Data Lakes for Pooling T1D Patients CGM Data to Enable EHR Integration and Prevent Data Silos: A Patient Pathway & Clinical Workflow Analysis

Jonathon Dilworth
School of Computer Science
University of Manchester, UK
jonathon.dilworth@postgrad.manchester.ac.uk

***Abstract*—Modern developments in healthcare data collection and management aim to improve patient outcomes. For instance, consider type 1 diabetes (T1D) patients, who benefit from enhanced continuous glucose monitoring (CGM) devices. These devices can provide time-sensitive alerts and facilitate automated data collection for patient-practitioner retrospectives, greatly improving patients' ability to manage their disease. However, CGM vendors often package their devices with proprietary software and rely on third-party infrastructure. This leads to interoperability challenges across clinical systems, fragmented workflows, and increased cognitive burden for clinicians, who grapple with a heterogeneous device landscape encumbered by data silos. This work aims to begin addressing these concerns by first analysing T1D patient data pathways through process modelling. By translating a traditional swimlane (business process) diagram to a formal graph-based representation, we propose a technique for identifying data-flow bottlenecks in complex multi-actor networks. Following this, we discuss an approach to centralised ingestion of CGM data based on a data lake-like architecture (inspired by the Scottish DataLoch initiative). We hope our investigation provides a refreshing perspective on some of the many issues arising from the fragmented nature of the healthcare industry.**

***Index Terms*—Continuous Glucose Monitoring, EHR, Data Lakes, Healthcare Interoperability, Clinical Workflows, Type 1 Diabetes, FHIR**

I. Introduction

In modern clinical practice, technology plays a vital role in effective workflows, helping to deliver beneficial patient outcomes [1]. Within this work, we focus on type 1 Diabetes (T1D), which has seen a surge in the development of medical devices for continuous glucose monitoring (CGM). These CGMs significantly improve patient outcomes by patient and carer alerts, signaling out-of-range incidents and autonomously recording glucose levels for patient-clinician retrospectives [2], [3]. These devices provide patients with improved disease management and help clinicians through providing ambulatory glucose profiles (AGPs). Thus, patients benefit from better care and lifestyle choice guidance [4], [2], [5].

Nonetheless, these advancements do not come without their share of challenges. For instance, CGM vendors often package their devices with proprietary software and rely on third-party computational infrastructure. This introduces problems tied to effective data sharing, such as a lack of interoperability mechanisms with NHS EHR systems. Moreover, the heterogeneous product landscape contributes to increasingly fragmented clinical workflows, disrupting patient experience, complicating data pathways, and leading to the proliferation of healthcare data silos [6].

Within this work, we aim to investigate the flow of clinical data within the context of T1D patient pathways. By focusing on CGM data, we seek to identify process bottlenecks through two data modelling techniques. First, we employ business process (swimlane) modelling in Section. III, then translate this model to a formal graph-based representation in Appendix. A. Drawing from graph theory, actor network theory [7] and computer science, we provide an analysis technique for identifying problematic data-flow bottlenecks in Appendix. A-B. By leveraging insights based on these analyses, we discuss associated challenges (Section. IV), identifying an over-reliance on third-party vendors and their platforms. Finally, we propose a potential solution inspired by the DataLoch [8] initiative: a data lake-like architecture as a scalable foundation for CGM data collection and data silo mitigation.

II. Background

Diabetes Type 1A (T1D) is a chronic disease characterised by the loss of pancreatic beta cells through an autoimmune response, resulting in lifelong insulin deficiency [9]. T1D diagnosis often occurs earlier in life, the disease cannot be prevented, and its cause is thought to be a combination of genetic and environmental factors [10]. Patients can manually mediate their insulin levels through insulin injections, preventing hyperglycemia ('high sugar, presence in blood'). Hypoglycemic (low blood sugar) states are also avoided through blood glucose (BG) monitoring and dietary control [9]. Poor BG mediation can result in co-morbidities related to the kidneys, cardiovascular system and other vital organs.

The financial implications of Diabetes cost the NHS an estimated £6.2 billion per year [11]. This cost clearly shows that improving care could result in significant savings, notwithstanding better patient outcomes and less stress on behalf of healthcare providers.

Advancements in CGM technology have significantly improved patient care, and UK NICE guidelines now recommend CGMs for all T1D patients [2]. As of 2024, the NHS is supplying patients with hybrid closed-loop systems (a CGM combined with an insulin pump), further highlighting the effectiveness of CGM technology in modern diabetes care [12].

III. Patient Pathway: CGM Data Sharing

To better understand issues with CGM-based clinical workflows (e.g. fragmentation, data silos), we describe the data journey for a general-case T1D patient interaction using our swimlane diagram shown in Figure 1, our actor list (Table I) and subsequently discuss key features (subsection III-A). For additional insights, we construct a network-based representation of our actor universe (Figure 7) using methods described in Appendix A-A.

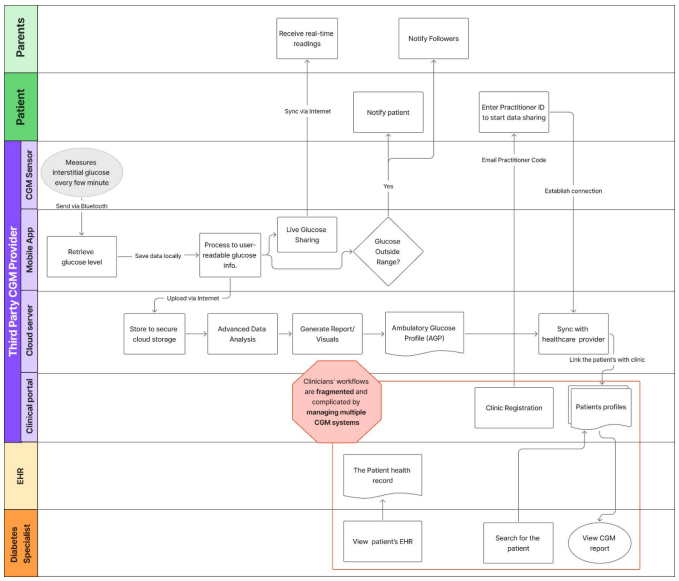

Figure 1. A swimlane diagram *(business process model)* modelling the processes involved in sharing CGM data between T1D patients and clinicians

Through our swimlane, we identify three groupings:

1) patient-related
2) provider-related
3) clinic-related

We describe each in turn below.

A. Patient: CGM Device and Mobile App

The CGM device sensor continuously measures glucose levels and automatically transmits readings to the mobile app every few minutes, collecting a data stream. Daily reading counts can vary but typically range between 96 and 1440, depending on the CGM device and its configuration [13]. The app (or, more specifically, the device it runs on) also performs local data processing, enabling data sharing (e.g. with parents and other followers). This local processing facilitates alert triggering (to multiple potential recipients) when glucose levels are outside the normal range. Syncing collected data with the CGM provider's cloud server also occurs via an internet connection.

Table I
Breakdown of T1D CGM Patient Actor Universe, defined as $\mathcal{A}$

Actor	Type	Grouping
Patient	Human	$\{patient\}$
Parent	Human	$\{patient\}$
CGM	Hardware-Software	$\{patient, provider\}$
App	Software	$\{patient, provider\}$
Receiver (Phone)	Hardware-Software	$\{patient, provider\}$
Server	Hardware-Software	$\{provider\}$
Clinical Portal	Software	$\{provider, clinic\}$
EHR	Hardware-Software	$\{clinic\}$
Clinician	Human	$\{clinic\}$
patient	Individual *(network)*	*patient* $\subseteq \mathcal{A}$
provider	Organisation	*provider* $\subseteq \mathcal{A}$
clinic	Organisation	*clinic* $\subseteq \mathcal{A}$

B. Provider: Cloud Platform

Once CGM data has been synced, advanced analytics can be performed for visualisation and use in report generation (typically an AGP report, see Fig. 2). The clinical portal serves as a third-party (CGM provider) platform, facilitating the management of patient profiles and data-sharing linkage between clinics and patient devices [14]. For instance, a clinician would register with a third-party CGM provider via their cloud platform and receive a practitioner code tied to their clinical ID. The clinician could then share an invitation with a patient, who would be required to verify their practitioner code before the clinic could access their CGM data [15].

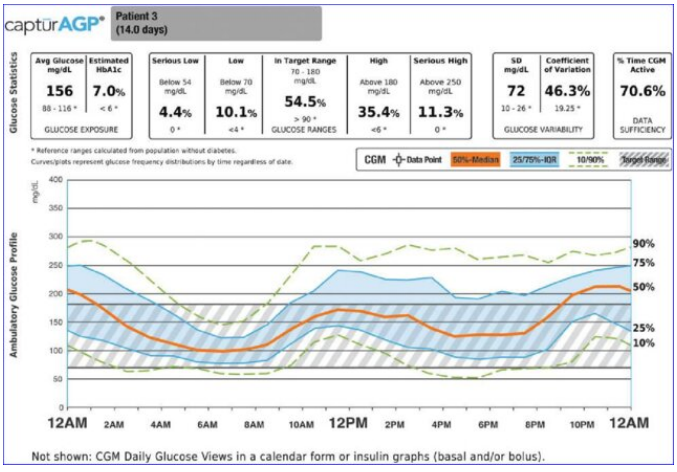

Figure 2. An example Ambulatory Glucose Profile (AGP) report, displaying key metrics such as 'time in range' and 'data sufficiency' [16]

C. Clinic: Clinician Portal and EHR

After establishing patient-clinic data linkage (data sharing), a clinician might review detailed graphs, daily metrics or the AGP (which shows glucose distribution and key data metrics over the last 14 to 90 days [17]) via the provider's cloud portal. They might then opt to print the report, screen-grab sections of it, or manually document their findings in some other medium before manually entering these details into the patient's EHR (where lab results, medication lists and other clinical notes are stored) as attachments or clinical notes.

IV. Insights and Challenges

A. Data Capture

Multiple CGM devices from various manufacturers (Fig. 3, Fig. 4) are available to T1D patients today, each with proprietary hardware and software [18]. This device heterogeneity means that different CGMs capture data at different frequencies with varying accuracies, utilise different file formats, and are packaged with diverse configurations. This factionalised ecosystem creates problems, which essentially boil down to device heterogeneity mapping to data heterogeneity, which makes standardising clinical workflows difficult [19]. For instance, different devices introduce different variance thresholds across data, impacting macro-level CGM data quality and clinicians' ability to interpret data effectively. It has been shown that the mean absolute relative difference (MARD) in accuracy discrepancies between various CGMs is measured between 10% and 20% [5].

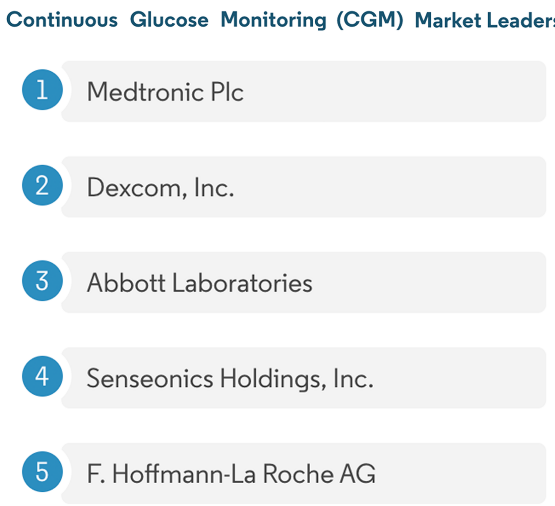

Figure 3. European market leaders in the CGM ecosystem [20]

Additionally, since not all CGMs automatically sync data with their provider's cloud servers (e.g., flash CGMs must be manually scanned), the patient must ensure that their data has been appropriately synced before consultations. Not all patients are technically literate, so minor technical problems can complicate this [3], [21]. These issues contribute to the potential for gaps in their CGM data before key appointments, leading to data completeness issues, compounding data accuracy problems, potentially altering clinical decision-making and negatively impacting patient outcomes [22].

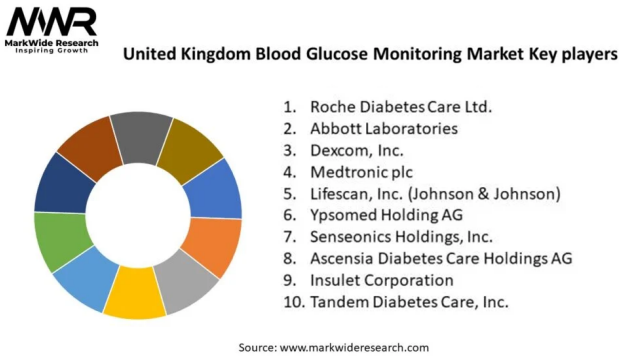

Figure 4. CGM vendor ecosystem in the United Kingdom as of 2025 [23]

B. Data Sharing

The steps enabling CGM data sharing require coordination and can feel clunky, especially when users are not technically literate [24]. Patient data does not miraculously flow from the CGM into the patient's EHR. The healthcare organisation must have an account on the provider's platform, the clinicians must be authorised to access the platform, clinicians must send the patient an invitation to begin sharing, the patient must authorise this request, and so forth. Suppose anything goes wrong during this process (non-delivery of an e-mail or network outages, for instance). In that case, it can lead to patient frustration or even communication issues between the patient and their clinician, and the entire procedure must be repeated [21]. This initial hurdle can often disrupt workflows [19] and frustrate patients during their journey, affecting their perspective of their care.

Once patients have effectively begun sharing their CGM data with a healthcare organisation, clinicians must also deal with a lack of interoperability between third-party vendors' products and the NHS's EHR system [19], [18]. As they typically have multiple patients, they must manage numerous logins and maintain lists of platforms on which their patients are registered. This additional management places increased cognitive load on the clinician, requiring them to switch tasks more often during patient appointments, introducing inefficiencies, delays and, at times, transcription errors, as data must be manually copied from the provider's cloud portal to the NHS's EHR system [18]. Furthermore, the prevalence of all these third-party systems effectively results in data silos [6] and raises additional questions regarding data privacy, security and ownership [18].

C. Human Factors

Whilst a discussion of human factors has been implicit within our discussions of data capture and data sharing, we draw further attention to overall user experience, outlining the need for effective interfaces which make the CGM monitoring and data sharing processes intuitive and accessible to all patients, regardless of their computer literacy [3], [21], [24]. In addition, adequate education will be required for those unfamiliar with the use of computers. However, a greater concern relates to low-income families who cannot afford, for instance, smartphones (a socioeconomic issue that we consider a human factor) that are required for specific models of CGM, as the NHS has come under scrutiny in the past for effectively limiting accessibility to CGM-based healthcare because the patients do not have smartphone devices [25].

Information overload on behalf of clinicians must also be considered when designing a solution that aims to integrate CGM data into EHR systems [26]; if data is dumped into the EHR (i.e. individual CGM readings sent as they are recorded), this could result in an increased cognitive load on clinicians and have the opposite effect to that which is intended (a more seamless experience for clinicians). Staff training and change management should also be considered when introducing new systems into the healthcare organisation, as sudden changes in staff operations often yield negative results and pushback [27] [28].

V. Solutions in Brief

A. Standards: Protecting Against Data Heterogeneity

Manufacturers are encouraged to use standardised outputs for their CGM devices. Whilst building for NHS EHR systems is complex, initiatives such as FHIR [29] make using HL7 (a data exchange standard used by the NHS) more convenient [30]. At least one open-source real-world initiative to use FHIR for CGM data sharing exists: The Argonaut Project [31] (an offshoot of the iCoDE standards project [32]). Though US-centric, their efforts are recognised as a significant example of providing a solution to integrating CGM data into EHR systems. Furthermore, standardised codes such as SNOMED CT [33] and LOINC [34] should label patient glucose data, facilitating consistent data representations throughout the device ecosystem. Typically, failure to adopt these clinical codes is a non-issue since IEEE and ISO 11073 [35] advise their adoption. However, interoperability does not necessarily translate to interpretability [18] (i.e. equivalent encoding protocols do not guarantee semantic equivalency), and, as discussed below, the governing jurisdiction is responsible for mandating the degree to which any manufacturer has to adhere to these standards.

B. Policy: Regulatory Frameworks for Interoperability

Establishing standards (HL7, FHIR, SNOMED CT, LOINC) allows government policy to use regulatory frameworks (e.g. Digital Technology Assessment Criteria [36]) to enforce constraints on CGM manufacturers. For instance, regulators could mandate a specific level of compliance, ensuring the interoperability of CGM devices with EHR systems is straightforward [37]. However, this is infeasible due to the complexities of having multiple vendors for multiple downstream applications (e.g. EHR vendors). Furthermore, regulatory capture may be a valid concern, as a few select companies that meet the burden of regulators may monopolise the market, leading to a lack of innovation [38], [39].

C. Privacy: Asymmetric Encryption in CGM Systems

Since sensitive healthcare data is streamed, batched and uploaded to private cloud infrastructure, legitimate concerns surround data privacy and ownership [40] [41]. We suggest offsetting these concerns through public-private cryptography, protecting data payloads during transit. For example, an NHS-controlled web service could generate a public-private key pair [42] during clinic registration with third-party CGM providers. Clinicians could then share the public key with patients, who register it in their mobile app, securing their data before syncing it with the vendor's cloud server. However, this would result in payloads that must be analysed client-side or by an intermediary NHS service. If not approached correctly, this solution could cause more problems than it solves (e.g. further fragmentation of clinical workflows by presenting non-summarised CGM data to clinicians).

VI. Data Lakes as a Scalable Foundation for CGM Integration and Data Silo Mitigation

The burden of enabling compatibility between CGM devices and the NHS EHR system has been left, for the most part, within the purview of the device manufacturers or private companies to solve. Whilst some hospitals are upgrading their EHR systems [43], [44] (which helps solve this), many cannot perform this transition due to a lack of available resources. Thus, clinicians outside of early pilots within specific ecosystems continue to struggle with multiple third-party portals to access their patient CGM data, and data silos persist. Thus, we offer an approach to reduce data silos and improve the accessibility (and ownership) of CGM data whilst providing new opportunities to researchers.

A. Details of our Proposed Solution

The core idea is to establish an NHS-based initiative for housing all CGM-related data in a single data repository, a data lake. 'Data Lake' is a polysemant often used to describe various architectural patterns tied to data stores. We define our use of the term in alignment with Manco et al. [45], who designed a data lake architecture (Fig. 5) to allow for:

- ingestion of heterogeneous data (streaming and batching)
- data manipulation, processing and analysis
- fault tolerance, high availability
- meeting clinical requirements
- data access patterns

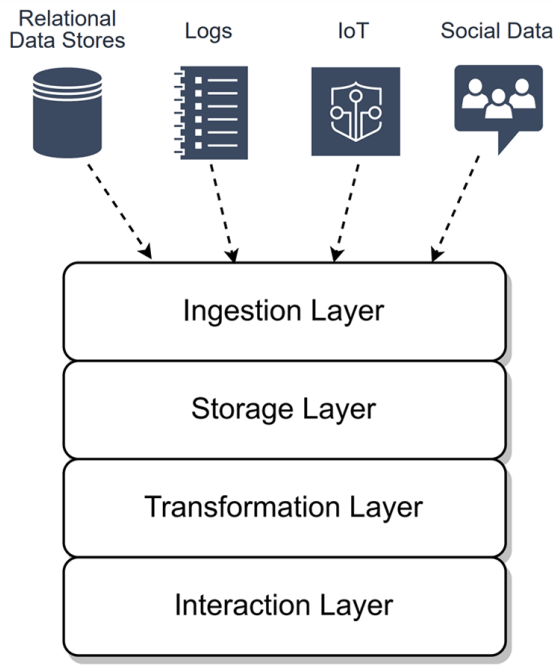

Figure 5. The data lake architecture described in [46], utilised by Manco et al. [45] for their HEALER system, on which our proposed solution is based.

This NHS-based initiative would mandate device manufacturers to push patient CGM data payloads to the data lake and their cloud services simultaneously, establishing a single data repository for all CGM data nationwide. This would allow the NHS to continue to rely on third parties in the short term whilst developing its own set of data pipelines for post-payload processing and integration into existing EHR systems (Fig. 6), mitigating against CGM data silos.

No additional data processing would be required on behalf of the vendor; this is a deliberate design decision to shift the responsibility for enabling EHR integration from vendors back onto the NHS. Effective ETL processes can be implemented to autonomously integrate data from the lake into patient EHRs, resolving the problems of fragmented clinical workflows. As this system continues to mature, the requirement to use CGM vendor's proprietary web portals will diminish (allowing the NHS to resume the role of custodian of its patient's data), transcription errors will be eliminated, and additional research efforts could use this newly available data source. These factors would improve patient outcomes whilst solving many previously discussed issues in CGM-based clinical workflows.

B. Data Lakes in Healthcare

Dagliati et al. [49] discuss multiple existing data lake service providers that support healthcare organisations and augment their efforts in areas such as genomics and epidemiology. They identify important considerations for data interpretability, governance and HCI (web portals).

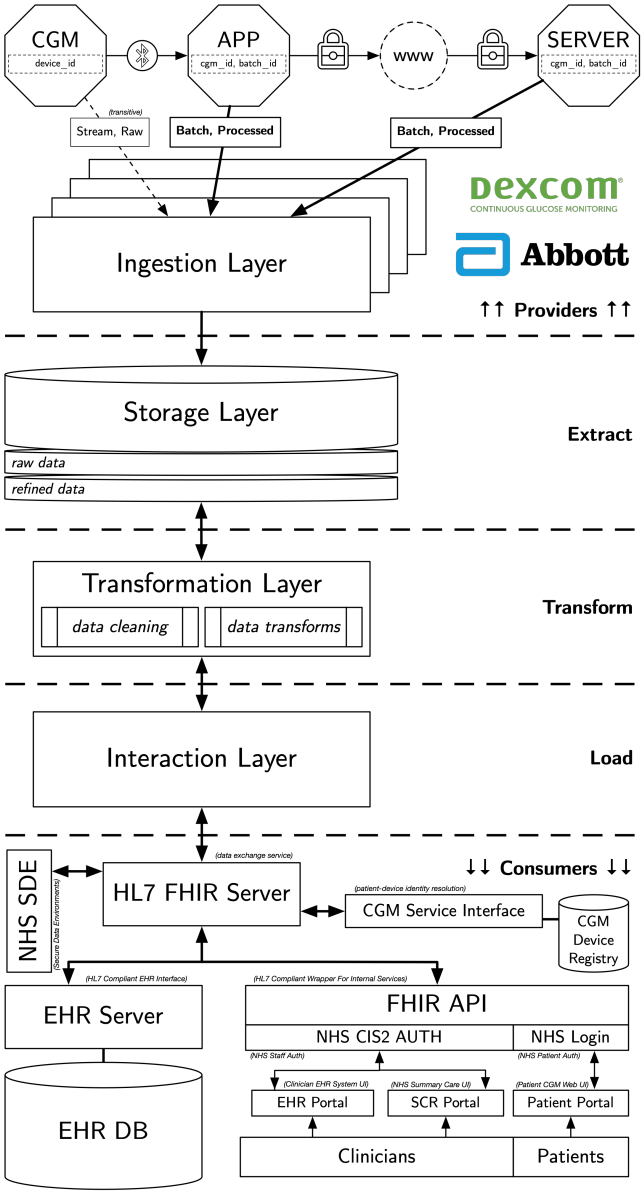

Figure 6. The proposed data lake architecture. Data providers, CGM devices, receivers (APP) and third-party services (SERVER) stream or batch data to the ingestion layer. Separate pools for 'raw' and 'refined' data allow for cleaning, then transformation during ETL. The interaction layer fetches and forwards data (grouped by CGM device id) to an HL7 FHIR data exchange server [47]. The CGM device registry resolves device IDs for CGM-EHR integration. NHS FHIR API and Secure Data Environment [48] facilitate downstream services (data consumers).

Bathani [50] discusses Extract-Transform-Load operations for scalable data lakes in healthcare analytics, claiming that:

> *"Enhanced ETL processes can facilitate timely access to high-quality data, enabling health care professionals to make informed decisions that ultimately improve patient outcomes."*
> *— Bathani (2021) [50]*

Tilala et al. [51] argue that while EHR systems have pro-

duced vast amounts of data for healthcare organisations, these organisations have a distinctive lack of data processing skills, resulting in inefficient workflows. They propose data lakes as a durable and meaningful solution for real-time data processing in healthcare for immediate clinical decision-making.

The most immediately relevant (and analogous) example of a data lake that matches our proposal that is in use today is a Scottish-based healthcare initiative from the University of Edinburgh and NHS Lothian. DataLoch [52], [8] captures data from as many healthcare interactions as possible, then pools and mines it for insights. DataLoch establishes a precedent to demonstrate that this approach is viable, and replicating their approach would be an appropriate first step. The additional components (such as ETL data-provider-consumer pipelines from lake to EHR) can always come later.

C. Stakeholders

NHS IT infrastructure teams responsible for managing NHS login, NHS digital and services such as SPINE [53] would need to be involved from the outset, likely shouldering some of the development efforts and consulting on integration into the larger NHS IT ecosystem.

Standardisation bodies such as HL7 UK (which manages data exchange standards) and FHIR representatives [47], [29] would be key stakeholders in consulting on designing interfaces in producer-consumer data interactions.

CGM device manufacturers involvement will be fundamental to enabling this initiative, as they will be the data providers. However, we anticipate pushback in efforts to protect their market dominance; as such, we suggest numerous potential strategies. First, formal agreements can be drafted to offer subsidisation for providing their efforts in facilitating this initiative. Second, data lake participation may be cited under device procurement contracts. Finally, a regulatory mandate may be enacted.

EHR vendors that provide the systems the NHS operates for patient records will need to work closely with standardisation bodies to build appropriate interfaces that can consume the payloads for clinicians. Their software acts as data consumers positioned at the receiving end of any ETL pipelines.

Clinicians will be key stakeholders in understanding their requirements when designing the ETL operations and the content of the final payloads.

Patient representatives and groups like Diabetes UK [54] must be worked with to ensure that the correct messaging and privacy concerns are effectively communicated. They are entitled to opt out of any procedures regarding the use of their data.

D. Training

Ideally, transitioning to using an alternative data repository (e.g., the data lake) should be seamless; it should primarily act as an alternative data provider. However, we note that EHR integration will result in changes regarding HCI. Patients should also be able to query their data from the lake. Thus, training will be required for clinicians who are expected to interface with patients' CGM data. Educational initiatives for patients must be implemented to ensure they fully understand how to view their data and adjust their data-sharing preferences. IT maintenance staff, engineers and researchers will also be required to undergo training specific to interfacing with data lakes, i.e. understanding OLTP, OLAP and related technical concepts.

E. Metrics & Modalities

According to Bergenstal (2018) [55], 14 days of CGM data provides the requisite data sufficiency for extrapolation over three months, allowing clinicians to gain meaningful insights. Thus, data should be recorded after 14 days of successful CGM measures. It follows that measuring the frequency of writes to the EHR (as writing an AGP summary to the EHR indicates a successful report has been logged) is a valuable metric that measures effectiveness—in addition, comparing successful ETL executions to manual clinician transcription provides a ratio representing utility growth. These measures can be calculated for macro-level performance (i.e. average measures of entire system use) in addition to specific cases that may be sensitive to other constraints.

Performing ETL procedures over an entire data lake is not instantaneous and takes some unit of time, which scales according to the available compute. Since compute is not infinite, and we do not anticipate that the NHS's access to vertical or horizontal scaling is as available as that of CGM vendors operating on cloud-based infrastructure, it is certainly possible that CGM vendors will outperform the proposed approach for data timeliness. As such, we expect to measure latency (or delay) in the delivery of summary AGP reports.

Other metrics we consider for measuring the effectiveness of the proposed solution are the total time saved by clinicians and patient satisfaction scores. Additional metrics to make robust technical assessments are drawn from our formal model, which can be found under Appendix A-B: reliability, mode of transit, sensitivity and exposure.

Rather than simply evaluating the system with quantitative formal methods, we also look to acquire feedback through other modalities: interviewing key stakeholders (clinicians and patients) and providing questionnaires.

Challenges we expect to face during this process relate to human factors, such as individuals concerned about providing honest feedback in fear of workplace reprisals or damaging trust in working relationships (conflict avoidance) [56]. Additionally, positive feedback may be provided due to modern socialisation, social norms and unconscious (confirmation) bias [57]. Assessments should offset these challenges through a mixture of multiple metrics and modalities.

VII. Discussion & Conclusions

By modelling interactions for T1D patients using CGM devices, we have identified several issues with clinical workflows and the patient journey: fragmentation, data privacy, ownership and silos.

The identified cause of these issues stems from the absence of in-house data processing capabilities for streaming patient data to the NHS IT ecosystem and reliance on private companies to provide third-party infrastructure and proprietary platforms. Whilst we consider these issues surprising, they are also worrisome as we question whether these entities are appropriate custodians of patient data. Additionally, we find that existing solutions to data sharing for CGM are inadequate, leading to increased management workload for clinicians and suboptimal patient outcomes.

We believe shifting responsibility for operationalising these procedures back to the NHS is increasingly necessary. By reviewing initiatives such as DataLoch, establishing a framework to accomplish this goal is viable.

We hope our extensive analysis, formal modelling methods, and proposed solution will serve as a basis for future work.

References

[1] R. Wachter, "Making IT work: harnessing the power of health IT to improve care in England," *National Advisory Group on Health Information Technology in England*, vol. 71, 2016.

[2] S. M. Ng, "NICE and NHS England leads the way to improve diabetes care with access to continuous glucose monitoring for people with type 1 diabetes," *BMC Medicine*, vol. 21, no. 1, p. 295, Aug. 2023. [Online]. Available: https://doi.org/10.1186/s12916-023-03014-2

[3] G. Sabben, C. Telfort, M. Morales, W. S. Zhang, J. C. Espinoza, F. J. Pasquel, and K. Winskell, "Technology and Continuous Glucose Monitoring Access, Literacy, and Use Among Patients at the Diabetes Center of an Inner-City Safety-Net Hospital: Mixed Methods Study," *JMIR Diabetes*, vol. 9, p. e54223, Oct. 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11522655/

[4] "Rationale and impact | Diabetes (type 1 and type 2) in children and young people: diagnosis and management | Guidance | NICE," Aug. 2015, publisher: NICE. [Online]. Available: https://www.nice.org.uk/guidance/ng18/chapter/rationale-and-impact

[5] R. Lin, F. Brown, and E. I. Ekinci, "The ambulatory glucose profile and its interpretation," *Med J Aust*, vol. 217, no. 6, pp. 295–298, Sep. 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9805111/

[6] R. Avula, "Overcoming data silos in healthcare with strategies for enhancing integration and interoperability to improve clinical and operational efficiency," *Journal of Advanced Analytics in Healthcare Management*, vol. 4, no. 10, pp. 26–44, 2020.

[7] "Actor-network mapping," Mar. 2011. [Online]. Available: https://weadapt.org/knowledge-base/adaptation-decision-making/actor-network-mapping/

[8] "DataLoch | Case Studies." [Online]. Available: https://ddi.ac.uk/case-studies/dataloch/

[9] P. Lambert and P. J. Bingley, "What is Type 1 Diabetes?" *Medicine*, vol. 30, no. 1, pp. 1–5, Jan. 2002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1357303906006360

[10] K. M. Gillespie, "Type 1 diabetes: pathogenesis and prevention," *CMAJ: Canadian Medical Association journal = journal de l'Association medicale canadienne*, vol. 175, no. 2, pp. 165–170, Jul. 2006.

[11] "Failure to prevent diabetes complications costing £6.2 billion, charity warns - The Diabetes Times." [Online]. Available: https://diabetestimes.co.uk/failure-to-prevent-diabetes-complications-costing-6-2-billion-charity-warns/

[12] N. H. S. England, "NHS England » Hybrid closed loop technology." [Online]. Available: https://www.england.nhs.uk/diabetes/digital-innovations-to-support-diabetes-outcomes/hybrid-closed-loop-technology/

[13] "FreeStyle for Every Need | FreeStyle Libre | Abbott." [Online]. Available: https://www.freestyle.abbott/uk-en/discover/benefitsforme.html

[14] "LibreView Quick Start Guide for professionals." [Online]. Available: https://pro.libreview.io/articles/qsg

[15] "Sharing Patient Data with Doctor or Clinic | Dexcom Provider." [Online]. Available: https://provider.dexcom.com/how-do-patients-share-data-my-clinic

[16] D. Slattery and P. Choudhary, "Clinical Use of Continuous Glucose Monitoring in Adults with Type 1 Diabetes," *Diabetes Technology & Therapeutics*, vol. 19, pp. S–55, May 2017.

[17] L. Czupryniak, G. Dzida, P. Fichna, P. Jarosz-Chobot, J. Gumprecht, T. Klupa, M. Mysliwiec, A. Szadkowska, D. Bomba-Opon, K. Czajkowski, and others, "Ambulatory glucose profile (AGP) report in daily care of patients with diabetes: practical tips and recommendations," *Diabetes Therapy*, vol. 13, no. 4, pp. 811–821, 2022, publisher: Springer.

[18] J. Espinoza, N. Y. Xu, K. T. Nguyen, and D. C. Klonoff, "The Need for Data Standards and Implementation Policies to Integrate CGM Data into the Electronic Health Record," *J Diabetes Sci Technol*, vol. 17, no. 2, pp. 495–502, Mar. 2023.

[19] J. Espinoza, P. Shah, and J. Raymond, "Integrating continuous glucose monitor data directly into the electronic health record: proof of concept," *Diabetes technology & therapeutics*, vol. 22, no. 8, pp. 570–576, 2020, publisher: Mary Ann Liebert, Inc., publishers 140 Huguenot Street, 3rd Floor New ....

[20] "Continuous Glucose Monitoring (CGM) Market Share | Size Analysis, and Growth Trends Report." [Online]. Available: https://www.mordorintelligence.com/industry-reports/continuous-glucose-monitoring-market

[21] T. Hall, M. K. Warman, T. Oser, M. K. Filippi, B. Manning, J. K. Carroll, D. E. Nease, E. W. Staton, and S. Oser, "Clinician-Reported Barriers and Needs for Implementation of Continuous Glucose Monitoring," *J Am Board Fam Med*, vol. 37, no. 4, pp. 671–679, Jul. 2024. [Online]. Available: http://www.jabfm.org/lookup/doi/10.3122/jabfm.2024.240049R1

[22] N. den Braber, C. I. R. Braem, M. M. R. Vollenbroek-Hutten, H. J. Hermens, T. Urgert, U. S. Yavuz, P. H. Veltink, and G. D. Laverman, "Consequences of Data Loss on Clinical Decision-Making in Continuous Glucose Monitoring: Retrospective Cohort Study," *Interact J Med Res*, vol. 13, p. e50849, Jul. 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11325125/

[23] "United Kingdom Blood Glucose Monitoring Market 2025-2034 | Size,Share, Growth." [Online]. Available: https://markwideresearch.com/united-kingdom-blood-glucose-monitoring-market/

[24] M. S. Lanning, M. L. Tanenbaum, J. J. Wong, and K. K. Hood, "Barriers to Continuous Glucose Monitoring in People With Type 1 Diabetes: Clinician Perspectives," *Diabetes Spectr*, vol. 33, no. 4, pp. 324–330, Nov. 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7666603/

[25] T. Thomas, T. T. Health, and i. correspondent, "Deprived families unable to make most of NHS diabetes devices, say charities," *The Guardian*, Dec. 2023. [Online]. Available: https://www.theguardian.com/society/2023/dec/09/deprived-families-nhs-diabetes-devices

[26] E. Asgari, J. Kaur, G. Nuredini, J. Balloch, A. M. Taylor, N. Sebire, R. Robinson, C. Peters, S. Sridharan, and D. Pimenta, "Impact of Electronic Health Record Use on Cognitive Load and Burnout Among Clinicians: Narrative Review," *JMIR Medical Informatics*, vol. 12, no. 1, p. e55499, Apr. 2024, company: JMIR Medical Informatics Distributor: JMIR Medical Informatics Institution: JMIR Medical Informatics Label: JMIR Medical Informatics Publisher: JMIR Publications Inc., Toronto, Canada. [Online]. Available: https://medinform.jmir.org/2024/1/e55499

[27] L. Hughes, "Many NHS staff struggle to use electronic records effectively, report finds," *Financial Times*, Apr. 2025.

[28] C. McCrorie, J. Benn, O. A. Johnson, and A. Scantlebury, "Staff expectations for the implementation of an electronic health record system: a qualitative study using normalisation process theory," *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, p. 222, Nov. 2019. [Online]. Available: https://doi.org/10.1186/s12911-019-0952-3

[29] "FHIR UK Core." [Online]. Available: https://digital.nhs.uk/services/fhir-uk-core

[30] "Health Level Seven International - Homepage | HL7 International." [Online]. Available: https://www.hl7.org/index.cfm

[31] Josh Mandel, "Integrating Continuous Glucose Monitoring with FHIR: Argonaut Project Update (Sept 2024)," Oct. 2024. [Online]. Available: https://www.youtube.com/watch?v=5sReITG-B4w

[32] N. Y. Xu, K. T. Nguyen, A. Y. DuBord, D. C. Klonoff, J. M. Goldman, S. N. Shah, E. K. Spanakis, C. Madlock-Brown, S. Sarlati, A. Rafiq, A. Wirth, D. Kerr, R. Khanna, S. Weinstein, and J. Espinoza, "The Launch of the iCoDE Standard Project," *J Diabetes Sci Technol*, vol. 16, no. 4, pp. 887–895, May 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9264445/

[33] "SNOMED CT." [Online]. Available: https://digital.nhs.uk/services/terminology-and-classifications/snomed-ct

[34] "Home." [Online]. Available: https://loinc.org/

[35] *11073-10407-2022 - IEEE/ISO International Standard–Health informatics–Device interoperability. Part 10407, Personal health device communication–Device specialization–Blood pressure monitor / Institute of Electrical and Electronics Engineers.* New York, USA: IEEE, 2022, backup Publisher: Institute of Electrical and Electronics Engineers, author, issuing body. Publication Title: 11073-10407-2022 - IEEE/ISO International Standard–Health informatics–Device interoperability. Part 10407, Personal health device communication–Device specialization–Blood pressure monitor.

[36] "Digital Technology Assessment Criteria (DTAC) - Key tools and information - NHS Transformation Directorate." [Online]. Available: https://transform.england.nhs.uk/key-tools-and-info/digital-technology-assessment-criteria-dtac/

[37] "Applicability of DCB 0129 and DCB 0160." [Online]. Available: https://digital.nhs.uk/services/clinical-safety/applicability-of-dcb-0129-and-dcb-0160

[38] "Private healthcare market investigation - GOV.UK." [Online]. Available: https://www.gov.uk/cma-cases/private-healthcare-market-investigation

[39] R. Witcomb, "7. In accordance with section 157(3)(c) of the Act, the CMA now releases HCA from the Interim Undertakings."

[40] K. E. Britton and J. D. Britton-Colonnese, "Privacy and Security Issues Surrounding the Protection of Data Generated by Continuous Glucose Monitors," *J Diabetes Sci Technol*, vol. 11, no. 2, pp. 216–219, Feb. 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5478039/

[41] P. Randine, M. Pocs, J. G. Cooper, D. Tsolovos, M. Muzny, R. Besters, and E. Årsand, "Privacy Concerns Related to Data Sharing for European Diabetes Devices," *J Diabetes Sci Technol*, vol. 19, no. 3, pp. 611–619, May 2025, publisher: SAGE Publications Inc. [Online]. Available: https://doi.org/10.1177/19322968231210548

[42] S. Josefsson and I. Liusvaara, "Edwards-Curve Digital Signature Algorithm (EdDSA)," Internet Engineering Task Force, Request for Comments RFC 8032, Jan. 2017, num Pages: 60. [Online]. Available: https://datatracker.ietf.org/doc/rfc8032

[43] M. Burns, "Challenges and successes in implementing an integrated electronic patient record (HIVE) at the Manchester University National Health Service Foundation Trust, England: 1000+ legacy systems, 10 hospitals, one electronic patient record," *Health Information Management Journal*, vol. 53, no. 1, pp. 20–28, 2024, publisher: SAGE Publications Sage UK: London, England.

[44] "NHS England announces Epic EPR integration with NHS App – HTN Health Tech News," Mar. 2025. [Online]. Available: https://htn.co.uk/2025/03/19/nhs-england-announces-epic-integration-with-nhs-app/

[45] C. Manco, T. Dolci, F. Azzalini, E. Barbierato, M. Gribaudo, L. Tanca, and others, "HEALER: A data lake architecture for healthcare," in *2023 Workshops of the EDBT/ICDT Joint Conference, EDBT/ICDT-WS 2023*, vol. 3379. CEUR-WS, 2023, pp. N–A.

[46] T. Hlupić, D. Oreščanin, D. Ružak, and M. Baranović, "An overview of current data lake architecture models," in *2022 45th jubilee international convention on information, communication and electronic technology (MIPRO)*. IEEE, 2022, pp. 1082–1087.

[47] "What Is RDF?" [Online]. Available: https://www.ontotext.com/knowledgehub/fundamentals/what-is-rdf/

[48] "Secure Data Environment." [Online]. Available: https://digital.nhs.uk/services/secure-data-environment-service

[49] A. Dagliati, A. Malovini, V. Tibollo, and R. Bellazzi, "Health informatics and EHR to support clinical research in the COVID-19 pandemic: an overview," *Briefings in bioinformatics*, vol. 22, no. 2, pp. 812–822, 2021, publisher: Oxford University Press.

[50] R. Bathani, "Optimizing Etl Pipelines for Scalable Data Lakes in Healthcare Analytics," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 9, no. 10, pp. 17–24, 2021.

[51] M. Tilala, S. Pamulaparthyvenkata, A. D. Chawda, and A. P. Benke, "Explore the Technologies and Architectures Enabling Real-Time Data Processing within Healthcare Data Lakes, and How They Facilitate Immediate Clinical Decision-Making and Patient Care Interventions," *European Chemical Bulletin*, vol. 11, pp. 4537–4542.

[52] "Home - Continuous Glucose Monitoring v1.0.0-ballot." [Online]. Available: https://build.fhir.org/ig/HL7/cgm/

[53] "Spine." [Online]. Available: https://digital.nhs.uk/services/spine

[54] "Diabetes UK - Know diabetes. Fight diabetes. | Diabetes UK." [Online]. Available: https://www.diabetes.org.uk/

[55] R. M. Bergenstal, R. W. Beck, K. L. Close, G. Grunberger, D. B. Sacks, A. Kowalski, A. S. Brown, L. Heinemann, G. Aleppo, D. B. Ryan, T. D. Riddlesworth, and W. T. Cefalu, "Glucose Management Indicator (GMI): A New Term for Estimating A1C From Continuous Glucose Monitoring," *Diabetes Care*, vol. 41, no. 11, pp. 2275–2280, Nov. 2018.

[56] "Frontiers | Abusive Supervision and Job Dissatisfaction: The Moderating Effects of Feedback Avoidance and Critical Thinking." [Online]. Available: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2017.00496/full

[57] "10 Examples Of Unconscious Bias In The Workplace And How to Avoid Them." [Online]. Available: https://www.easyllama.com/blog/unconscious-bias-in-the-workplace

[58] M. Bradway, M. Muzny, &#197, and E. Rsand, "Use of a Data-Sharing System During Diabetes Consultations," in *Telehealth Ecosystems in Practice*. IOS Press, 2023, pp. 223–227. [Online]. Available: https://ebooks.iospress.nl/doi/10.3233/SHTI230783

[59] A. Albanese-O'Neill, "Roadmap to the Effective Use of Continuous Glucose Monitoring by Diabetes Care and Education Specialists as Technology Champions," *Diabetes Spectrum*, vol. 36, no. 4, pp. 288–298, Nov. 2023. [Online]. Available: https://doi.org/10.2337/dsi23-0006

[60] "Cartesian products of metric spaces." [Online]. Available: https://lavalle.pl/planning/node187.html

[61] B. Milewski, *Category theory for programmers*. Bartosz Milewski, 2019.

[62] "5.1.1 Metric Spaces." [Online]. Available: https://lavalle.pl/planning/node184.html

[63] A. E. Sizemore and D. S. Bassett, "Dynamic graph metrics: Tutorial, toolbox, and tale," *NeuroImage*, vol. 180, pp. 417–427, Oct. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811917305645

APPENDIX A

A. Constructing an Actor Network Model

First, through observing the T1D CGM patient actor table (Table I) we note three distinct *localities*, $\mathcal{L}$, such that:

$$\mathcal{L} = \{patient, provider,\ clinic\} \subseteq \mathcal{P}(\mathcal{A})$$
$$\implies \forall x \in \mathcal{L} : x \subseteq \mathcal{A}$$

We would like to model this swimlane as an actor network (a graph) so that we can provide an alternative view (visualisation) of the network and annotate edges. Thus, we define a graph:

$$G = \langle V, E \rangle, \quad V = \mathcal{A}, \quad E \subseteq V \times V$$

and model data flow from the swimlane as a set of tuples:

$$e_i = (u_i, v_i)\ \ \forall e \in E$$

For additional insights, we consider where the data is stored and how it is transmitted. We define a set of data repositories, $\mathcal{D}$ and transmission labels, $\mathcal{T}$:

$$\mathcal{D} = \{buffer, DB_{phone}, DB_{server}, view, DB_{clinic}, ID\}$$

$$\mathcal{T} = \{stream, batch, alert, manual\}$$

$$transmit : E \to \mathcal{T},\ \ locate : V \to \mathcal{D}\ \ :\ \ E, V \in G$$

Then define: $resolve : Path(G) \to d : d \in D$. This provides a fairly neat alternative formalism to model data flow, whom data flows between (electronic, human) and how data is transferred (streamed, batched, manually transcribed). We simply let $resolve$ map to the identity for human actors, providing the network shown in Fig. 7.

B. An Enriched Actor Network Model: Metrics

While modifying the representation from the previously introduced swimlane (Fig. 1) to our graph (Fig. 7) we were able to identify that the data flows through the clinician, to the EHR **manually** *and presumably, is transcribed*. However, this does not necessarily mean much until we consider how one might operationalise this model *(i.e. how we might apply this model to concretely measure a physical system and identify issues)*.

We recognise several potential metrics during our modelling process (Fig. 1, Table I, Fig. 7) and our review of related works discussed throughout [19], [18], [58], [59], [5], [17] *(noting the arrival of data at nodes, V, reliable forwarding and appropriately security as important in improving related patient pathways)*: 1) frequency of transmission (ω), 2) reliability of transmission (κ), 3) the transfer medium (τ), 4) the sensitivity of the data in transit and, 5) the source and target destination. We provide a definition for each of these metric spaces (as their own subspace) [60] under Table II and a metric space, $\mathcal{M}$:

$$\mathcal{M} = \underbrace{\mathbb{R}_{\geq 0}}_{\omega} \times \underbrace{[0,1]}_{\kappa} \times \underbrace{\{0,1,2\}}_{\tau}$$

Thus, we measure the space at any $e \in E$ through:

$$\mu(e) = \big(\omega(e), \kappa(e), \tau(e)\big)$$

And define binary operations $\forall m \in \mathcal{M}$ between any $(\mu_{src}, \mu_{target}) \leftarrow e : \forall e \in E$ for $\mathcal{M}$ under Table III. Then, for any path:

$$\pi = (e_1, e_2, ..., e_n) \in Path(G), \quad \mu_i = \mu(e_i)$$

This allows us to annotate edges in G with values for ω, κ & τ and simply fold ($\langle \mathcal{M}, \oplus, identity \rangle$) across:

$$\mu(\pi) = \mu_1 \oplus \mu_2 \oplus ... \oplus \mu_n$$

Using an identity:

$$identity = (\infty, 1, 0)$$

This provides the basis for our proposed method for modelling actor networks and calculating potential pain points that could yield suboptimal patient outcomes. Note that the approach is devised through insights drawn from [61], [62], [63].

Allow us to consider assigning rules (shown below Table III) or risk metrics based on custom thresholds. We may then fold over paths in G, bringing potential issues to attention.

Table II
METRIC SPACE DEFINITIONS

Metric	Symbol	Description & Range
Frequency	ω	transmissions per day: $\mathbb{R}_{\geq 0}$
Reliability	κ	probability of delivery: $[0, 1]$
Transfer Medium	τ	$\{stream, batch, manual\} \to \{0, 1, 2\}$
Sensitivity	ϵ_1	has patient identity: $\{0, 1\}$
Encrypted	ϵ_2	**not** encrypted: $\{0, 1\}$

Table III
BINARY OPERATIONS $\oplus$ OVER $\mathcal{M}$

Metric	$\mu_{src} \oplus \mu_{dest}$	Intuition
Frequency	$min(\omega_{src}, \omega_{target})$	minimum ω: bottleneck
Reliability	$\kappa_{src} \times \kappa_{target}$	probabilities multiply
Transfer Medium	$max(\tau_{src}, \tau_{target})$	slowest transfer: bottleneck
Sensitivity	-	-
Encrypted	-	-

A quick note on ϵ_1 & ϵ_2:

$$(\epsilon_1^{src} \wedge \epsilon_1^{target}) \wedge (\epsilon_2^{src} \wedge \epsilon_2^{target})$$

That is:

```
if ( sensitive @ (source,target)
    && !encrypted @ (source,target) ):
      // PROBLEM
    else:
      // we're OKAY.
```

C. Annotated Actor Network Model

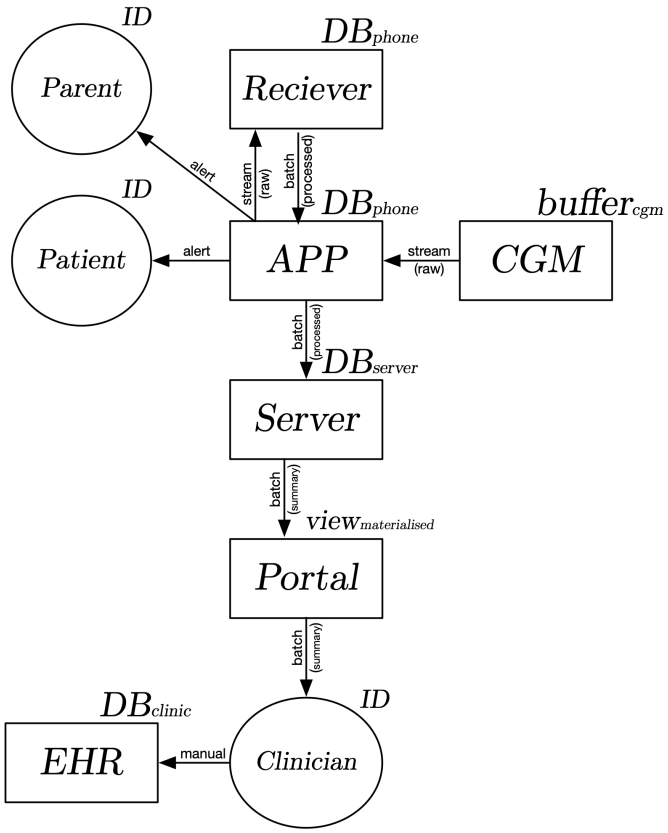

Figure 7. A labelled actor network model for our swimlane