ICT for Health Science Research
A. Shabo (Shvo) et al. (Eds.)

doi:10.3233/978-1-61499-959-1-120

International Standard Accident Number: A Master Case Index Linking Accident & Emergency with Medical Data

Hoang Phi LE[a,1], Siegfried HACKEL[b], Andreas GUENTHER[c] , Raimar GOLDSCHMIDT[d], Melhem DAOUD[a] and Thomas M DESERNO[a]
[a] *Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School, Braunschweig, Germany*
[b] *Physikalisch-Technische Bundesanstalt, Braunschweig, Germany*
[c] *Emergency Medical Services Braunschweig, Braunschweig, Germany*
[d] *Municipal Hospital Braunschweig, Braunschweig, Germany*

Abstract. Unconnected data silos make it difficult to view a patient's accident and emergency case in a comprehensive way, i.e. across all sectors. This includes the emergency medical services (EMS), medical data in a hospital's electronic health record (EHR) as well as event data recorders (EDRs) collecting information about the circumstances of the accident and emergency event. In this paper, we propose a conceptual architecture which introduces a novel case-based record linkage approach and the international standard accident number (ISAN) as a master case index for linking data from EDR, EMS, and EHR.

Keywords. identity management, master patient index, record linkage, secondary use of data, health services research, accident research, emergency medicine

1. Introduction

As in other areas of health care, the emergency care chain is characterized by a heterogeneous system landscape and siloed data. Thus, it is difficult to view a patient's accident and emergency case, including his or her treatment and care and the corresponding outcomes in a comprehensive way, i.e., across all sectors. This includes the emergency medical services (EMS) collecting, e.g., rescue data, initial findings, and continuous health monitoring data during the patient's transport. Medical data is recorded in the hospital forming an electronic health record (EHR). In addition to data from the EMS and EHR, data about the circumstances of the event itself can be included for more elaborate analysis. This type of technical or medical data collected even before the prehospital phase potentially becomes more readily available with the growing Internet of Things and sensor-enhanced information systems in the human environment such as smart homes, vehicles and wearables (event data recorders, EDRs).

In this paper, we introduce the international standard accident number (ISAN) and an architecture for linking data from EDR, EMS, and EHR.

[1] Corresponding Author: Hoang Phi Le, Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School, Mühlenpfordtstr. 23, 38106 Braunschweig, Germany, Email: hoang.le@plri.de.

2. Methods

We reviewed existing methods for linking data silos in the medical field. The scope of the review did not include methods for syntactic and semantic integration of data. Issues such as unreliable linkage with demographic data as well as potentials for improvement of demographics-based record linkage approaches taking advantage of available case data in the accident and emergency context were identified. Based on these results, a conceptual architecture for the ISAN system was collaboratively developed with stakeholders of the emergency care chain, such as EMS and hospitals.

3. Master patient index and demographics-based record linkage

In the absence of a standardized unique identification number in many countries, be it for patients or their cases, the common approach to link information in the medical field are master patient indexes (MPI) storing cross-references between local IDs.

The international Integrating the Health Enterprise (IHE) initiative defines the PIX integration profile, which standardizes workflows, involved actors and applicable communication protocols to link patient data from different data sources in the clinical context. An IHE-compliant MPI allows transmission of patient identity information, cross-referencing and usually a unique master ID linked to the corresponding local patient IDs [1]. Following the paradigm of informational separation of powers, identifying data and additional personal (medical or technical) data must be managed by independent organizational entities.

While the PIX profile does not explicitly dictate how the local IDs should be linked to each other, (research) projects or software products implementing the profile usually rely on record linkage with demographic data in order to cross-reference patient IDs from different systems (e.g., E-PIX [2], OpenEMPI[2]).

Record linkage refers to the method of linking data from different data sources that belong to an entity (e.g., person, object, event) using appropriate key variables [3-5]. In the medical field, this entity is often the patient. March et al. identified several software tools (e.g., G-Link, LinkSolv, DataMatch) for patient based record linkage, which solely rely on demographic data [6]. While some demographic data is robust over time (e.g. birth date), other might change (e.g. last name, address). There is less diversity in names in some nationalities reducing its value as an identifying attribute. Names, addresses etc. may be written or abbreviated in different ways (synonyms) or are faultily entered due to phonetic similarities or typographical errors. While there are approaches to take these things into account, such as fault-tolerant probabilistic linkage, the question arises as to whether more reliable data points for linkage can be used. Attributes such as the case's timestamp and location of emergence can be used for more robust linkage using less identifiable information.

4. The ISAN as a master case index

The ISAN system is based on established MPI and related concepts [1-2] and consists of an ID management center (generation and handling of ISANs as globally unique IDs), a trustee (data supplier), data consumers, and the data sources (Fig. 1).

[2] https://www.openempi.org/

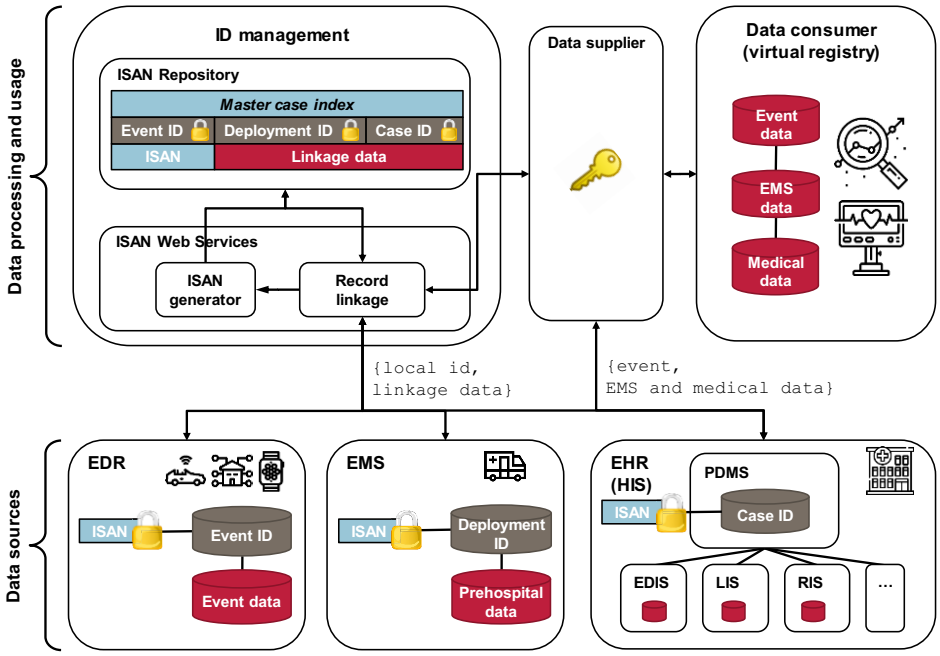

Figure 1. Architecture of the ISAN system.

4.1. Data sources

The data sources typically use proprietary and local IDs for referencing event and treatment data:

- EDR: Smart homes, smart cars, smart wearables, or smart implants will autonomously generate alarms in the near future. The European eCall system is one example.
- EMS: The deployment ID is used to link rescue data, initial findings, diagnoses etc.
- EHR: A variety of data is stored in a hospital information system (HIS) linking the emergency department information system (EDIS) with, e.g., the laboratory and radiology information system (LIS and RIS, respectively) and assigning a case ID in the patient data management system (PDMS).

4.2. ISAN repository and web services

A master index is an agreement upon a unique identifier. The Internet Engineering Task Force (IETF) defines a uniform resource namespace for universally unique identifiers (UUIDs) as a 128-bit character set [7]. According to concepts of the TMF, a German umbrella organization for medical research networks, the ISAN as a UUID corresponds to a first-level pseudonym [8].

The ISAN is requested at the ID management. The corresponding token, which is returned upon a request, is cryptographically secured and – if appropriately implemented in the source systems – allows linked access to data through the trustee. The token itself does not support direct linkage of records. Re-identification is possible for real-time patient care whenever identified information is needed. The ISAN is generated when an

accident or emergency is recorded and then linked to the local ID (or a proxy) of the requesting source system.

The proposed architecture also allows the ex post generation of ISANs for cases that already exist in the source systems. If an ISAN already exists, the case reference between the requesting system to the existing ISAN must be established. For example, in a scenario with automatic eCall, the vehicle (EDR) detects an accident, generates the alarm and requests an ISAN. The local ID and other case identifying attributes (linkage data) are sent to the ID management with the ISAN request. Linkage data such as timestamp and location exist in the systems. However, inaccuracies and deviations in measurement of this data must be considered in the record linkage algorithm. Demographic data is considered for scenarios where case attributes are insufficient for linkage, e.g. if the patient delivers himself to the hospital.

4.3. Data consumers and data supplier

Data consumers are health care providers, research facilities, automotive manufacturers and suppliers, as well as other stakeholders. The data to be productively used in the consumer systems are encrypted, de-identified and passed through an independent trustee, which is representing the data supplier.

5. Discussion

A holistic and temporal view on the accident and emergency case allows stakeholders ranging from healthcare providers along the care chain to vehicle manufacturers and smart home companies to improve their services and products and to create a foundation for data-driven and cross-sectoral research. The relevance of a technological solution for connecting data silos in the accident and emergency domain is reflected in several projects (e.g., [9-10]). Most of them have in common that they do not use a MPI or other standardized means to link data. Either a record linkage algorithm was fitted to the available data or proprietary interfaces in the existing software systems were exploited. This results in isolated methodical solutions that are only applicable within the respective project. Given the lacking availability of a unique identifier, the ISAN offers a standardized way to link accident and emergency data. However, syntactic and semantic interoperability of the data itself are not directly addressed. With increasing universal commitment to support standards such as openEHR[3] or HL7 FHIR[4], those problems might be solved in the future (see, e.g., the German medical informatics initiative[5]).

Compared to common patient-centric approaches for record linkage using only demographic data, our case-based approach offers new possibilities for improved matching. So far, EDR data is rarely considered for linkage. Case data like timestamp and location can be generated automatically avoiding entry and spelling errors, which frequently occur on demographic data.

Linking case data implicitly allows a temporal view on the accident and emergency event from its emergence to its treatment. Our master case index is more differentiated

[3] https://www.openehr.org/
[4] https://www.hl7.org/fhir/
[5] http://www.medizininformatik-initiative.de/en

than MPI. The cases can be mapped to a MPI using demographic data, i.e., a patient is represented by his cases.

It should be mentioned that the conditions for data linkage are more favorable in some countries. A unique identification number in the form of a social security number (civil registration number) is issued, e.g., at birth in Canada, Sweden and Denmark [6].

6. Future work

The architecture proposed in this paper has to be further specified. For example, the minimum set of data for case linkage has to be defined based on the systems analysis of a representative sample of stakeholders. The actual record linkage process using this data has to be specified. In recent years, research into privacy-preserving record linkage approaches has received attention especially in medical informatics [3].

The handling of data by the data supplier and its communication with the data sources and consumer systems needs standardization. An application programming interface for developers and researchers must be provided.

Feasibility can only be explored and value can only be generated by the technology through actual use. Data sources must connect to the system and request ISANs for their recorded events and cases. Finally, applications making use of the potential of linked information must be developed, which is on the road map of the Working Group Accident & Emergency Informatics of the International Medical Informatics Association (IMIA)[6].

References

[1] IHE International. IHE IT Infrastructure (ITI) Technical Framework Volume 1 (ITI TF-1) Integration Profiles Revision 15.0; 2018. Available from: https://www.ihe.net/uploadedFiles/Documents/ITI/IHE_ITI_TF_Vol1.pdf.
[2] Bialke M, Bahls T, Havemann C, Piegsa J, Weitmann K, Wegner T, Hoffmann W. MOSAIC – a modular approach to data management in epidemiological studies. Methods Inf Med. 2015;54:364–71.
[3] Christen P. Data matching: Concepts and techniques for record linkage, entity resolution, and duplicate detection: Springer Science & Business Media; 2012.
[4] Fellegi IP, Sunter AB. A theory for record linkage. J Am Stat Assoc. 1969;64:1183–210.
[5] March S, Stallmann C, Swart E. Datenlinkage. In: Swart E, editor. Routinedaten im Gesundheitswesen: Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven. 2nd ed. Bern: Huber; 2014.
[6] March S, Antoni M, Kieschke J, Kollhorst B, Maier B, Müller G, et al. Quo vadis Datenlinkage in Deutschland? Eine erste Bestandsaufnahme. Gesundheitswesen. 2018;80:e20-e31.
[7] Leach P, Mealling M, Salz R. A Universally Unique IDentifier (UUID) URN Namespace; 2005. Available from: https://tools.ietf.org/html/rfc4122.
[8] Pommerening K, Drepper J, Helbing K, Ganslandt T. Leitfaden zum Datenschutz in medizinischen Forschungsprojekten: Generische Lösungen der TMF 2.0. Berlin: MWV; 2014.
[9] Slagman A, Keil T, Moeckel M. INDEED – Inanspruchnahme und sektorenübergreifende Versorgungsmuster von Patienten in Notfallversorgungsstrukturen in Deutschland. In: Amelung VE, Eble S, Hildebrandt H, Knieps F, Laegel R, Ozegowski S, et al., eds. Innovationsfonds. Berlin: MWV; 2017.
[10] Gimpel H, Regal C, Meindl O, Dormann H, Rashid A, Eder P. Qualitätscockpit in der Notfallmedizin. Management & Krankenhaus kompakt Notfall- und Intensivmedizin. 2018:18–9.

[6] https://www.plri.de/en/forschung/projekte/imia-wg-a-ei