Security on medical data sharing
(a literature review)

Dayana Spagnuelo, Gabriele Lenzini
University of Luxembourg
Email: {dayana.spagnuelo, gabriele.lenzini} @uni.lu

Keywords—literature review, security, medical data, electronic health records.

I. Introduction

Medical records (e.g., test results and health reports) are about patients. Hospitals and healthcare institutions generate them after a patient's visit. Today they are digitized, stored electronically, and accessed remotely by professionals.

European directives suggest that patients should access these records too. Besides, they say, patients should have control over these data and be informed if and when their records are shared and how secure they are [1]. These requirements are hard to be met.

From a patient's perspective, the viewpoint of this paper, it may be easier to address at least one of such requirements: to inform patients about how secure their data are. This is a property usually referred as *transparency*, but a clear meaning of the word is still missing. According to [2] transparency ought to be regarded as an additional feature that qualifies security. So, security can be said to be transparent when is intelligible to human user. It opposes an opaque security, which holds technically but without the user's being aware of it. Thus, transparency is a socio-technical security property.

Transparency, is not a new term. It has been proposed in relation to *Transparency Enhancing Tools* (TETs) [3]. These are usually browser extensions that read out web server's privacy policies and inform users concisely, for instance, that a web server records the user's whereabouts and may sell the user's data to third parties. TETs have been discussed in relation to electronic health records [2], but no concrete solution has been proposed. Transparency in the medical domain is still an unfulfilled requirement.

Contribution. We survey the literature in medical data sharing and discusses what are the main security concerns in it. We intend also to figure out whether transparency is debated in that domain, in relation to which other properties, and which meaning and role are given to it.

II. Methodology and tools

We browsed the state of the art by searching for papers via "Findit.lu"[1]. This is the largest library portal in Luxembourg, and it is entirely dedicated on searching for electronic contents. It indexes a large number of important scientific digital libraries such as, among many others, LNCS, the ACM Digital Library, IEEEXplore, ScienceDirect, Scopus, and Medline.

[1]The portal is accessible via www.bibnet.lu, or directly at, www.findit.lu

We queried for "Security" and "Medical Data Sharing", and we looked for papers containing them in the title, in the abstract, in the list of keywords, and in the entire body. We chose "Security" because it is a general term: we expect that a paper that addresses more precise security properties will also mention "security" somewhere its text. We chose "Medical Data Sharing" to refine our domain to papers that discuss sharing medical data.

First, we queried without constraints on the year of publication. We got as many as 526 articles, too many for us to be able to read or scan them all. Thus, we restricted the focus to the last ten years, from 2004 to now. Excluding the repeated results and the papers not available for download, our pool shrank down to a total of 75 papers. We read the abstract and skimmed through the content of all of them. It turned out that 20 papers were about medical data sharing but with no focus on "security": the word appeared to be mentioned but the concept is not discussed. We discarded those papers and, after this skimming, we were left with a pool of 55 papers.

We organized our findings around one question: "*what particular security property the paper is about?*". To answer this question helped us to classify the papers depending on the property, or properties, they debate. It also helps us to understand whether transparency is considered as a security requirement and, if it is, in relation to which other property.

III. Major findings

Answering our main question, and so looking into what security properties our pool of papers is about, lead us to identify eight main security categories, each concerning policies, tools, or techniques meant to guarantee, preserve, or enforce a specific property. The 8 categories are the following: *Privacy*, concerning to provide anonymity to the data owner or to empower her to define who can operate on the data; *User authentication*, concerning to enhance the way in which users are authenticated electronically; *Access control*, concerning better ways to define who can access medical data and in what circumstances; *Data authenticity*, concerning to prove that the data origin is authentic, that is coming from the source as it is claimed; *Data Integrity*, concerning solutions to guarantee and prove that the data have not been manipulated or tampered with; *Confidentiality*, concerning to prevent the disclosure of data content to non-authorized third parts; *Auditability*, concerning to help the data owner to retrieve information clarifying how her data is being used; *Transparency*, concerning to guarantee openness about security policies and processes.

Most of the surveyed papers argue about data confidentiality (see Figure 1). This property is invoked in relation

to protect the data transmitted in open channels, such as the internet, or stored in open data bases, such as the cloud. One comment is mandatory: in the pool "confidentiality" there are 27 papers, namely [4]–[30]. Some of those were, per keywords, first gathered under "privacy". A closer look revealed that they are using the term inappropriately since their concern is mainly about encrypting data. But, encryption *per se* is insufficient to guarantee that the user's personal and sensitive information remains private during the whole data life cycle; more sophisticated techniques have to be in place for privacy to be protected. Thus, we decided to re-classify those works as being about confidentiality, adding those up to the ones already in that category.

Confidentiality is constantly discuss together with data integrity and data authenticity. That is because encryption is the technique that is more often adopted to enforce confidentiality in medical systems and the same technique is also proposed for data authenticity and integrity. In a total of 16 papers about data integrity (i.e., [5], [6], [8]–[10], [12], [13], [16], [18], [22], [23], [25], [29], [31]–[33]) only three works do not discuss confidentiality. We observed a very similar scenario with the category data authenticity. Only three works do not discuss confidentiality, out of 9 papers discussing data authenticity (i.e., [8], [9], [12], [22], [25], [29], [31]–[33]). Also, all works that examine data authenticity discuss data integrity too.

After confidentiality, the second and third most discussed security properties are privacy and access control. We found out that 20 works discuss privacy (the correct interpretation of this term) [14], [20], [25], [26], [30], [33]–[47], and that 19 papers discuss access control [11], [13], [19], [22], [23], [25], [29], [34], [37], [41]–[43], [45], [48]–[53].

User authentication seems not a major concerns as it is present only in 3 papers [13], [37], [54]. We do not have enough data to justify this lack of interest in authentication, but we can speculate on it. An hypothesis we have is that most of the works give for granted that medical data are accessed only by professionals and that they are considered trustworthy. Similarly, we claim that the lack of interest in user authentication may indicate that there is not yet a widespread concern about opening the access of the health data to patients. This is, indeed, a requirement that only very recently has been debated and brought to the attention of the society. If concrete actions to open up access to patients were taken

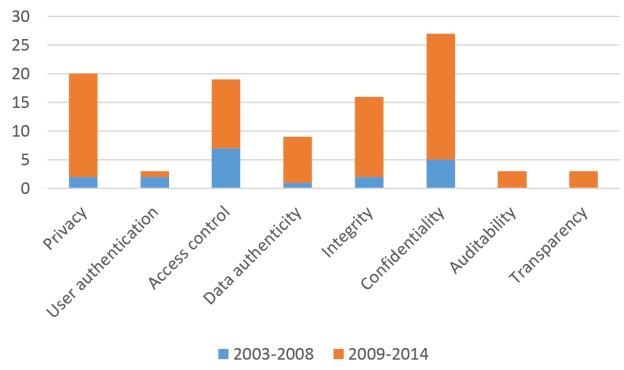


Fig. 1. Number of papers published per category from 2004 to now. We distinguished the first from the second 5 years.

into consideration, it would, we expect, raise more attention about identification and authentication. Indeed the works which discuss such a feature have identification and authentication as their main requirement (e.g., see [55]). A similar speculation, i.e., that the patient-centred approach is not yet under the bull's-eye in medical data security, concerns also the last two properties, transparency – the one of interest for this paper– and auditability. Auditability is subject of discussion of only 3 papers [33], [47], [56], *ex equo* (so to speak) with transparency which is mentioned as well in 3 papers [36], [42], [47].

Transparency is regarded as openness about policies and processes (we quote, "*there should be openness and transparency about policies, procedures, and technologies that directly affect individuals and/or their individually identifiable health information*" [36]) as well as a predisposition to increase responsibility and therefore presented with accountability (we quote, "*Transparency and accountability will be critical to helping society manage the privacy risks that accumulate from expeditious progress in communication, storage, and search technology*" [47]). Relevantly for this work, Routsalainen et al [42] propose transparency as the property to be informative towards patient. In fact they point out the lack of transparency since " *[the] patient is not automatically aware which professionals or entities are processing her EHR and for what purposes. [The] patient are not aware of all disclosures of the content of her EHR".*

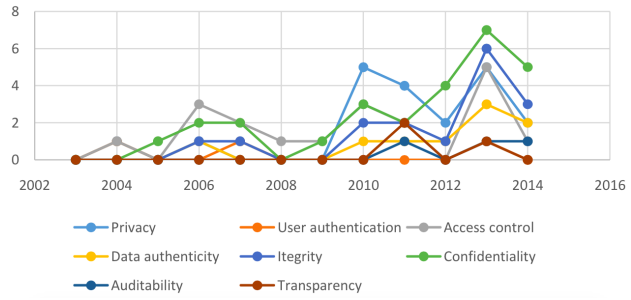
IV. Discussion and Conclusion

Our review has an obvious limitation: it considers papers that matched only two key-phrases, "security" and "medical data sharing". However, "security" is a generic terms under which we were able to find papers discussing more specific properties and requirements. "Medical data sharing" is our target, so this choice is justified. Still one could question why we did not searched for synonyms, and whether, in not doing so, we missed some important papers. Our searching on the whole body of the paper, however, was sufficient to catch works about electronic health records, bio-medical data, health care information systems, health-grid. Therefore, we judged the choice of our key-phrases sufficiently good for our scope.

This survey, organized around the works published in the last 10 years, shows that confidentiality and privacy are the major concerns in security for medical data (see also Figure 2). This comes with no surprise. About transparency, the survey shows that this requirement has just began to be addressed; all the considered papers see transparency related to inform users and make policies and processes openly available. This seems to be the interpretation of "transparency" in the medical domain, a meaning which matches what we propose. However, there is no formalization of it and no standard solution that makes a medical system compliant to it.

We also observed that the majority of papers were published in the last 5 years, which endorses the hypothesis that security is a relative young concern in medical systems engineering. Although we already had some hint of it, after having looked at the recent growth of interest as this survey reports, it is evident that there is still little attention from the security community towards auditability, transparency, and user authentication, at least in relation to medical data systems. (We

Fig. 2. Number of papers per year per category



did not searched into the literature of auditability and checked for use cases on medical data (e.g., as in [57]). Auditability and transparency are essential wherever humans need to be informed about practices in sharing sensitive personal data. No solution exists to comply with current EU regulations on this. Our first impression is that both categories are relatively understudied in the medical sectors. We expect a growth in attention to these properties as the idea of user empowerment will get more popular. User authentication seems suspiciously undervalued in the papers we surveyed. It is hard, from the data we have, to infer why. It may be that there are already good-enough authentication solutions to which medical systems can resort to. But, if we have to attempt another explanation, we are keen to suppose that current medical data are accessed mainly by professionals and that these roles are assumed to be trustworthy. Authentication is therefore implemented by simple login and password. Similarly as what we claimed while discussing transparency, if the EU directive suggesting to let users access their medical data should take off, we expect the problem of user authentication to became a pillar for the working of other several security features, and to foster a renewed interest.

REFERENCES

[1] E. P. E. Commission, "EU Directive 95/46/EC - The Data Protection Directive - IP/12/46 - 25/01/2012," October 2005 and 2012.

[2] A. Ferreira and G. Lenzini, "Can Transparency Enhancing Tools support patients accessing Electronic Health Records?" in *Proc. of the 3rd World Conference on Information Systems and Technologies, to be held at Ponta Delgada, So Miguel, Azores, Portugal, 1 - 3 April 2015*, 2015, (to appear).

[3] M. Janic, J. Wijbenga, and T. Veugen, "Transparency Enhancing Tools (TETs): An Overview," in *Socio-Technical Aspects in Security and Trust (STAST), 2013 Third Workshop on*, June 2013, pp. 18–25.

[4] C. A. Cassa, R. A. Miller, and K. D. Mandl, "A novel, privacy-preserving cryptographic approach for sharing sequencing data." *JAMIA*, vol. 20, no. 1, pp. 69–76, 2013.

[5] F. E.-Z. A. Elgamal, N. A. Hikal, and F. E. Z. Abou-Chadi, "Secure medical images sharing over cloud computing environment," *International Journal of Advanced Computer Science and Applications(IJACSA)*, vol. 4, no. 5, 2013.

[6] S. H. Han, M. H. Lee, S. G. Kim, J. Y. Jeong, B. N. Lee, M. S. Choi, I. K. Kim, W. S. Park, K. Ha, E. Cho, Y. Kim, and J. B. Bae, "Implementation of Medical Information Exchange System Based on EHR Standard," *Healthcare informatics research*, 2010.

[7] R. Kettimuthu, R. Schuler, D. Keator, M. Feller, D. Wei, M. Link, J. Bresnahan, L. Liming, J. Ames, A. Chervenak, I. Foster, and C. Kesselman, "A data management framework for distributed biomedical research environments," in *Proceedings of the 2010 Sixth IEEE International Conference on e-Science Workshops*, 2010, pp. 72–79.

[8] P. Rewagad and Y. Pawar, "Use of digital signature and rijndael encryption algorithm to enhanced security of data in cloud computing services," *IJCA Proceedings on Emerging Trends in Computer Science and Information Technology*, no. 2, pp. 5–7, April 2012.

[9] H. Satoh, N. Niki, K. Eguchi, H. Ohmatsu, M. Kusumoto, M. Kaneko, R. Kakinuma, and N. Moriyama, "Teleradiology network system using the web medical image conference system with a new information security solution," in *Proc. SPIE*, 2013.

[10] D. Thilakanathan, S. Chen, S. Nepal, R. Calvo, and L. Alem, "A platform for secure monitoring and sharing of generic health data in the cloud," *Future Gener. Comput. Syst.*, pp. 102–113, Jun. 2014.

[11] F. Al-Nayadi and J. H. Abawajy, "An authorization policy management framework for dynamic medical data sharing," in *The International Conference on Intelligent Pervasive Computing*, Oct 2007, pp. 313–318.

[12] R. Basavegowda and S. Seenappa, "Electronic medical report security using visual secret sharing scheme," in *15th International Conference on Computer Modelling and Simulation*, April 2013, pp. 78–83.

[13] F. Al-Nayadi and J. H. Abawajy, "An authentication framework for e-health systems," in *IEEE International Symposium on Signal Processing and Information Technology*, Dec 2007, pp. 616–620.

[14] W. K. Seng, R. Besar, and F. Abas, "Collaborative support for medical data mining in telemedicine," in *2nd Information and Communication Technologies*, vol. 1, 2006, pp. 1894–1899.

[15] K. Chida, G. Morohashi, H. Fuji, F. Magata, A. Fujimura, K. Hamada, D. Ikarashi, and R. Yamamoto, "Implementation and evaluation of an efficient secure computation system using 'r' for healthcare statistics," *Journal of the American Medical Informatics Association*, vol. 21, pp. e326–e331, 2014.

[16] T. Ermakova and B. Fabian, "Secret sharing for health data in multi-provider clouds," in *IEEE 15th Conference on Business Informatics*, July 2013, pp. 93–100.

[17] M. A. Hajjaji, S. Ajili, A. Mtibaa, and E.-B. Bourennane, "A new system for watermarking based on the turbo-codes and wavelet 5/3," in *13th International conference on Sciences and Techniques of Automatic control & computer engineering*, Tunisia, Dec. 2012.

[18] M. A. Hajjaji, A. Mtibaa, and E. bey Bourennane, "A watermarking of medical image: New approach based on "multi-layer " method," 2011.

[19] S. Hameed, H. Yuchoh, and W. Al-Khateeb, "A model for ensuring data confidentiality: In healthcare and medical emergency," in *4th International Conference On Mechatronicsw*, May 2011, pp. 1–5.

[20] A. Hossain, S. Ferdous, S. Islam, and N. Maalouf, "Rapid cloud data processing with healthcare information protection," in *IEEE World Congress on Services*, June 2014, pp. 454–455.

[21] W. Lee, S. Kim, M. Noh, and H. Kim, "A virtualized network model for wellness information technology research," in *International Conference on IT Convergence and Security*, Dec 2013, pp. 1–3.

[22] J. Liu, X. Huang, and J. K. Liu, "Secure sharing of personal health records in cloud computing: Ciphertext-policy attribute-based signcryption," *Future Generation Computer Systems*, 2014.

[23] S. N. Bharti Ratan Madnani, "Attribute based encryption for scalable and secure sharing of medical records in cloud computing design and implementation," 2013.

[24] M. Mohanty, P. Atrey, and W. T. Ooi, "Secure cloud-based medical data visualization," in *Proceedings of the 20th ACM International Conference on Multimedia*, 2012, pp. 1105–1108.

[25] R. Neame, "Effective sharing of health records, maintaining privacy: a practical schema," 2013.

[26] I. Nwankwo, S. Hanold, and N. Forgo, "Legal and ethical issues in integrating and sharing databases for translational medical research within the eu," in *IEEE 12th International Conference on Bioinformatics Bioengineering*, Nov 2012, pp. 428–433.

[27] L. Seitz, J. M. Pierson, and L. Brunie, "Encrypted storage of medical data on a grid," in *Methods Inf Med*, 2005, pp. 198–201.

[28] Y. Tian, H. Lei, L. Wang, K. Zeng, and T. Fukushima, "A fast search method for encrypted medical data," in *IEEE International Conference on Communications Workshops*, June 2009, pp. 1–5.

[29] P. M. Vieira-Marques, R. J. Cruz-Correia, S. Robles, J. Cucurull, G. Navarro, and R. Marti, "Secure integration of distributed medical

data using mobile agents," *IEEE Intelligent Systems*, no. 6, pp. 47–54, Nov 2006.

[30] P. de Vlieger, J. Y. Boire, V. Breton, Y. Legre, D. Manset, J. Revillard, D. Sarramia, and L. Maigne, "Sentinel e-health network on grid: developments and challenges," *Stud Health Technolol Inform*, vol. 159, 2010.

[31] S. A. K. Mostafa, N. El-Sheimy, A. S. Tolba, F. M. Abdelkader, and H. M. Elhindy, "Wavelet packets-based blind watermarking for medical image management," *The open biomedical engineering journal*, vol. 4, pp. 93–98, 2010.

[32] G. Coatrieux, C. Quantin, F.-A. Allaert, B. Auverlot, and C. Roux, "Watermarking - a new way to bring evidence in case of telemedicine litigation," 2011.

[33] V. Goudar and M. Potkonjak, "A robust watermarking technique for secure sharing of basn generated medical data," in *IEEE International Conference on Distributed Computing in Sensor Systems*, May 2014, pp. 162–170.

[34] H. B. Rahmouni, T. Solomonides, M. C. Mont, and S. Shiu, "Privacy aware access controls for medical data disclosure on european healthgrids," 2010.

[35] G. Haddow, A. Bruce, S. Sathanandam, and J. C. Wyatt, "nothing is really safe': a focus group study on the processes of anonymizing and sharing of health data for research purposes," *Journal of Evaluation in Clinical Practice*, vol. 17, no. 6, pp. 1140–1146, 2011.

[36] K. K. Kim, D. McGraw, L. Mamo, and L. Ohno-Machado, "Development of a privacy and security policy framework for a multistate comparative effectiveness research network," 2013.

[37] H. Lambert and C. F. Leonhardt, "Federated authentication to support information sharing: Shibboleth in a bio-surveillance information grid," *Proceedings of the 18th International Congress and Exhibition*, vol. 1268, no. 0, pp. 135–140, 2004.

[38] S. Lohiya and L. Ragha, "Privacy preserving in data mining using hybrid approach," in *Fourth International Conference on Computational Intelligence and Communication Networks*, Nov 2012, pp. 743–746.

[39] T. Neubauer and J. Heurix, "A methodology for the pseudonymization of medical data," *International Journal of Medical Informatics*, pp. 190–204, 2011.

[40] C. Quantin, M. Fassa, E. Benzenine, D.-O. Jaquet-Chiffelle, G. Coatrieux, and F.-A. Allaert, "The mixed management of patients' medical records: responsibility sharing between the patient and the physician," *Studies in health technology and informatics*, vol. 156, p. 189200, 2010.

[41] H. B. Rahmouni, T. Solomonides, M. C. Mont, and S. Shiu, "Privacy compliance and enforcement on european healthgrids: an approach through ontology," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 368, no. 1926, pp. 4057–4072, 2010.

[42] P. Ruotsalainen, B. Blobel, P. Nyknen, A. Seppl, and H. Sorvari, "Framework model and principles for trusted information sharing in pervasive health," 2011.

[43] M. Jafari, R. Safavi-Naini, C. Saunders, and N. P. Sheppard, "Using digital rights management for securing data in a medical research environment," in *Proceedings of the Tenth Annual ACM Workshop on Digital Rights Management*, 2010, pp. 55–60.

[44] A. Solanas, A. Martinez-Balleste, and J. Mateo-Sanz, "Distributed architecture with double-phase microaggregation for the private sharing of biomedical data in mobile health," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 6, pp. 901–910, June 2013.

[45] D. Weerasinghe and R. Muttukrishnan, "Secure trust delegation for sharing patient medical records in a mobile environment," in *7th International Conference on Wireless Communications, Networking and Mobile Computing*, Sept 2011, pp. 1–4.

[46] P. K. Katarzyna Pasierb, Tomasz Kajdanowicz, "Privacy-preserving data mining, sharing and publishing," 2013.

[47] R. Gajanayake, R. Iannella, and T. Sahama, "Sharing with care: An information accountability perspective," *IEEE Internet Computing*, no. 4, pp. 31–38, July 2011.

[48] T. Tashiro, S. Date, S. Takeda, I. Hasegawa, and S. Shimojo, "Practice and experience of building a medical application with permis-based access control mechanism," in *The Sixth IEEE International Conference on Computer and Information Technology*, Sept 2006, pp. 71–71.

[49] A. Gaignard and J. Montagnat, "A distributed security policy for neuroradiology data sharing," *Stud Health Technol Inform.*, vol. 147, pp. 257–262, 2009.

[50] T. Tashiro, S. Date, S. Takeda, I. Hasegawa, and S. Shimojo, "Architecture of authorization mechanism for medical data sharing on the grid," *Studies in health technology and informatics*, vol. 120, p. 358367, 2006.

[51] Y. feng Jiang, S. yue Zhang, Z. Huang, M. qing Liu, L. Yin, and J. ping Niu, "Access control for rural medical and health collaborative working platform," *The Journal of China Universities of Posts and Telecommunications*, no. 0, pp. 7 –10, 2013.

[52] S. Langella, S. Hastings, S. Oster, T. Pan, A. Sharma, J. Permar, D. Ervin, B. B. Cambazoglu, T. M. Kurç, and J. H. Saltz, "Sharing data and analytical resources securely in a biomedical research grid environment," *JAMIA*, vol. 15, no. 3, pp. 363–373, 2008.

[53] J. Stevovic, F. Casati, B. Farraj, J. Li, H. Motahari-Nezhad, and G. Armellin, "Compliance aware cross-organization medical record sharing," in *IFIP/IEEE International Symposium on Integrated Network Management*, May 2013, pp. 772–775.

[54] M. Kavitha and T. K. Anjana, "Password authentication scheme based on shape and text for secure sharing of phr using abe in cloud," 2013.

[55] A. Ferreira, G. Lenzini, C. Santos-Pereira, A. B. Augusto, and M. E. Correia, "Envisioning secure and usable access control for patients," in *IEEE 3rd International Conference on Serious Games and Applications for Health (SeGAH 2014)*, Rio de Janeiro, Brazil, May 2014.

[56] R. Nithiavathy, "Data integrity and data dynamics with secure storage service in cloud," in *International Conference on Pattern Recognition, Informatics and Mobile Engineering*, Feb 2013, pp. 125–130.

[57] M. A. C. Dekker, "Flexible Access Control for Dynamic Collaborative Environments," Ph.D. dissertation, CTIT PhD.-thesis series ISSN 1381-3617, Number 09-159, IPA Dissertation series, University of Twente, 2009.