếp theo là các hình thức chia sẻ dữ liệu qua các ứng dụng trực tuyến như Google Docs với $\bar{X}=2.65$ ($\text{ĐLC} = .548$), và qua các thiết bị lưu trữ ngoài với $\bar{X}=2.45$ ($\text{ĐLC} = .613$).

Bảng 6. Cách thức giảng viên chia sẻ dữ liệu nghiên cứu

Cách thức chia sẻ DLNC	Số lượt	Trung bình	Độ lệch chuẩn
Gửi dữ liệu qua email	148	2.89	.333
Sử dụng các ứng dụng của Google (Google Docs, Google Sheets,...)	141	2.65	.548
Dùng thiết bị lưu trữ ngoài (USB, CD/DVD, ổ đĩa di động)	144	2.45	.613
Sử dụng các công cụ thư mục chia sẻ như Dropbox...	124	2.06	.747
Sử dụng các công cụ quản lý trích dẫn và tài liệu tham khảo, lưu trữ (Refworks, BibTex, EndNote, Papers, Stata)	128	1.68	.72

Kết quả phỏng vấn cho thấy các nhận định tương đồng khi tất cả các GV cho rằng

họ thường xuyên chia sẻ dữ liệu qua email cá nhân vì đây là công cụ rất thông dụng, dễ sử dụng và phần lớn GV cũng xác định họ thường chia sẻ dữ liệu qua ứng dụng miễn phí.

“Tôi thường xuyên chia sẻ dữ liệu trong nhóm nghiên cứu trên Google Drive, Onedrive, nhưng do dùng miễn phí nên gặp khó khăn khi chia sẻ dung lượng lớn. Tôi lo ngại việc chia sẻ dùng chung rất dễ mất dữ liệu vì không đặt mật khẩu cho tệp tin chia sẻ được và vô tình hoặc cố ý xóa mất dữ liệu”.

4. THẢO LUẬN VÀ KẾT LUẬN

Các hoạt động QL DLNC được tìm hiểu trong nghiên cứu này bao gồm việc thu thập và tạo lập dữ liệu, việc sử dụng các cơ sở hạ tầng công nghệ để lưu trữ, bảo mật dữ liệu, hoạt động tổ chức và chia sẻ dữ liệu. Thứ nhất, kết quả khảo sát trình bày ở phần trên cho thấy, dữ liệu được tạo lập dưới dạng văn bản chiếm tỷ lệ cao nhất, phản ánh đặc thù của các hoạt động nghiên cứu trong lĩnh vực KH XH & NV vốn thường có đối tượng và phương pháp nghiên về các vấn đề của con người, cuộc sống và xã hội thông qua các phương pháp phân tích và suy luận định tính. Ngoài ra, việc GV thu thập DLNC dưới dạng đa phương tiện, các dữ liệu thống kê, dữ liệu liên quan đến hệ thống thông tin địa lý cũng cho thấy các lĩnh vực chuyên môn của GV tại Trường, liên quan đến các khoa học chuyên ngành như địa lý, đô thị học, báo chí và truyền thông. Kết quả nghiên cứu thống nhất với các nghiên cứu trước đó của Akers và Doty [1], Majid, Foo và Zhang [9], Weller và Monroe-Gulick [20] khi nhận định bản chất của loại dữ liệu trong KH XH & NV có thể dựa trên hành vi, nhận thức, sở thích, thái độ và quan điểm của con người. Nghiên cứu cho thấy rằng, dữ liệu thường được tạo lập và phản ánh đặc thù và đa dạng của các hoạt động nghiên cứu trong lĩnh vực KH XH & NV.

Thứ hai, việc xác định và hiểu về cơ sở hạ tầng công nghệ mà GV sử dụng để lưu trữ DLNC sẽ giúp đề xuất những cải tiến có thể được thực hiện cho QL DLNC. Kết quả nghiên cứu cho thấy hầu hết GV thường xuyên lưu trữ

dữ liệu dưới các định dạng khác nhau trong các thiết bị cá nhân như máy tính và ổ cứng di động; tiếp theo đó là các dịch vụ lưu trữ trực tuyến. Kết quả này thống nhất với nghiên cứu trước đó của Majid, Foo và Zhang (2018) [9], cũng xác định ưu tiên của đa số người tham gia khảo sát từ mọi lĩnh vực là lưu trữ DLNC trên thiết bị cá nhân và dịch vụ lưu trữ đám mây. Những ưu điểm bao gồm tính tiện lợi, khả năng sở hữu cá nhân và có thể chủ động kiểm soát việc lưu trữ. Việc xác định cơ sở hạ tầng công nghệ sử dụng để lưu trữ DLNC có thể giúp đề xuất những cải tiến cho QLNLNC.

Thứ ba, kết quả nghiên cứu cũng giúp nhận diện những vấn đề về thực trạng việc bảo mật, bảo quản và tổ chức dữ liệu của các GV tham gia nghiên cứu. Theo đó, GV đã ý thức về vấn đề bảo mật dữ liệu nhưng có kiến thức và kỹ năng bảo mật ở mức độ khác nhau, nhìn chung còn ở mức độ hạn chế. Đa số giảng viên tự quản lý dữ liệu, dẫn đến việc cách thức tổ chức và quản lý sẽ đa dạng, dựa trên các nguyên tắc tổ chức được đặt ra bởi cá nhân GV mà chưa dựa trên các tiêu chuẩn chung. Kết quả nghiên cứu ủng hộ các nghiên cứu trước đó của Manu [10], Schopfel và Prost [13], Piracha và Ameen [12], trong đó nhận diện những hạn chế trong lưu trữ, tổ chức và bảo quản dữ liệu, bao gồm việc tổ chức dựa trên nguyên tắc cá nhân hơn là theo các quy tắc chung. Điều này dẫn đến những khó khăn trong quá trình tìm kiếm và sử dụng DLNC, đặc biệt là khi dữ liệu tổ chức không có cấu trúc rõ ràng nên cần có các hỗ trợ để lưu trữ và tổ chức dữ liệu hiệu quả.

Thứ tư, nghiên cứu cho thấy GV đã nhận thức được tầm quan trọng của việc chia sẻ DLNC và thường thực hành việc chia sẻ mang tính cá nhân qua các ứng dụng miễn phí. Dữ liệu phỏng vấn cũng làm nổi bật một số thách thức và lo ngại của GV trong quá trình chia sẻ dữ liệu, bao gồm dung lượng dữ liệu được chia sẻ, những vấn đề về an toàn bảo mật dữ liệu. Kết quả nghiên cứu thống nhất với nghiên cứu trước đó khi nhận diện những thách thức của chia sẻ dữ liệu do những hạn chế về các công cụ chia sẻ, những lo ngại của các nhà nghiên cứu về khả năng mất mát,

lạm dụng và hiểu lầm về dữ liệu [8, 14, 28]. Những vấn đề này đặt ra những yêu cầu cần được giải quyết để tối ưu hóa quá trình quản lý và chia sẻ DLNC trong trường đại học.

Nghiên cứu đã thực hiện đánh giá thực tiễn QLNLNC của GV tại ĐHKHXHNV, ĐHQG-HCM về các nội dung: các loại DLNC được tạo lập hoặc thu thập, việc lưu trữ dữ liệu, bảo mật, quản lý dữ liệu và chia sẻ dữ liệu. Kết quả nghiên cứu không chỉ giúp định hình hướng phát triển trong QLNLNC mà còn đặt ra những thách thức cụ thể cần giải quyết để tối ưu hóa quá trình quản lý và chia sẻ DLNC của GV tại Trường, cung cấp dữ liệu thực tiễn cho các bên liên quan, trong đó có thư viện đại học trong việc cung cấp các dịch vụ cơ sở hạ tầng để QLNLNC và cung cấp các khóa huấn luyện giúp nâng cao kiến thức, kỹ năng và năng lực thực hành QLNLNC cho GV. Đề tài cũng mở ra hướng nghiên cứu tiếp theo trong tìm hiểu nhu cầu của GV trong việc sử dụng các dịch vụ hỗ trợ QLNLNC tại Trường.

TÀI LIỆU THAM KHẢO

1. Akers, K. G., & Doty, J. (2013). Disciplinary differences in faculty research data management practices and perspectives. *International Journal of Digital Curation*, 8(2), 5-26. doi:10.2218/ijdc.v8i2.263
2. Ashiq, M., Usmani, M. H., & Naeem, M. (2022). A systematic literature review on research data management practices and services. *Global Knowledge, Memory and Communication*, 71(8/9), 649-671.
3. Aydinoglu, A. U., Dogan, G., & Taskin, Z. (2017). Research data management in Turkey: Perceptions and practices. *Library Hi Tech*, 35(2), 271-289.
4. Bunkar, A. R., & Bhatt, D. D. (2020). Perception of Researchers; Academicians of Parul University towards Research Data Management System; Role of Library: A Study. *DESIDOC Journal of Library & Information Technology*, 40(3), 139-146.
5. Chigwada, J., Chiparousha, B., & Kasiroori, J. (2017). Research Data Management in Research Institutions in

- Zimbabwe. *Data Science Journal*, 16(31), 1-9. doi:10.5334/dsj-2017-03.
6. Cox, A.M. & Pinfield, S. (2014). Research Data Management and Libraries: Current Activities and Future Priorities. *Journal of Librarianship and Information Science*, 46(4), 299-316.
7. Elsayed, A.M. and Saleh, E.I. (2018), "Research data management and sharing among researchers in Arab universities: an exploratory study". *IFLA Journal*, 44 (4), 281-299.
8. Joo, S. and Peters, C. (2019), "User needs assessment for research data services in a research university". *Journal of Librarianship and Information Science*, 52 (3), doi:10.1177/0961000619856073.
9. Majid, S., Foo, S., & Zhang, X. (2018). Research data management by academics and researchers: Perceptions, knowledge and practices. In *Maturity and Innovation in Digital Libraries: 20th International Conference on Asia-Pacific Digital Libraries, ICADL 2018, Hamilton, New Zealand, November 19-22, 2018, Proceedings 20* (pp. 166-178). Springer International Publishing.
10. Manu, T. R. (2018, August). Researchers' perceptions on research data management: a survey. In *International Conference on Exploring the Horizons of Library and Information Science: From Libraries to Knowledge Hubs*. Bangalore: Document Research and Training Center (DRTC) (pp. 115-128).
11. M'kulama, A. C., & Akakandelwa, A. (2021). Research Data Sharing and Reuse Through Open Data: Assessing Researcher Awareness and Perceptions at the Zambia Agricultural Research Institute (ZARI). In *Open Access Implications for Sustainable Social, Political, and Economic Development* (pp. 284-306). IGI Global.
12. Piracha, H. A., & Ameen, K. (2018). Research data management practices of faculty members. *Pakistan Journal of Information Management & Libraries*, 20, 60-75.
13. Schopfel, J., & Prost, H. (2016). Research data management in social sciences and humanities: A survey at the University of Lille (France). Truy cập từ <https://lilloa.univ-lille.fr/bitstream/handle/20.500.12210/71273/https://hal.univ-lille.fr/hal-01395816/document?sequence=2>
14. Stamatopols, A., Neville, T. and Henry, D. (2016). Analyzing the data management environment in a master's-level institution. *The Journal of Academic Librarianship*, 42 (2), 154-160. doi: 10.1016/j.acalib.2015.11.004.
15. Tripathi, M., Chand, M., Sonkar, S.K. and Jeevan, V.K.J. (2017). A brief assessment of researchers' perceptions towards research data in India. *IFLA Journal*, 43(1), 22-39.
16. Unal, Y., Chowdhury, G., Kurbanoglu, S., Boustany, J., & Walton, G. (2019). Research data management and data sharing behaviour of university researchers'. *Information Research: an international electronic journal*, 24(1).
17. Vela, K. and Shin, N. (2019). Establishing a research data management service on a health sciences campus. *Journal of eScience Librarianship*, 8 (1). doi: 10.7191/jeslib.2019.1146.
18. Vilar, P. and Zabukovec, V., (2019). Research data management and research data literacy in Slovenian science. *Journal of Documentation*, 75 (1). doi: 10.1108/JD-03-2018-0042.
19. Wang, X., Duan, Q., & Liang, M. (2020). Scholars' data reuse behaviors in disciplinary context: A meta-synthesis study. *iConference 2020 Proceedings*.
20. Weller, T. & Monroe-Gulick, A. (2014). Understanding methodological and disciplinary differences in the data practices of academic researchers. *Library Hi Tech*, 32(3), 467-482. doi: 10.1108/LHT-02-2014-0021.
- (Ngày Tòa soạn nhận được bài: 15-11-2023; Ngày phản biện đánh giá: 10-12-2023; Ngày chấp nhận đăng: 15-01-2024).*