



Information Resilience: the nexus of responsible and agile approaches to information use

Shazia Sadiq¹ · Amir Aryani² · Gianluca Demartini¹ · Wen Hua¹ · Marta Indulska¹ · Andrew Burton-Jones¹ · Hassan Khosravi¹ · Diana Benavides-Prado³ · Timos Sellis^{2,4} · Ida Someh¹ · Rhema Vaithianathan¹ · Sen Wang¹ · Xiaofang Zhou⁵

Received: 21 February 2021 / Accepted: 3 December 2021 / Published online: 16 January 2022
© The Author(s) 2022

Abstract

The appetite for effective use of information assets has been steadily rising in both public and private sector organisations. However, whether the information is used for social good or commercial gain, there is a growing recognition of the complex socio-technical challenges associated with balancing the diverse demands of regulatory compliance and data privacy, social expectations and ethical use, business process agility and value creation, and scarcity of data science talent. In this vision paper, we present a series of case studies that highlight these interconnected challenges, across a range of application areas. We use the insights from the case studies to introduce Information Resilience, as a scaffold within which the competing requirements of responsible and agile approaches to information use can be positioned. The aim of this paper is to develop and present a manifesto for Information Resilience that can serve as a reference for future research and development in relevant areas of responsible data management.

Keywords Information Resilience · Data quality · Responsible data science · Effective information use · Value creation

Amir Aryani, Gianluca Demartini, Wen Hua, Marta Indulska, Andrew Burton Jones, Hassan Khosravi, Diana Benavides Prado, Timos Sellis, Ida Someh, Rhema Vaithianathan, Sen Wang and Xiaofang Zhou are ordered alphabetically with equal contributions

✉ Wen Hua
w.hua@uq.edu.au

Shazia Sadiq
shazia@itee.uq.edu.au

Amir Aryani
aaryani@swin.edu.au

Gianluca Demartini
g.demartini@uq.edu.au

Marta Indulska
m.indulska@business.uq.edu.au

Andrew Burton-Jones
abj@business.uq.edu.au

Hassan Khosravi
h.khosravi@uq.edu.au

Diana Benavides-Prado
diana.benavides.prado@aut.ac.nz

Timos Sellis
tsellis@fb.com

1 Introduction

Advancements in machine learning (ML) and artificial intelligence (AI) are being valued at contributing up to US\$15.7 trillion [122] to the global economy by 2030. Walsh et al. [156] state that ‘AI is enabled by data’ and highlights the need for robust mechanisms for ‘generating, sharing and

Ida Someh
i.asadi@business.uq.edu.au

Rhema Vaithianathan
r.vaithianathan@uq.edu.au

Sen Wang
sen.wang@uq.edu.au

Xiaofang Zhou
zxf@cse.ust.hk

- ¹ The University of Queensland, Brisbane, Australia
- ² Swinburne University of Technology, Melbourne, Australia
- ³ Auckland University of Technology, Auckland, New Zealand
- ⁴ Facebook, New York, USA
- ⁵ The Hong Kong University of Science and Technology, Kowloon, Hong Kong

using data in a way that is accessible, secure and trusted' if advances in AI are to be supported and realised. Multiple industry reports and recent media coverage identify data gone wrong as the biggest risk for AI and other emerging technologies [87]. The impact that bad data products and processes can have on emerging technologies, including but not limited to AI, is recognised not only as a threat from mundane cyber-criminals but, increasingly, from sophisticated well-funded entities, raising concerns of national security [60].

At the same time, despite the scarcity of data science talent [51], it is widely known that data scientists spend over 80% of their time tackling problems related to data access, linkage, and cleaning [19]. Many incidents have occurred due to inadequate handling of data complexity, for example, propagation of biases in data into speech analyses [3], or data discrepancies and algorithmic assumptions that resulted in public harm [54]. Analytical outcomes from data-driven intelligent systems often fall short of expectations due to such issues, or due to time and cost over-runs caused by under-estimation of data curation and preparation needs. Supporting data scientists with the right tools and technologies to overcome these issues in a responsible, auditable, and proactive way is an area of significant need, without which the security and efficiency of data pipelines will remain vulnerable to risks and failures.

Further, data pipelines constructed by data engineers and scientists often lack explainability due to the use of complex computational, statistical, and machine learning techniques [131]. This results in difficulties in creating repeatable data-driven solutions and a lack of trust in the analytical results. Although there are notable recent technical advancements in explainability [93], a practice of transparency needs to be created in data analytics workforce [152]. The lack of alignment within the various stakeholders in the data pipeline is a parallel problem highlighted due to disciplinary, functional, and cultural divides within organisational groups. In particular, this can create a substantial disconnect between how consumers think their data should be handled and how it is actually treated [2]. Whereas several legislative and regulatory frameworks are emerging globally [49], currently most are untested for adequacy, and are creating an increasing compliance burden on organisations with limited technical capability and/or organisational capacity to ensure compliance.

Data re-purposing [175], and the resulting distance between the design and use intentions of the data, is a fundamental issue behind many of these problems. It also presents an unprecedented opportunity for organisations to create new value from existing data assets, which was previously not possible. These new settings demand an extension of the long standing history of contributions to data management by the database research community, through efficient storage and retrieval systems, and well-defined data modelling

and design principles, into embracing the central importance of effective information use [26] from the perspective of a socio-technical organisation.

By bringing together additional perspectives from social scientists and business experts in this paper, we approach the above-mentioned problems from the perspective of organisational information use, arguably signifying a shift in legacy and current research in (responsible) data management. We posit that the organisational information use perspective is necessary to manage paradoxical capabilities that, on the one hand, control, govern, and tightly manage the use of organisational data assets and, on the other, democratise and distribute data and analytics capabilities through automated decision systems [142].

We propose *Information Resilience* as a scaffold within which the conflicting needs of organisational information use can be positioned, and define it as:

The capacity of organisations to create, protect, and sustain agile data pipelines, that are capable of detecting and responding to failures and risks across their associated value chains in which the data is sourced, shared, transformed, analysed, and consumed.

Although Information Resilience has been mentioned previously in the context of risk assessment, incident management, and interoperability [148], we foresee the need for a broader and more integrated approach to Information Resilience to tackle current data management risks and failures in an end-to-end coverage of the data pipeline, from data creation or acquisition to data consumption. It is this end-to-end coverage that is the key issue for managers with responsibility over data and data use. What has been missing to date, in both research and practice, has been a concept that can provide the necessary scaffolding to bridge the ends of the pipeline and integrate the key issues under one central idea. We propose that Information Resilience is a concept that can achieve this aim. In the language of interdisciplinary researchers [58], it is a bridging concept, i.e. 'a concept which combines different ideas and elements related to a specific phenomenon' and offers 'a degree of breadth combined with specificity in focusing on a single phenomenon'.

Our aim is to develop a manifesto for Information Resilience using an empirical lens [129]. To this end, we first present a series of case studies that highlight these interconnected challenges. Subsequently, we analyse the learnings from the case studies to draw insights that can help identify key competencies and functions for Information Resilience for organisations, within which the competing requirements of responsible and agile approaches to information use can be positioned. The major contributions of the paper are summarised as:

- We provide empirical evidence of the socio-technical challenges of effective use of information assets with

three case studies based on projects involving end-users and domain experts.

- We identify five key functions that are collectively critical to achieve Information Resilience, namely responsible use of data assets, data curation at scale, algorithmic transparency, trusted data partnerships, and agility in value creation from data.
- We develop a manifesto for Information Resilience that can help position and drive future research activity in responsible data management.

2 Case studies

Recent years have seen many successes [100] and failures in the adoption of advanced technologies for data-driven decision making. There are a number of factors that influence effective information use—prior studies [143] indicate that these include a variety of people, technology, and process aspects. One of our aims in this paper is to provide empirical evidence of the socio-technical challenges of effective information use. Accordingly, we searched for case studies that represented end-to-end coverage of the data pipeline and were based on empirically validated projects involving end-users and domain experts. Our selection of case studies was also guided by a need to ensure a sufficiently diverse coverage of requirements and challenges that would allow us to identify and explain the variety of functions needed by organisations to adequately build Information Resilience. The three selected case studies provided the required diversity across different types of data, analytical models and algorithms, and stakeholders groups. The cases are presented below and each provides an end-to-end experiential report on how the data was sourced, accessed, and shared, how it was analysed and how the analytical solutions were adopted and used. The analysis of the case studies identified a number of functions, collectively needed for Information Resilience, which are summarised in Table 1, and further discussed in Sect. 3.

2.1 Child welfare services

Violence and maltreatment against children are a worldwide concern. It is estimated that half of all children aged 2–17 years experience some form of physical, sexual, or emotional violence or neglect every year [77]. Child welfare and protection agencies are faced with a difficult task of alleviating the effect of maltreatment suffered by children by intervening with support and services, while also preventing new events of maltreatment from happening by identifying families at high risk. Referral screening is one example of the challenging decisions faced by child welfare agencies, to determine which families are in need of further investigation

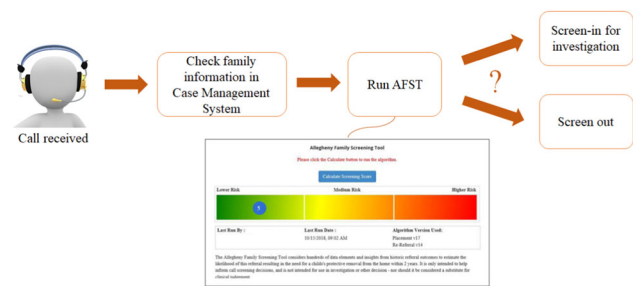


Fig. 1 Decision making supported by AFST

and/or provision of services. Predictive risk modelling has emerged as a useful means of identifying families at high risk of abuse or neglect [30]. Research has also shown that it is possible to use factors known at the time of birth to identify children at high risk of experiencing abuse and neglect [153].

The Allegheny Family Screening Tool (AFST) is a decision support tool implemented at Allegheny County, Pennsylvania in 2016 [30,154]. AFST relies on a predictive risk model that uses a range of input features regarding allegations of child abuse and neglect. It helps workers decide whether a call should be investigated further. When a call is received, a call screener uses a case management system to record the details of individuals (children, parents, alleged perpetrator, etc.) involved in a maltreatment referral. An AFST score between 1 and 20 is automatically presented to the call screener who can make a decision to screen-in or screen-out that referral. The decision-making process supported by AFST is depicted in Fig. 1.

AFST uses existing administrative data concerning children and adults named in a maltreatment referral to generate the risk score. This data is available to Allegheny County through its data warehouse, and reflects records originating from a variety of sources. Input features used to calculate the AFST risk score consider child welfare records, jail records, juvenile probation records, behavioural health records, and birth records. The AFST risk score relies on a LASSO model [150] built on historical data of referrals and placements out of home for around 46,000 calls received by Allegheny County hotline between 2010 and 2014.

With over 4 years of usage, AFST has shown value as an aid for referral screening decisions. For example, the implementation of AFST and associated policies increased accuracy of cases screened-in for investigation, and to reduce disparities in terms of outcomes for similar children from age and race/ethnic subgroups [64]. Below we outline our key learnings and insights from this case study.

A crucial challenge in child welfare is the social licence for using data that is routinely collected for providing support and services to families. There are low levels of trust with regard to the systems used by child welfare services agencies, and this distrust can be projected onto the data

assets [21], resulting in a lost opportunity to utilise the data for supporting those families. Although Allegheny County already had a high-trust relationship with its community, the County did not take social license for granted. Right from the start the community was informed and consulted about the intention to start using their data for building predictive risk models. The families received an explanation of the tool and expected value for the Agency in using it. These community engagements gave the County and the research team sound understanding of the concerns of the community and what measures had to be taken to provide comfort to those families as the tool was implemented.

During the validation and deployment of AFST, it was evident that responsible use of data assets should involve an ethical analysis of the components and implications of an automated decision-making support tool such AFST. The roll-out of this tool included an independent ethical review that provided a set of recommendations and guidelines. The value of an independent ethical review cemented further the importance of ethical management of consent and social licence in the use of data assets for public good [154].

Agency leadership was key for engagement of internal staff at Allegheny County. Via an open call and competitive selection, the Agency showed a strong desire to make better use of its integrated data systems to support decision making, and expressed an openness to machine learning tools providing one potential pathway. The Agency leadership was further evident from existing good data governance practices which facilitated access to all feature data that was housed within the County's integrated data warehouse, while requiring the research team to conduct a broad range of post-modelling analysis to examine how the proposed tool would change call screening decisions. Some of these post-modelling analysis showed that the Agency's practice was resulting in screening decisions that led to screen-outs of almost one-third of children who would have been scored at the highest risk by the AFST; additional data confirmed that these children were subsequently re-referred for maltreatment at very high rates. Meanwhile, the Agency's practice led to almost half the children with a low-risk score being screened-in for investigation, with very few having any investigative findings that would have justified that decision [154].

As a form of external validation encouraged by the Agency before the adoption of the tool, the research team linked the maltreatment referral data to local paediatric hospitalisation records to show that the children who were classified as high risk by the AFST were also significantly more likely to be admitted to hospital for injuries (children who were classified in the highest 5% of risk by the AFST were 20 times more likely to be admitted to hospital for injuries than children who scored in the lowest 50% of risk) [153].

AFST has a continual release of versions as data becomes obsolete or new data becomes available. The continual

improvement requires tight communication between domain experts and model experts to identify, test, and react to potential changes that may affect the ways in which the tool works. Keeping the domain experts in the loop of data identification, curation, and model validation has been instrumental in ensuring the effective use of the tool, and its subsequent impact on practices and processes.

On the other hand, once faced with a new technology, human decision makers such as AFST users commonly want to understand and interpret the inner-workings of the technology. However interpretable models such as LASSO, or state-of-the-art interpretable methods such as LIME [126] and SHAP [95], seem to be far from providing the types of explanations that decision makers in child welfare require. These explanation methods fail to consider co-dependencies between features, relative importance that different features of the model have for decision makers, and explanations that effectively provide humans with more dynamic case-level information that is richer than the information they already have. Since the input features in AFST describe aspects of the families that human decision makers could interpret, for example their demographics, or their history of interaction with child welfare services, among others, decision makers could use these factors to take protective actions especially with high-risk families. The relevance of data literacy thus emerged as an important consideration that was well embedded in the approach for AFST, where the deployment included extensive training for staff by providing them with a basic understanding of predictive risk models, usage of AFST, and the expected changes in practice.

2.2 Adaptive educational systems

Adaptive educational systems (AESs) make use of data about students, learning processes, and learning products to adapt the level or type of instruction for each student. To effectively adapt to the learning needs of individual students, AESs rely on learner models that capture an abstract representation of a student's ability level based on their performance and interactions with the system [45]. The adaptive engine of an AES utilises information from the learner model to recommend items from a large repository of learning resources that best match the current learning needs of a student. These resources are commonly created by domain experts, which makes AESs expensive to develop and challenging to scale [10].

RiPPLE [85] is an AES that takes the crowdsourcing approach of partnering with students, also referred to as learnersourcing [84], to create the resource repository. Figure 2 uses screenshots of the platform to demonstrate some of its main functionality. Figure 2a shows the personalised practice interface in RiPPLE. The upper part allows the students to view their knowledge state as an interactive visualisation widget [5], and the lower part displays learning resources rec-

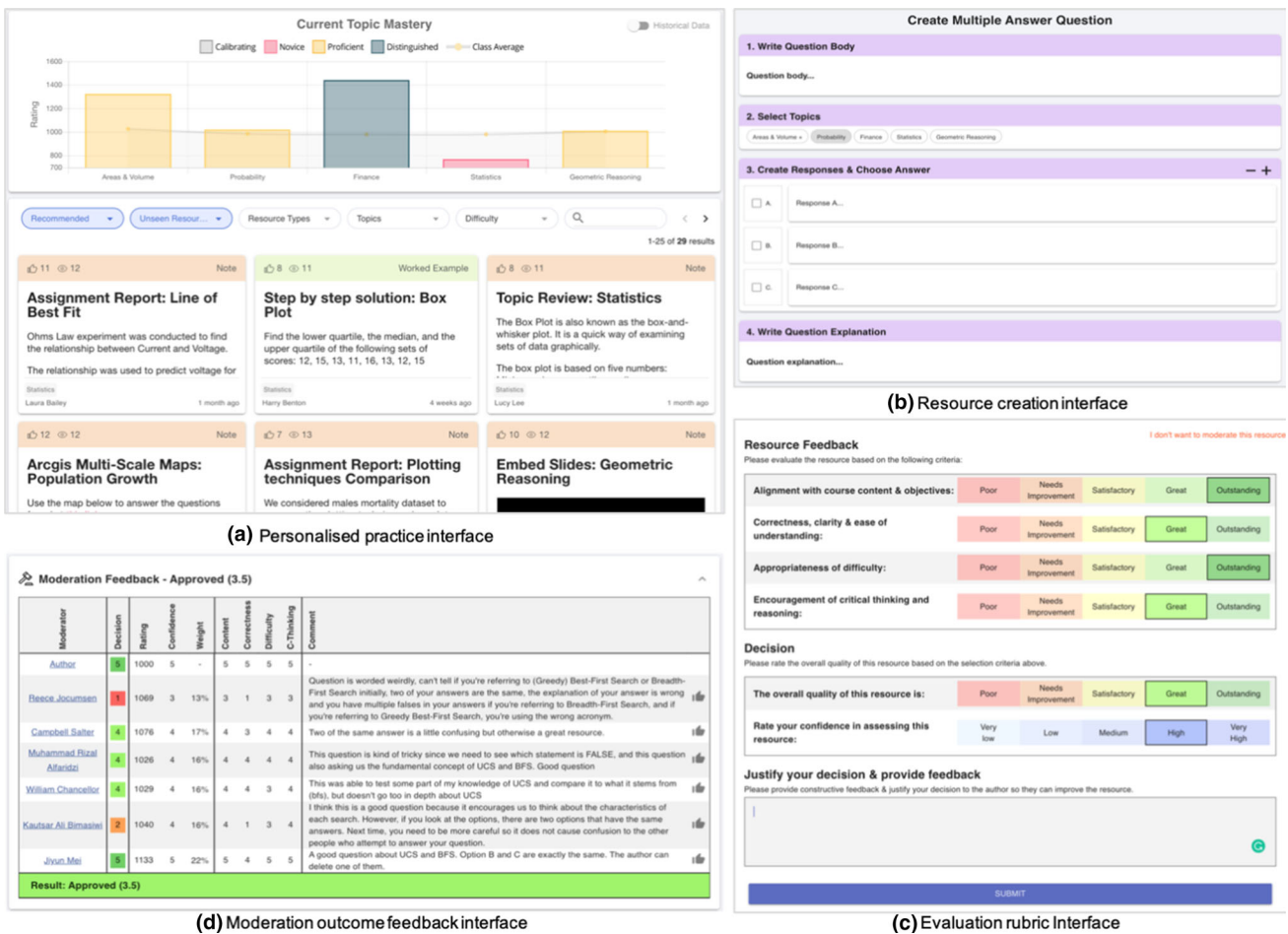


Fig. 2 Four of the main interfaces of RiPPLE

ommended to a student based on their learning needs using the recommender system outlined in [83]. Figure 2b illustrates an example of the interface used for creating learning resources, such as a multiple answer question. To effectively utilise a learnersourced repository of content, there is a need for a selection process to separate high-quality from low-quality resources. RiPPLE uses an evaluation process based on moderation and consensus, where students act as moderators to review and evaluate existing resources, as illustrated in Fig. 2c. Figure 2d shows an example of how evaluations and the inferred outcome are shared with the author, moderators, and instructors. Authors of resources are encouraged to update their resources based on the feedback provided before the resources are added to the course repository.

Advancements in learnersourced AESs are guided by the following three questions. How can learnersourcing systems: (1) accurately and transparently assess the quality of students' contributions? (2) be designed to incentivise a large portion of the student population to offer high-quality contributions?, and (3) empower instructors with actionable and explainable insights to provide oversight? Below we outline

our key learnings and insights from our experience, which spans development and adoption of RiPPLE in over 100 courses across a range of disciplines including Medicine, Pharmacy, Psychology, Education, Business, Computer Science, and Biosciences with roughly 25,000 students who have authored over 50,000 learning resources and over 2 million interactions on these resources. A more in-depth analysis of the case study is available in [84].

RiPPLE uses a consensus approach to automatically predict the quality of a resource. Predictive models have been extensively used in learning analytics tools and have demonstrated promising results in the automatic identification of students in need of assistance [78,97]. At the same time, there are increasing concerns about using predictive models without human oversight in decision-making tasks that affect individuals [71].

A detailed explanation of the consensus approach used in RiPPLE and its strengths and shortcomings are discussed in [39]. Despite our efforts to make the decision-making process accurate, many factors, such as dealing with poor quality or imbalanced data, hyperparameter tuning, and not know-

ing when to retrain the algorithm with new data, may bias or reduce the accuracy of the model. Borrowing from the existing literature on hybrid human–machine information systems [43], we have taken some preliminary steps towards enabling users to provide human judgement such as allowing users to provide feedback on automatic decisions and whether they think the right decision was made.

A common phenomenon, called participation inequality or the 90-9-1 rule [106], has been observed in many systems that rely on users to create content. Participation inequality suggests that 90% of users are lurkers (i.e. observe but do not contribute); 9% of users contribute from time to time and 1% of users participate a lot and account for most contributions. While previous work has reported on challenges related to engaging students in learningsourcing [86,163], implications of participation inequality for learningsourcing systems, and best practices for incentivising a larger portion of students to engage with learningsourcing activities are largely unknown.

Following best practices outlined by the ‘students as partners’ approach [101], RiPPLE introduces mutually beneficial learning partnerships between/within learners and experts for providing high-quality learning at scale. Our approach has focused on utilising students’ contributions and data towards the development of open learner models (OLMs) [23] that capture an abstract representation of a student’s knowledge state. By and large, existing learner models are grounded in psychometrics and approximate a student’s knowledge state solely based on their performance on assessment items. The OLM used in RiPPLE aggregates and builds students’ knowledge state and competencies from assessment, engagement, and learningsourcing activities. This approach requires a significant effort in data preparation (both technically as well as administratively) given the diverse sources from which such multi-modal data can be collected, such as the institutional learning management system, student administrative data, a variety of tools including RiPPLE. However, the extended OLM has demonstrably resulted in creating trust through disclosure, and has consequently helped students monitor and regulate their learning, understand how their involvement in higher-order learning tasks has impacted their learning and promoted their contributions in RiPPLE [4].

RiPPLE relies on explainable recommendations for two tasks: recommending learning resources to students and recommending resources to be spot-checked by instructors. Educational recommender systems commonly operate as a ‘black box’ and give students no insight into the rationale of the recommendation. While the literature suggests that the use of explainable AI (XIA) [157] is not always wanted or necessary [25], the use of machine learning algorithms with black-box outcomes is particularly inadequate for educational settings, where educators strive to provide extensive feedback to enable learners to develop their own vision, reasoning, and appreciation for inquiry and investigation

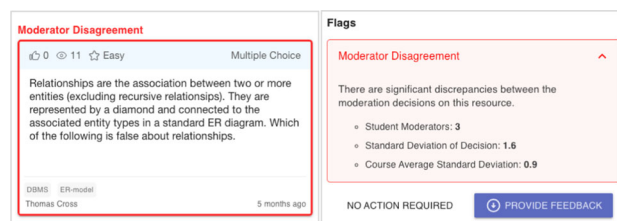


Fig. 3 Spot-checking example

and fairness. In RiPPLE, we have complemented the recommendation engine with a transparent and understandable OLM that provides justification for the recommendations, while showing students their mastery level on each topic. Results from a randomised control experiment suggest that the addition of the OLM to provide justification for the recommendations has had a positive effect on student engagement and their perception about the effectiveness of the system [6].

RiPPLE relies on evaluations from students to judge the quality of the student-generated content. However, this method poses the problem of ‘truth inference’ since the judgements of students as experts-in-training cannot be wholly trusted. Given the limited availability of instructors, RiPPLE incorporates a spot-checking algorithm [160] to identify resources that would benefit the most from being reviewed by an expert. At a high-level, the spot-checking algorithm in RiPPLE employs a range of human-driven metrics (e.g. high-disagreement in moderation evaluations, a high ratio of downvotes in comparison to upvotes) and data-driven metrics (e.g. assessments items that have a low discrimination index or questionable distractors where the popular answer is not the one proposed by the author). When flagging a resource for spot-checking, RiPPLE uses absolute and relative points of comparison to help instructors make sense of the recommendation. Figure 3 demonstrates an example of how a resource being flagged due to moderator disagreement and providing the relative points of comparison has helped to significantly improve the efficiency of expert review process [39,84].

2.3 Illness severity prediction in ICUs

With recent advances in deep learning, significant attention has been directed to development of predictive models to conduct representative tasks in Intensive Care Units (ICUs). This case study is based on the development and deployment of a sentinel system based on interpretable deep models, which analyses multi-variate medical sequential data from multiple sources in Intensive Care Units (ICUs) and predicts illness severity and diagnoses for each critical patient in a real-time manner [28,159]. This monitoring system can provide sufficient evidence for the time-critical decisions required in such

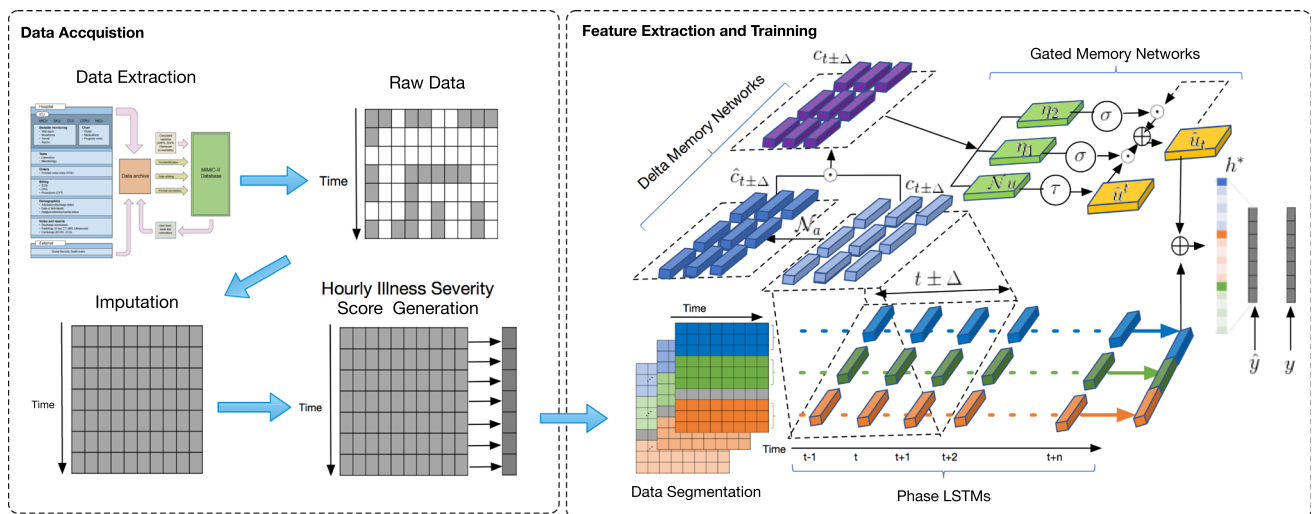


Fig. 4 Workflow of the medical time-series analysis

dynamic and changing environments. For ICU caregivers, the facts and reasoning behind a prediction are the most important criteria when deciding what medical actions to take. The workflow of the system is depicted in Fig. 4. Below we outline our key learnings and insights from this case study.

The source of the medical time-series data was a multi-source, multi-variate sequential database, including bedside sensors, lab tests, and medical notes. In our study, patients under the age of 15 or with ICU stays of less than 24 h are excluded. Each ICU stay instance is then treated as an independent data observation. From these records, we extracted 41 physiological variables from multiple sources and assigned each to six different organ systems according to suggestions by clinical experts. We note that the setup of the database signifies a movement within the health system towards a data-driven mindset. Further, a significant advantage of the database for the study was that it eliminated patient data privacy concerns due to the anonymous nature of patient records. Nonetheless, it had a number of data quality issues that needed to be handled before the data could be used in the prediction model. A number of inconsistencies, errors, and missing values existed in the data. These could be attributed to noise due to instability in the sensors or clerical mistakes in human data entry, such as duplicate records or inconsistent values. For example, some medical doctors used Celsius, while others measured patient temperature with Fahrenheit. Some physiological features were misreported occasionally or were missing. All of these quality issues are detrimental to the down-stream learning and analytical tasks. To improve the quality of the integrated data, we conducted manual data curation to perform data cleaning and data imputation processes, which are rule-based scripts (i.e. SQL scripts) when extracting data from multiple sources, before data integration. Specifically, for features measured by different units,

we explicitly applied unit conversion ensuring the values of a specific feature are in the consistent range. For abnormal and missing features, we used a forward-fill imputation strategy to replace them, assuming they should be same as the last measurements. If the features were not previously recorded, we replaced them with median values of the features. The process of identifying data quality problems and transforming/ preparing the data to enable the subsequent steps was time-consuming and required a mix of automated and manual tasks. Our experience highlights the need for repeatable and efficient methods in the data preparation steps.

Previous works on predictive tasks [118], typically treat all multi-variate time-series variables as an entire input stream without considering the correlations between the physiological variables. However, human organs are highly correlated to each other and to a patient's deterioration. When one or two organs start to malfunction, others tend to follow over a short period. Thus, exploiting correlations between medical time-series variables can further improve classification performance for ICU prediction tasks. On the other hand, feeding all the physiological features from different organ systems into a black-box model to train a decent classifier or a regressor can create hurdles for subsequent interpretations, even though some specific methods (e.g. attention mechanism) could be integrated to provide preliminary explanations on the model's behaviours. This is because the interpretation of a complicated system as a whole is more challenging than the interpretation of a subsystem that has a single function. In the light of this, we designed a new phased LSTM for the medical time series to effectively learn physiological features for each organ. The learning procedure for each organ system is regarded as an independent learning task. To exploit feature correlations that exist across multiple organ systems, we fused all the memories of the

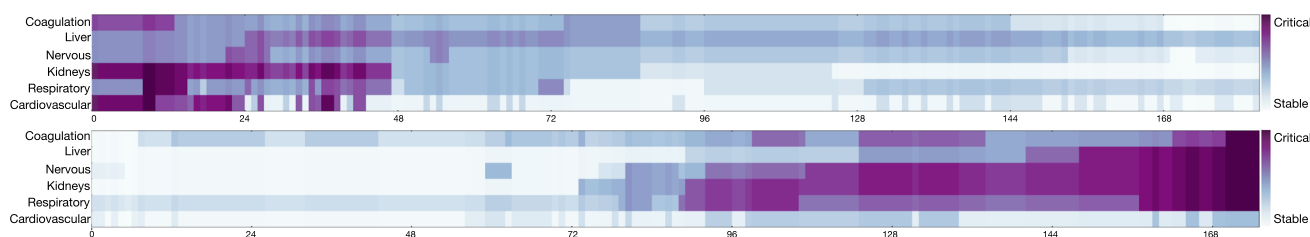


Fig. 5 Fused memories from all single-task LSTMs. The length of stay in hours is plotted on the *x*-axis; the *y*-axis denotes the coefficients in fused memory w.r.t. each organ system. The heatmaps indicate the

severity of the organ system's condition. The two panels show the different journeys of Patient A (ID: 80030) (top) and Patient B (ID: 45767) (bottom)

multi-task LSTMs and captured the cross-tasks interactions. As some feature correlations between multiple organ systems are asynchronous (e.g. a deterioration in the fraction of inspired oxygen can asynchronously affect cerebral blood flow), we added a parameter in the memory fusion mechanism [170] to consider how much temporal information should be included. It is worth noting that we can enlarge this parameter to exploit asynchronous feature correlation across tasks over a longer temporal range. Last, we adopted attention mechanism in the framework to learn non-uniform weights across different organ systems so as to optimise the ultimate learning target—illness severity prediction.

To demonstrate the interpretability of the system, we visualised attention weights in the framework for two patients in Fig. 5. The output scores show the attention variations for the two patients over time. A darker colour signifies a worse condition for the specific organ system. In Patient A's first 48-h stay, our sentinel system made alerts for the malfunctions in the kidney and cardiovascular systems, even though the overall SOFA scores were not so high as to draw attention from medical staff. The consequences of this situation were reported by the doctor in the radiology report made 72 h later (noteevents.row_id=1200211): 'evaluate for obstruction causing acute renal failure'. With Patient B, our system paid the most attention to the respiratory system in the early days of admission. After 72 h, the doctor diagnosed 'respiratory failure' in the medical report (noteevents.row_id=105617). Medical attention to the respiratory system was dramatically increased after 80 h, which was recorded in the medical report as 'Large right pleural effusion is increasing' after 144 h. From the above results, our system showed sufficient capacity as a highly accurate early warning system to clinicians. Beyond a prediction, the system provides an explanation and an indication of which organ systems are causing problems well before traditional forms of monitoring signal an issue. Initially, clinicians and caregivers were reluctant to investigate an early warning from an intelligent model as the interpretation of the signal was hard to understand. The important prediction results were at risk of being ignored, causing irreversible consequences

for patients. With our interpretable results, clinicians could narrow down the problem which assisted in timely decision making.

3 Functions of Information Resilience

We have proposed Information Resilience as a means of unifying the conflicting demands on organisations in their efforts to create value from their data assets. The need to bring together perspectives from social scientists and business experts alongside computer scientists and data management researchers is evidenced in the presented case studies. The analysis of the case studies surfaced a variety of required functions and competencies as summarised in Table 1. We note the emergence of five clusters of functions that can faithfully capture these diverse socio-technical challenges, namely (1) Responsible Use of Data Assets, (2) Data Curation at Scale, (3) Algorithmic Transparency, (4) Trusted Data Partnerships, and (5) Agility in Value Creation from Data. In this section, we synthesise our learnings from the three case studies and elaborate on each of the function clusters by providing a summary of the state-of-art as well as open questions.

3.1 Responsible use of data assets

We acknowledge that in the current data landscape, the use of data is often disconnected from its creation. As a result, the design and use intentions are no longer aligned. A failure to guarantee that data is used and exploited for the right (or at least intended) purposes can have disastrous consequences. It is also putting increasing pressure on both private and public sector organisations to ensure that their data is protected from misuse. Several recent incidents have highlighted these consequences, such as the Cambridge Analytica scandal faced by Facebook, wherein a failure to protect the use of their collected data led to a crisis heavily reported in the media and with economic and reputation implications for the company [149]. In the public sector, Robodebt, a scheme

Table 1 Summary of functions identified from case studies

	Child welfare services	Adaptive educational systems	Illness severity prediction in ICUs
Responsible use of data assets	<p>Social Licence: Community engagement and ethical review</p> <p>Purposeful Analytics: Agency leadership to support data-driven decision making</p> <p>Data Literacy: Training staff about predictive risk models and expected changes in practice</p>	<p>Purposeful Analytics: Institutional strategy for improving the experiences and learning outcomes of large online student cohorts</p> <p>Data Literacy: Use of open learner models (OLMs) to help students understand their academic strengths and weaknesses and better regulate their learning</p>	<p>Social Licence: Community acknowledgement of evidence-based medical practice</p> <p>Purposeful Analytics: Use of health data within ICUs for prediction of length of stay, deterioration, and mortality</p>
Data Curation at Scale	<p>Human Oversight: Domain experts in the loop of new data identification, curation and model validation</p> <p>Repeatable and Verifiable Data Curation: Streamlined feature data extraction from integrated data warehouse</p>	<p>Repeatable and Verifiable Data Curation: Embed RiPPLE in institutional systems for streamlined data flow</p> <p>Human Oversight: Use of human judgement to build authentic consensus on quality of learning resources; Spot-checking by experts to improve accuracy of moderation</p> <p>Data Worker Behaviour: Using (OLMs) to understand behaviour in content creation</p>	<p>Understanding Impact of Data Curation: Inconsistencies in unit measurement and missing physiological features</p> <p>Repeatable and Verifiable Data Curation: Data transformations on anonymised data</p> <p>Human Oversight: Clinicians' decision making amplified (not replaced) with an early warning system</p>
Algorithmic Transparency	<p>Assurances and Explainability: Exploration of different approaches to explainability between researchers and experts, accompanied by training and improved data literacy</p>	<p>Fairness and Transparency: Use of best practice in students as partners approach</p> <p>Assurances and Explainability: Explainable recommendations to students (with OLMs) and instructors (with spot-checking indicators)</p>	<p>Assurances and Explainability: Visualisation of attention weights to explain model behaviour by providing the learned medical insights</p>
Trusted Data Partnerships	<p>Data Discovery: Ease of access to the county's integrated data warehouse</p>	<p>Data Discovery: Access to institutional datasets and platforms</p>	<p>Privacy and Security: Anonymised patient data</p>

Table 1 continued

	Child welfare services	Adaptive educational systems	Illness severity prediction in ICUs
	Data Governance: Existing data governance facilitated access to feature data	Data Linkage and Provenance: Linking of multi-modal data from learning systems, tools, and student administrative databases	Data Discovery: Access to multi-source database
			Data Linkage and Provenance: Linking of multi-variate sequential data from multiple medical data sources
Agility in Value Creation from Data	Organising for Data-driven Work: Regular communication between call centre users, leadership staff, and model experts	Organising for Data-driven Work: Support by developer team for instructors, students, and learning analytics stakeholders in day-to-day use of RiPPLE	Organising for Data-driven Work: The need for interpretability from clinicians was prioritised by the developer team
	Integrating Analytics with Processes and Products: Managing impact on practices and processes through authentic engagement with domain experts	Integrating Analytics with Processes and Products: Integrating algorithmic insights and recommendations to augment users' capabilities	Integrating Analytics with Processes and Products: Integrating the explainable insights into the predictions to assist medical experts' critical and urgent decisions
	Effective Data Use: Leadership support for effective use of AFST tool	Effective Data Use: Adoption of RiPPLE in 100+ courses	Data Value Capture and Measurement: Early warning system reduced time to action for clinicians
	Data Value Capture and Measurement: Continuous measurement of the impact of decision-making support tool	Data Value Capture and Measurement: Creation of over 40K learning resources by students and measurable improvement in learning outcomes	

launched by Department of Human Services for fraud prevention and debt recovery, considered a policy fiasco [164] and exposed serious consequences of ill-defined data matching across government agencies coupled with reduced human oversight.

Social Licence Based on our case studies, we stipulate that even though data may be secured and the right privacy measures are taken (functions of the Trusted Data Partnerships cluster), organisations, especially public sector departments, still need to make sure they have the social licence to use data for specific purposes, raising the question of *how to effectively engage with the relevant stakeholders*. Social License to Operate (SLO) is well recognised in corporate knowledge and refers to 'the perceptions of local stakeholders that a project, a company, or an industry that operates in a given area or region is socially acceptable or legitimate' [124]. We note that having a social license is not a legal obligation

but an ethical one and generally refers to the 'acceptance granted to a company or organisation by the community' [147], for example the acceptance of evidence-based medical practice was observed in our ICU case study. At the same time, a lack of social licence can result in significant community dis-satisfaction about the use of data [52,108]. An open challenge is *how to resolve conflicting interests* from multiple stakeholders. In AFST case, Allegheny County made significant effort to inform internal and external stakeholders via community meetings, where external stakeholders included advocacy groups, service providers, court staff, and consumer groups. These meetings included discussion of sensitive topics, and an evaluation of the process showed that stakeholders considered their involvement as an indication of transparency [14] and resulted in stakeholder support through the lifetime of the project.

Nevertheless, social licence is highly contextual. Decisions as to who to involve depend on the context, and require a multi-faceted approach, i.e. engaging stakeholders, through an iterative process [91]. A participatory/democratic approach that puts emphasis on people and communities engaging directly with developers and governments in the social licensing process could be a means to ensure social licence [111], but implies that organisational mechanisms for encouraging interaction between developers, researchers, public sector and the society should be established [111], along with external independent oversight [91]. Further research is needed to investigate *if a participatory approach is the right framework*.

Purposeful Analytics In addition to community and external stakeholder acceptance, we note the importance of internal acceptance and a shared understanding of purpose and outcome. The importance of this function surfaced in the Child Welfare case study where the Agency had a clear internal purpose and understanding of how AFST would support their existing business processes and what outcomes they expected in terms of increasing the accuracy of call screening decisions [14]. Similarly, the institutional strategy for improving student experiences and outcomes in the Adaptive Educational Systems case study, as well as the accuracy and timeliness of the warning system in the Illness Severity Prediction case study, defined a clear purpose for the development of analytical approaches. These experiences highlight the importance of *organisational leadership and agency towards purposeful analytics*. Purposeful analytics is well-known in business circles to appropriately direct resources and investment [41]. However, there is a parallel imperative for the organisation's leadership to change the narrative from what *can* be done to what *should* be done when developing their data analytics agenda. This raises the question of *what an analytics agenda looks like*, as well as *how internal capability can be enhanced and maintained to respond to the analytics agenda*.

Data Literacy A clear internal purpose would not be possible if literacy around the use of data and analytical models is not developed internally. Data literacy is relative to the organisational teams and capabilities as well as the analytics agenda, and hence it is challenging to *define data literacy objectives and needs*. Communication and collaboration between human decision makers and tool development teams are fundamental for achieving decision making that leads to optimal results [67]. For example, call screening staff at the Child Welfare Agency received training on how to use AFST that included, step-by-step explanation of how the process supported by the tool would work and exploration of cases to demonstrate how the risk score was associated with some of the factors typically observed during call screening. Subsequent evaluations revealed the benefits of the training for staff engagement and accurate decision-making processes

[14]. Whereas the benefits are obvious, the *mechanisms by which these collaborations can be encouraged and realised* remain an open challenge. We also note that while human-in-the-loop of algorithmic tools is an intensive area of research [171], how humans and algorithms interact and achieve better decisions is context dependent [69], as evidenced by the use of OLMs in promoting self-regulation in Adaptive Educational Systems. This issue is particularly severe in systems that provide scores or rankings, e.g. a tool that scores a candidate's job performance based on body language and speech patterns should be identified by (savvy) users as 'fundamentally dubious' [105]. An initial step towards a successful adoption of analytical tools and solutions requires a level of *data and algorithmic literacy that can help users understand issues of fairness and bias* [62,142].

In summary, we reiterate the importance of the organisational information use perspective in responsible data management. We identify three stand-out functions for organisations to consider, i.e. social licence, purposeful analytics, and data literacy, that can create and support capacity for responsible use of data assets through principled approaches. Table 2 summarises the current literature and open challenges in each function.

3.2 Data curation at scale

Data curation is a multi-faceted problem that spans economic and incentive models, social structures, governance and standards, organisational culture, technological support, and workforce development. It is a key step to ensure data being used to power data-driven applications is fit for purpose [128]. Data curation may include a large variety of activities such as format transformations, de-duplication, identification of illegal or missing values, or entity resolution. Not surprisingly, it is being increasingly recognised that the task of exploring, curating, and preparing data to make it analytics-ready takes up to 80% of data workers time [104,145]. In particular, we point to three areas of data curation at scale, *viz.* data quality discovery, information extraction, and data linkage, that have been noted in current literature as being particularly demanding .

Data quality has been an area of research for over 2 decades [128], with contributions from computer science, statistics, information systems, and respective domain areas. It has been widely acknowledged that one cannot manage data quality without first being able to measure it meaningfully [90]. Therefore, discovering the quality of a dataset is a fundamental task in most, if not all, data-driven projects [18]. Quality of data is typically assessed against certain stated requirements [80], which are elicited from data users. In the current data landscape, users of such re-purposed data lack knowledge of the quality requirements. The body of knowledge on how to evaluate the quality of datasets that exhibit

Table 2 Functions relating to responsible use of data assets

Functions	Relevant cases	Selected refs	Open challenges
Social Licence	Child Welfare Services; Illness Severity Prediction	[91,111]	How to engage stakeholders? How to resolve conflicting interests? Does a participatory approach provide the right framework?
Purposeful Analytics	Child Welfare Services; Adaptive Education Systems; Illness Severity Prediction	[14,41]	How to achieve agency and organisational leadership for purposeful analytics? What does an analytics agenda look like? How to enhance and maintain internal capability for responding to analytics agenda?
Data Literacy	Child Welfare Services; Adaptive Education Systems	[14,41,111]	How to define data literacy objectives? How to encourage collaboration of developers and end-users? How to exploit data literacy to tackle fairness and bias?

characteristics typical of re-purposed data [175] is critically lacking [33], and hence the task of data quality discovery and evaluation of the fitness for use of a given dataset for a defined analytical purpose remains highly challenging.

There is an abundance of data types, i.e. structured (relational records, XML) and unstructured (texts, trajectories, images). To prepare data for specific analytical tasks and models, a number of information extraction tasks [132] have to be undertaken that aim to extract structured information (e.g. entities and their relationships) from unstructured data. Existing approaches to information extraction are largely ignorant of the quality of data, assuming that the original data sources are reliable, standardised, and well-represented. A few recent attempts [46,119,161] have investigated noise-aware information extraction by examining the heterogeneity, ambiguity, uncertainty, and untrustworthiness of data. A number of issues are being tackled by current research to improve information extraction, in particular: 1) Data informality where data sources (especially those collected from Web documents) could be full of abbreviations, synonyms, grammar mistakes, and lack of sufficient contextual signals for information extraction [79], and thus even very fundamental text processing operations would encounter a significant performance degradation when applied on these informal data. This calls for new extraction techniques to be developed for these abundant yet informal data; 2) Temporal evolution is where associate knowledge (e.g. a patient's

health status, a user's friendship network) keep evolving over time, making it necessary to incorporate the temporal information into existing frameworks for information extraction; 3) Efficiency is a key consideration where the size of real-life datasets could easily reach TB-level. Moreover, when temporal evolution of data is considered, it is no longer sufficient to examine only a single snapshot of data. This dramatically increases the amount of data that needs to be processed in practice, posing great efficiency challenges on information extraction methods and algorithms.

Data linkage is the process of identifying, matching, and merging records that correspond to the same real-world entity from several datasets or even within one dataset. Admittedly, it is a fundamental step to make data useful as information about an entity can be enriched from different sources if they are reliably and efficiently linked together. It is relatively easy when common identifiers exist in the datasets. In practice, however, data is usually anonymised (i.e. identifiers removed) by third-party data-holders for privacy concerns, making it necessary to conduct entity linking (also called entity resolution, entity alignment, or record linkage) which matches multiple mentions of the same entity in different data sources [31]. Similar demands can also be found for other types of data, especially unstructured data (e.g. text data, multimedia data, spatiotemporal data, etc.). For instance, the user tracking services can benefit from linking a person cross different surveillance cameras [172,176], a user with multi-

Table 3 Functions relating to data curation at scale

Functions	Relevant cases	Selected refs	Open challenges
Understanding Impact of Data Curation	Illness Severity Prediction	[61,76,120,128,175]	How to measure quality when datasets are re-purposed? How to measure data fitness for purpose? How to best perform quality requirement elicitation?
Repeatable and Verifiable Data Curation	Child Welfare Services; Adaptive Education Systems; Illness Severity Prediction	[27,42,43,59,61,98,141]	How to design human-machine hybrid data curation methods? How to design evolution-aware data curation methods? How to perform data curation over anonymised data?
Human Oversight	Child Welfare Services; Adaptive Education Systems; Illness Severity Prediction	[32,43,104,123,129]	What is the role of human oversight in data curation processes? How is human intervention best deployed to improve data curation outcomes?
Data Worker Behaviour	Adaptive Education Systems	[27,72,73,104,114]	How do different data workers perform data curation differently? How to leverage data worker behaviour to estimate data quality?

ple phone numbers or social media accounts [29], or a taxi driver registered with different companies [81,82]. Similarly, the multi-source multi-variate time-series data of each patient needs to be combined in order to make accuracy predictions in the ICU case study. The challenges from these domains include how to cope with the loose structuredness, extreme diversity, high speed, and large scale of entity descriptions used by various applications [31]. Performing data curation activities at scale remains an open research challenge. Through our case studies, we identified four functions of data curation at scale, as discussed below (Table 3).

Understanding the Impact of Data Curation The widespread re-purposing [175] of data is a root cause of many data curation problems. In practice, there are three main categories of approaches to data curation, as data workers aim to prepare data for analytics, namely, ad hoc/manual [76,120], automated [68], and human-in-the-loop [89,113]. Manual approaches to data curation are still the predominant choice for data workers [15] and usually do not facilitate building well-defined processes for required data curation activities, and, hence, the impact of data curation (particularly transfor-

mations) on the quality of the data [120] remains unclear, as observed in the Illness Severity Prediction case study with the impact of inconsistent and missing data. Thus, it can lead to several potential issues, including bias effects brought in by the data worker during their analysis process, the problem of data reusability brought in by lack of transparency and documentation of the analysis or generative processes, and scalability and generalisability across different datasets and use cases [61,76].

It then becomes important to understand *how to measure quality when datasets are re-purposed* and, thus, define *measures of data quality that are a function of the intended use*. Additionally, there is a need to *collect and understand data quality requirements* relative to the intended use of the data. We thus envision moving away from objective and absolute data quality metrics to more purpose-specific definitions of data quality dimensions to improve understanding of the impact of data curation.

Repeatable and Verifiable Data Curation To overcome the adhocism and lack of repeatability of current manual practices, as well as issues with transparency and explainabil-

ity of automatic methods, human-in-the-loop data curation has increasingly become popular across different domains [48,98]. It is based on the assumption that certain tasks can be performed more effectively by humans compared to algorithms, and, hence, human-in-the-loop systems (e.g. [42,59,89]) aim at leveraging both the scalability of machine-based data processing as well as at increasing system effectiveness by selectively involving humans in the difficult cases that algorithms struggle with [43].

The related research community has been looking at how to address the lack of repeatability and verification of outcomes of data curation processes and how to increase the limited scalability by means of human-machine collaboration [43], in particular how to effectively outsource certain data curation tasks to a crowd of online data workers. Some successful examples of human-machine hybrid systems and crowdsourcing for data curation include Data Tamer [141], ZenCrowd [42], CrowdDB [59], and Qurk [99].

We note from all three case studies that repeatable and verifiable data curation was an important function and highlighted the importance of designing *human-machine hybrid data curation systems*. In AFST, an ongoing shift in the data feed as data becomes obsolete, or new data becomes available, was managed through the integrated data warehouse pointing to the need for *evolution-aware data curation methods*. The ability to keep track of the data transformations reinforced the need for mixed-method approaches to tackle challenges associated with verification and repeatability of data curation processes. Similarly, we note the streamlined data flow from institutional systems as key success criteria for RiPPLE. The Illness Severity Prediction case study further identified the need to design *data curation methods that effectively apply to anonymised data*.

Human Oversight Machine-based algorithms cannot effectively complete all data curation activities by themselves [129] and without human intervention [104]. The involvement of domain experts in the child welfare study as a clear example of this fact. The tight communication with domain experts provided a trusted verification step that allowed the tool to continue to deliver decision-making support for the department.

In fact, there are a number of potential issues arising from automatic processes, such as the requirement for substantial manual labelled training data to function effectively [123], issues of trust caused by insufficient understanding of the underlying black-box models (as highlighted by the Illness Severity Prediction case study), the lack of ground truth to evaluate the effectiveness of automated methods [129] (e.g. assessing the quality of learnersourced educational content), and difficulties in automating the detection of errors and cleaning tasks [32]. Given the diversity of data worker roles, a key challenge is the ability to *understand the role of human oversight in data curation processes*.

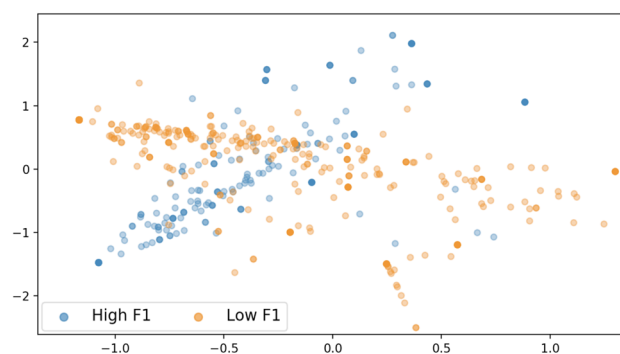


Fig. 6 2-D projection of multi-dimensional embeddings representation of data curation interaction behaviour representing different types of data curators [73]. Blue dots indicate high-quality data worker behaviour vectors and orange dots represent low-quality data worker behaviour vectors (colour figure online)

Thus, the need for human oversight may be the way to address these open issues, increase trust in the system outcomes, and to improve the overall quality. While the benefit of having human oversight is clear, there is a need to understand *how to best involve humans in combination with automated methods* so that such human involvement can be efficiently and effectively deployed.

Data Worker Behaviour Recent studies [104,114] have looked at data workers practices because understanding how data workers engage in manual data curation activities will improve and benefit the design of data curation systems and processes. Further studies have examined the behaviour of data workers in data curation tasks through fine-grained observations made based on a dataset of 50M data points collected from in-lab experiments on a variety of data curation tasks [27,72,73]. Key observations include the ability to distinguish expert and non-expert behaviours to understand which strategies expert data curators adopt. These allow them to perform data curation tasks more efficiently and effectively (e.g. efficiently searching for relevant code snippets online and reuse/adapt them for their data curation task at hand [73]). An open challenge in this regard is *how different data workers perform various curation tasks*.

Moreover, based on low-level behavioural data collected from data curation activities, it is possible to model data curators and to then make performance predictions and curator type classifications based on the observed behaviours (e.g. by building a behaviour embedding representation of data curation activities and comparing such representations in multi-dimensional spaces [72] (see Fig. 6). In the Adaptive Education Systems case study, the use of OLMs highlighted the value of understanding student behaviour. These studies assist in developing a better understanding of data worker behaviour and *how behaviour data can be leveraged to improve data curation*, through tool design, sharing of best practices, and capacity building in the workforce.

In summary, the four functions of data curation at scale highlight the need for organisations to develop capacity for these functions to achieve improvements in data preparation tasks and enable cost-effective data worker time usage and reducing the time-to-value for the organisation.

3.3 Algorithmic transparency

With the phenomenal increase in the volume of non-structured data in the Big Data era, Machine Learning (ML) algorithms have emerged as methods of choice in many applications despite their traditional ‘black-box’ nature aiming to achieve high generalisation performance based on a set of training data. We see deep learning being applied to an ever-expanding range of applications, including computer vision [146], speech recognition [44], recommender systems [158], game programs [136], healthcare [134], and so on.

Assurances and Explainability Models that are more transparent, more explainable, and produce results that are better aligned with human intuition, such as decision trees, are seen as worthy of greater trust, even though they often underperform by comparison. Explainability is the key as it serves the dual purpose of explaining how the impressive prediction accuracy was achieved, as well as providing a path to justify the prediction results, which instils trust in the user. These assurances are particularly critical when the decision can have critical consequences. For example, the consequences of an incorrect cancer diagnosis, whether positive or negative, are severe. In this context, the explanations behind the predictions can be just as essential as the predictions themselves. Interpretable machine learning can be divided into two groups: intrinsic models and post hoc models. The intrinsic models are born interpretable because of their simple yet transparent structures, while the post hoc models are to create a secondary model to interpret the original black-box model.

Data scientists, statisticians, and domain experts have been using intrinsically interpretable learning models for decades on various structured data. For instance, linear regression and logistic regression, which are mathematically explainable and trustworthy, have been widely used in many applications in bioinformatics [34,168]. The simple structure that sums up weights of all the input features in the mathematical equations contributes to their intuitive interpretability because values of the weight directly reflect the importance of the associated feature. In this glass-box model, you could clearly see the contributions from different features to the final output. Another advantage is incorporating specifically designed regularisers in the objective function, such as LASSO regulariser, which can enforce the weights to be sparse. Linear regression can fail when the data points are not linearly separable. Instead, decision trees, which test features with certain thresholds from the root nodes to the leaf nodes concluding the target value, are often used for nonlinear modelling while

preserving interpretability. The representation and interpretation of decision trees are natural for a human to understand and accept.

Another prevailing solution to interpretability is model agnostic [125], which involves extracting post hoc explanations by learning an interpretable model on the predictions of the black-box model. This family of models takes advantages of flexibilities in modelling, explanation, and representation, which make model-agnostic models work with any complex black-box models using diverse explanation and representation forms. For example, one popular model-agnostic method is Permutation Feature Importance (PFI) [20,57], which evaluates how much performance can deteriorate after a specific feature is permuted. The assumption is that an important feature can significantly drag down the performance if its value is randomly picked. Except for directly investigating feature importance, there is another stream of model-agnostic models, which is to find a model to approximate the original black-box model, questioning *how to mimic black-box model behaviours via intrinsically explainable models*. For example, in the Illness Severity Prediction case study, we emphasised the importance of interpretability of machine learning in a real-world healthcare application, where attention weights were visualised to interpret how the multiple deep models work on different organ systems.

The alternative model could either globally or locally explain the inputs and outputs from the target model. In a global manner, the goal is to train an intrinsically interpretable model (e.g. decision tree) to simulate the behaviours of a black-box model over all the data samples. We call this approximated model as a global surrogate model. When the data is not linearly separable, we can build a surrogate model based on a subset of the original data, not over the entire dataset. In this way, we change the focus of the interpretation and it is no longer a global one. Local interpretable model-agnostic explanations (LIME) [126] is one of the representative methods in this family. Another widely used local surrogate model, SHAP [95], calculates Shapley values [133] for each feature to offer an explanation for each instance by measuring its contribution to the overall prediction result. Comparing to global interpretable methods, the local interpretable approaches only concentrate the explanations of a model for a small group of samples, which could result in explanations inconsistent with the global ones. Thus, *how to provide consistent explanations at all levels* is difficult. Apart from the above challenges when developing explainable models, it is worth mentioning that *how to effectively integrate user feedback* is essential to some domains. We observed in the Child Welfare Services case study, that despite the use of intrinsic (e.g. LASSO) and model-agnostic models (e.g. LIME and SHAP) to yield explainable results, these were still insufficient to offer co-dependencies between features and dynamic case-level explanations. It was the

continuous engagement with stakeholders that resulted in achieving the necessary assurances for AFST to be successfully adopted. A similar experience was observed in the Adaptive Education Systems case study, where explainable recommendations improved the transparency of the results and engaged both students and instructors into mutually respectful and trusting partners working together to improve the learning resources together with the consensus algorithms.

Fairness and Transparency ML/DL algorithms are often applied in domains that require transparency, accountability, and fairness, because the predictions made by the algorithms could result in irreversible consequences. For example, when making decisions by policy makers in the fields of social science, education, healthcare, to name a few. In our Adaptive Education Systems case study, the explainable predictions produced by the transparent algorithms provided sufficient fairness guarantee and resulted in deep engagement of students as partners.

When it comes to deep models, there are many layers in the deep neural networks (DNN), which make these models mysterious, even though the performance is often impressive especially in non-structured data analysis, like image, video, multi-variate time series, as shown in the Illness Severity Prediction in ICUs case study. The demand for reducing the opaqueness in deep black-box models has been rising in recent years because the transparency of such high-performance models leads to increased trustworthiness and sheds light on addressing *how to guarantee there is neither implicit or explicit discrimination in the predictions*. In general, there are three types of specific methods to explain the deep models, including feature visualisation [177], model distillation to simulate the DNN behaviour via a smaller model, e.g. decision trees [174], graphs [173], and intrinsic methods such as attention mechanisms [16] and joint training frameworks [94].

It is encouraging to see that there is a growing awareness of the problems surrounding fairness and accountability in machine learning [7]. In general, the function of so-called Algorithmic Transparency is necessary to promote explainability and transparency into the design of machine learning algorithms, which are black-box and often mistrusted by domain experts, yet have the potential to be of high performance and value. It is worth noting that *how to trade off between model performance and interpretability* for black-box models is a promising research direction. Although algorithmic transparency has attracted much research attention, sensitive information protection is still a hurdle because making AI systems fully transparent can expose sensitive data, which leaves *how to protect information in the data* as a follow-up challenge. Table 4 summarises the current literature and open challenges in achieving algorithmic transparency.

3.4 Trusted data partnerships

Many lessons are to be learnt from the legacy of database research in distributed, federated, and multi-databases [110]. Similarly record linkage has been widely studied from statistical [166] and computational [47] perspectives (see also review of Data Linkage in Sect. 3.2). The so-called data deluge of the last decade has challenged well-established norms of data ownership and control, blurring the lines between data creation and acquisition. This has led to both technological as well as organisational challenges in establishing data partnerships. Data portals [107], markets [38], and collaboratives [144] have emerged that present unprecedented opportunities to harness the available data for scientific, commercial, or public good, while also raising a number of access, sharing, and appropriate use concerns. While principles for government and enterprise data access and sharing are still emerging (see, e.g. [1], the scientific community proposed the FAIR guiding principles for scientific data management and stewardship [165] outlining five principles that require data to be Findable, Accessible, Interoperable and Reusable. The FAIR framework has been widely acknowledged and adopted as a useful framework for sharing data that will enable maximum use and reuse of scientific data.

Privacy and Security We note that the presented case studies on child welfare and illness severity had the benefit of pre-existing data collections that provided the necessary input data for the developed prediction models. Challenges in aggregating multi-modal data to build student's knowledge state were observed in the case study on educational systems. These challenges are expected to increase multi-fold when considering cross-organisational data exchange or sharing. The foremost challenge to overcome in data partnerships, is that of satisfying privacy concerns. Privacy is usually referred to as the ability of an individual to control the terms under which personal information is acquired and used. It entails the protection of several data aspects such as collection, mining, querying, releasing, and sharing. Privacy protection has been extensively studied in the literature, and several well-known techniques have been proposed to anonymise tabular records stored in the database including k-anonymity [130], l-diversity [96], t-closeness [92], and differential privacy [50]. With the increasing popularity of intelligent systems applied in various domains (e.g. education, child safety, and intensive care as discussed in the case studies), extensive data analysis has become ubiquitous which, however, increases the risk of disclosing sensitive information of an individual, making it essential to design techniques for privacy protection in these heterogeneous types of data (database records, texts, images, user trajectories). The research questions that need to be considered in this direction include but are not limited to: 1) *How to adapt existing privacy models designed for tabular records to other unstructured data sources*, while

Table 4 Functions relating to algorithmic transparency

Functions	Relevant cases	Selected refs	Open challenges
Assurances and Explainability	Child Welfare Services; Adaptive Education Systems; Illness Severity Prediction	[34,95,125,126,168]	How to mimic black-box model's behaviours via intrinsically explainable models? How to provide consistent explanations at all levels? How to effectively integrate user feedback?
Fairness and Transparency	Adaptive Education Systems	[94,173,174]	How to guarantee there is neither implicit or explicit discrimination in the predictions? How to trade off between model's performance and interpretability? How to protect sensitive information in the data?

Table 5 Functions relating to trusted data partnership

Functions	Relevant cases	Selected refs	Open challenges
Privacy and Security	Illness Severity Prediction	[50,92,96,130]	How to deal with various data types? How to achieve privacy-preserving data linkage? How to balance privacy and utility?
Data Discovery	Child Welfare Services; Adaptive Education Systems; Illness Severity Prediction	[55,56]	How to support domain-specific data? How to improve search efficiency and scalability? How to incrementally add new data source?
Data Linkage and Provenance	Adaptive Education Systems; Illness Severity Prediction	[12,75,127,135,137,140,162]	How to deal with dynamic and heterogeneous data? How to efficiently index and update provenance data? How to support model interpretability with provenance data?
Data Governance	Child Welfare Services	[1,9,37]	How to govern data in use by multiple parties and stakeholders? How to translate governance standards to practice?

considering their specific data characteristics? For instance, some recent studies [74] have shown that the high sensitivity, uniqueness, and low anonymisability of human mobility data raise many new issues and concerns in privacy protection; 2) *How to achieve privacy-preserving data linkage*, noting that data linkage presents enormous opportunities for businesses to explore and leverage the value of linked and integrated data. However, privacy concerns impede sharing or exchanging data for linkage across different organisations. This calls for novel techniques for high-quality entity linking without revealing much sensitive information about these entities. 3) *How to balance privacy and utility*, an enduring issue of performance measurement, where a good privacy protection model should be able to balance two metrics: privacy (how much private information is leaked) and utility (how much information is retained/lost). On one hand, returning completely random data guarantees privacy but results in null utility. On the other hand, retaining raw data maximises utility but ensures no additional privacy. The privacy and utility metrics are usually application-specific, that is, different applications may require different semantics and levels of privacy and utility.

Data Discovery When dealing with large data ecosystems, it is reasonable to expect thousands of sources. Recent research has identified data discovery [55,56] as an important function within organisations that provides the ability to find relevant data to satisfy a particular analytical problem. All of our case studies showed the benefits of ease of access—from the county's integrated data warehouse, institutional educational datasets, and the multi-source sequential database for ICU patients.

Without the ease of access, the challenges of efficient data discovery would have severely limited the success in these case studies. From literature, we note that these challenges relate to 1) *Support for domain-specific data*, e.g. time-series data in the Illness Severity Prediction case may be better represented by their patterns rather than the original sequences, while highly unstructured text can be encoded using their vector representations. Hence, appropriate feature engineering is a major challenge to support accurate discovery of domain-specific data. 2) *Search efficiency and scalability* of data discovery is essentially a nearest neighbour search in the huge data ecosystem, which requires efficient indexing and algorithm design. 3) *As new data sources appear, we need to incrementally add the new sources into the data ecosystem* so that the discovery engine can successfully locate them given a task at hand.

Data Linkage and Provenance With several data sources, the ability to support data linkage (see Sect. 3.2) and data provenance (also called data lineage) becomes indispensable. Data provenance aims to monitor data evolution and trace errors back to the data source when necessary. It is indispensable to achieving trusted data partnership especially

when the loosely coupled data resources are brought together from different applications and organisations. Data provenance has been extensively studied in the past 30 years, starting from the very early form of tagging or annotation used in relational databases [162]. It is typically applied on structured data models (e.g. relational model) and declarative query languages with clearly defined semantics of individual operators (e.g. SQL), and existing provenance solutions usually focus on describing the origins (i.e. instance-based) and/or processing (i.e. query-based) of the data [75,137]. Several recent attempts have also extended data provenance to the blockchain systems [127], Internet of Things (IoT) [12,135], and textual data [140]. Despite these successes, we still observe several open challenges in data provenance based on our case studies: 1) *The data we need to process is usually dynamic and heterogeneous by nature* (e.g. bedside sensors, lab tests, and medical notes in the ICU example), and involves information provided by different vendors which need to be interoperated. However, trusted interoperability is still a major challenge in the development of data partnership across organisations. 2) Although more provenance information can help to better assess quality, ensure reproducibility, and reinforce trust of the data products, it is generally impossible in practice to record all relevant provenance information. In particular, tracing the provenance from a massive amount of data may require excessive resource consumption [12], which signifies the *need for designing novel techniques to efficiently index and update provenance metadata*. 3) Interpretability of machine learning models in highly demanded nowadays. It is natural to question why certain decisions are made by a model, especially when decisions are critical [24]. Provenance of the input and output data may help to *provide human friendly explanations for the model*, which is also a promising research direction for the future work.

Data Governance Together with the technological competencies for managing data partnerships, organisations also need a strong data governance function, which is a policy-centric approach to data models, data standards, data security and data life-cycle management, and processes for defining, implementing, and enforcing these policies [9]. The benefits of good data governance were realised in the child welfare case study through the integrated data warehouse demonstrating best practice for *governing data in use by multiple parties*. Supported strongly by practitioner community, there are a number of well-respected Data Governance frameworks, for example the DAMA framework defines Data Governance as ‘The exercise of authority, control and shared decision making (planning, monitoring, and enforcement) over the management of data assets’ [37]. Governance charters for public sector including the availability of government data for public and social good have been manifested in the form of government open data initiatives in many countries (see, e.g.

data.gov, data.gov.au, data.gov.uk), and further progressed into data transparency and accessibility regulations and legislation, e.g. [1]. However, *how to translate these standards to practice* remains an open challenge.

In summary, trusted data partnerships require four competencies within organisations, namely, (1) ability to satisfy privacy concerns, (2) efficient data discovery, (3) data provenance, and (4) a strong data governance function. Table 5 summarises the current literature and open challenges in each function.

3.5 Agility in value creation from data

Despite technical advancements in data science and machine learning, an open question remains related to pervasive use of data in practice. Many organisations are now attempting to develop into data-driven entities by infusing, adapting, and exploiting contemporary data sources and analytics tools within their business processes, products, and everyday work practices with the ultimate goal of enhancing firm value. Adopting a socio-technical perspective, the Information Systems (IS) field has long studied how organisations generate value from IT assets and capabilities [102,155]. While that literature offers helpful insights for understanding data's strategic value, the study of Information Resilience requires researchers to go beyond the safe confines of stable contexts within which well-understood structured data was owned and controlled by IT units and used internally mainly to describe and monitor firm performance. In the Information Resilience context, researchers need to understand the much more complex, unstable, and diverse reality that organisations now face: Data is not anymore withheld or jailed within IT, rather organisational silos are broken down and data is democratised and traded through platforms in order to facilitate and distribute data-enabled learning and discovery [65,70,151]. This has put data at the forefront and epicentre of organisations' value creation and brings about significant implications to how organisations are structured, operate and meet customer and employee demands.

Our investigation of the case studies and emerging literature identifies four inter-related sub-themes and challenges that together help organisations move towards the function of agility in value creation from data:

Organising for Data-Driven Work One central challenge for organisations that intend to pervasively use data for value creation relates to *structuring of the analytics teams* in ways that data and analytics capabilities can be integrated into every part of the organisation and deeply influence the fabric of the organisation [40]. This typically requires a systematic organisation-wide transformation to build proximity and facilitate effective interaction between analytics teams and business groups [139], but is slowed down because of organisational inertia which creates a natural drift for companies

to maintain their functional silos and their old approaches of conducting work. Organising for data-driven work, however, requires a shift away from a traditional function-based approach wherein a central analytics unit is established to serve varying needs of many business units [66], in most cases leading to a strained relationship between analytics and business groups [115,116]. Instead, data-driven work is rooted in a pervasive approach in which companies embrace Everyone's IT mindset [66] to distribute data, data ownership, and weave analytics capabilities into the fabric of the organisation. Such an approach faces the challenge of *breaking down organisational boundaries* to facilitate analytics and business groups to develop collaboration ties, and a common language to work together in local projects, but also in a way to enable *scaling of analytics approaches* and to *nurture data science talent organisation-wide*.

In the Child Welfare Services case, the Allegheny County dedicated data scientists to the project and centralised functions were only used as a supporting mechanism to build the County's integrated data warehouse, which then provided access to data on various relevant attributes. In the Adaptive Education Systems case, the educational centre hired machine learning academics and nurtured a team of learning analytics professionals to develop the data-enabled adaptive learning tool, RiPPLE. Both organisations invested in training to facilitate the adoption and pervasive use of the tools. The need for organisation re-structuring is also observed in the case of Illness Severity Prediction.

Integrating Analysis with Processes and Products Agility in data-driven work can also be challenged with the way in which analytical insights are delivered, incorporated, and integrated within processes and products. Long-standing IT business literature has advocated for a process perspective on value of IT resources and capabilities [102,112], in which data assets and capabilities generate business value when they are embedded within organisational processes, transforming them into higher-order data-driven business processes. Yet, it is not clear *how to re-design traditional processes with data science*. This perspective also needs to be expanded with the more recent movements towards digital innovation [117,169], particularly *how analytics-based features can be incorporated into products, how they can be developed and how organisations can charge for such features* [11]. Moreover, analytics outputs are changing from the traditional static reports or dashboards that existed separate to the processes to a new logic for process/product design and improvement, in which operational and informational experiences are blended into an integrated application or platform that promotes user engagement and empowers internal users or customers with relevant information and evidence for decision making [167]. Incorporation of data and analytics into product/process architecture (e.g. connected cars or smart fridges) renders them incomplete by design where more and

Table 6 Functions relating to agility in value creation from data

Functions	Relevant cases	Selected refs	Open challenges
Organising for Data-driven Work	Child Welfare Services; Adaptive Education Systems; Illness Severity Prediction	[66,115,116,139]	<p>How to structure data science teams to enable pervasive data use across organisations?</p> <p>How to break down barriers to data-driven work such as functional silos?</p> <p>How to optimise scaling and distribution of analytics resources across organisations?</p> <p>How to nurture data science talent and expertise organisation-wide?</p>
Integrating Analytics with Processes and Products	Child Welfare Services; Adaptive Education Systems; Illness Severity Prediction	[11,13,22,117,121,167,169]	<p>How to re-design business processes with data science?</p> <p>How to add data science features to products and services and how to price these features?</p>
Effective Data Use	Child Welfare Services; Adaptive Education Systems	[8,17,26,36,109]	<p>How do users and designers perceive data and algorithmic-driven representations and how that can inform designing for effective user?</p> <p>How to build faithful representations of data to maximise effective use?</p> <p>How effective use of data can be investigated when tasks are delegated to algorithmic agent or distributed between humans and algorithms?</p>
Data Value Capture and Measurement	Child Welfare Services; Adaptive Education Systems; Illness Severity Prediction	[53,143,151]	<p>How to objectively measure data and analytics value across organisational levels?</p> <p>How to enable data network effects within products and platforms?</p> <p>How to develop user-centric value metrics for data?</p>

more real-time data capture triggers ongoing learning and product/process evolution [63]. Our child protection case showed how a government department was able to formulate an externally facing data-driven service, changing ways of working to a uniform call screening approach and a new policy to do mandatory screen-ins based on data-driven risk scores.

In the Adaptive Education Systems case, a notable design related to how the tasks were divided between humans and algorithms. RiPPLE integrated the algorithmic insights and recommendations with the goal of augmenting users' capabilities and then provided built-in oversight mechanisms such that algorithmic decisions could be monitored and contested. The counter part is designing data-driven organisations that fully automate decision making, where algorithmic agents

decide and act independently [22]. There is growing evidence that this can create tensions by dividing tasks between human agents and algorithmic agents. While algorithms can perform highly structured tasks and process massive data sets in real time, humans usually fare better with less structured tasks, especially ones that require creativity and interpretation [13]. Optimally, human–machine configurations should leverage both agents’ strengths in a complementary manner. However, finding the right balance between automation and human involvement is not easy and practically best approaches are still emerging [121].

Effective Data Use Even with ideal organisational structures and processes in place, Information Resilience can fail if systems are used ineffectively. On the positive side, however, because data can contain unexpected insights, effective use of analytics systems can help organisations to reap unexpected gains. Despite the importance of effective use, many firms struggle to achieve it. Historically we know that, ‘...effective use is one of the greatest challenges for Business Intelligence systems ... Despite increasing investments in analytics systems, many organisations are still unable to attain the desired success ... due to underutilisation and ineffective use’ [8]. One theory that explains how organisations can use data and analytics systems more effectively is the Theory of Effective Use (TEU) [26]. TEU suggests that more effective use involves three dimensions: (a) transparent interaction: seamlessly accessing the representations offered by a system, (b) representational fidelity: obtaining more accurate representations from the system, and (c) informed action: taking actions based on faithful representations. The challenge for executives and researchers studying Information Resilience is to learn how each dimension can be attained.

Advances in data analytics offer challenges and opportunities in each of these dimensions. Users’ transparent interaction with their data and algorithms can be undermined because users are unable to navigate and understand increasingly complex data structures and opaque algorithms. Representational fidelity can be undermined because of the difficulty of knowing if the system faithfully reflects the real-world domain that the system purports to represent. This problem can occur due to the probabilistic nature of common ML algorithms, which render the worlds being represented ‘possible’ rather than ‘real’ worlds, *make faithful representations a challenge for effective data use*. Furthermore, informed action can be undermined because users can have difficulty understanding the predictions and recommendations they receive from advanced algorithms, reducing their trust in the system and their motivation to act upon the data. In the Adaptive Education Systems case study, effective use of the analytics systems was addressed by creation of a high-performance model, ensuring representation fidelity, as well as explainability of the model. For the latter, instructor recommendations were delivered together with justifications of

why instructors received certain recommendations. This was an important step along the way to creating a transparent interaction with the user. The project conducted experiments to ensure that the justifications improved engagement and acceptance, and ultimately facilitated effective use of the system. Similarly, the need for effective use of data analytics is also noticed in the case of Illness Severity Prediction.

Just as TEU can provide insights for achieving or failing to achieve effective use, studies of information resilience in contemporary organisations also offer opportunities for improving TEU. For instance, recent studies suggest that workers may not always need to use their systems directly but rather may delegate that use to other parts of their system (i.e. one system or subsystem uses the other system or subsystem), and likewise, the system could also delegate tasks to the user (i.e. the system uses the human) [17]. More research is needed to *understand the effective use of systems when tasks are delegated to algorithmic agents or distributed between humans and agents*. It is likely that researchers would need to look beyond representational gaps between the system and real world and also look to representational gaps in humans’ mental representations too [35]. More generally, research is beginning to show the importance of understanding *how users and designers work with representations*, how they value them, and how they try to leverage the value in practice [36,109].

Data Value Capture and Measurement Organisations expect a wide range of benefits from their data and analytic initiatives, including increased revenue growth, efficiency and productivity, and strategic planning [53]. Yet, these expectations are often met with challenges, both technical and organisational [143]. A thus-far enduring challenge is *how to quantify the value of data* [88], monetise solutions, and measure return on investment [103]. The adage of ‘you cannot manage what you cannot measure’ also presents a roadblock to increasing value creation and capture from solutions based on data assets. Recent studies have aimed to create better understanding of value realisation from data [143], and are increasingly shifting from traditional ROI-type value measurement to ‘data network effects’ [65,70,151] wherein data value scales as products learn from many users and engage in constant data-enabled learning and improvement. Further work is needed to develop new approaches for valuing data assets, their use and reuse, particular *how incumbent organisations can enable data network effects* within their products and platforms. This also needs to include *a user-centric approach for measuring data* value and how data value can be equally distributed among different stakeholders and its negative externalities monitored and accounted for [138].

In Child Welfare Services, evaluation of the analytics value was complex, with multiple internal and external stakeholders involved. The project was also high-stake because of

working with data on vulnerable people. In all three of our cases, measuring analytics value quantitatively and in a tangible manner has proven difficult and provides opportunities for further research.

Finally, the investigation of data value needs to account for the costs associated with establishing the data infrastructure organisationally, using data responsibly, building trusted platforms, and developing the information resilience capacity of the organisations. We argue that without investing in Information Resilience, the value of data will be short-lived or even produce harm (e.g. earlier mentioned Robodebt case in Australia) rather than benefiting organisations and communities. Organisations investing in developing Information Resilience functions will be able to manage the tensions in data use and achieve long-term sustained advantage.

Overall, to achieve agility in value creation, organisations need to pay attention to four aspects: (1) organising for data-driven work, (2) data-driven process design and improvement, (3) effective use of analytics systems, and (4) measuring value of analytics. Table 6 summarises the open challenges of this function.

4 Manifesto for Information Resilience

The work presented in this paper is the outcome of deliberations of a multi-disciplinary team consisting of social scientists, database researchers, business experts, computer scientists, mathematicians, and information systems researchers. In this section, we assemble our investigations in the form of a manifesto for Information Resilience, which outlines 17 principles based on our identified functions. The brevity of the manifesto is not intended to undermine the depth of challenges in each principle listed, rather to ensure that the breadth is not lost in detail.

Responsible Use of Data Assets

1. Obtain social licence for your analytics project
2. Pursue purposeful analytics
3. Educate your workforce and set high standards for data literacy

Data Curation at Scale

4. Explore your data and understand the impact of data curation on data quality
5. Ensure that there is human oversight of all data curation processes
6. Design and develop repeatability and verifiable data curation processes
7. Understand the behaviour of your data workers and share best practice

Algorithmic Transparency

8. Create fairness and transparency in the working of your algorithms and your people
9. Provide assurances and explainability of your black-box models

Trusted Data Partnerships

10. Satisfy privacy and security concerns of your partners and stakeholders
11. Implement efficient data discovery
12. Keep track of data provenance
13. Strengthen your data governance

Agility in Value Creation from Data

14. Structure your analytics teams and domain experts for effective collaborations
15. Transform your organisation into a data-driven entity that pervasively uses analytics across all business processes and services
16. Make effective use of your analytics systems
17. Define and measure your analytics value

It is not our intention to claim that the manifesto covers all possible challenges and functions for responsible and agile use of information. However, we hope that the scaffolding afforded by our notion of Information Resilience and the 17 functions therein will serve as a reference for future research and inspire often disconnected research communities to come together to collectively tackle these challenges.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Australian government: Data sharing principles
2. Australian competition and consumer commission: Digital platforms inquiry—final report (2019)

3. ABC/Reuters: TayTweets: Microsoft AI bot manipulated into being extreme racist upon release (2016)
4. Abdi, S., Khosravi, H., Sadiq, S.: Modelling learners in crowdsourcing educational systems. In: *Artificial Intelligence in Education*, pp. 3–9 (2020)
5. Abdi, S., Khosravi, H., Sadiq, S., Gasevic, D.: A multivariate ELO-based learner model for adaptive educational systems. In: *EDM*, vol. 12, pp. 228–233 (2019)
6. Abdi, S., Khosravi, H., Sadiq, S., Gasevic, D.: Complementing educational recommender systems with open learner models. In: *LAK*, pp. 360–365 (2020)
7. Ahmad, M.A., Teredesai, A., Eckert, C.: Fairness, accountability, transparency in AI at scale: Lessons from national programs. In: *FACCT*, p. 690 (2020)
8. Ain, N., Vaia, G., DeLone, W.H., Waheed, M.: Two decades of research on business intelligence system adoption, utilization and success—a systematic literature review. *Decis. Support Syst.* **125**, 113 (2019)
9. Al-Ruithe, M., Benkhelifa, E., Hameed, K.: Key dimensions for cloud data governance. In: *FiCloud*, pp. 379–386 (2016)
10. Alevin, V., McLaughlin, E.A., Glenn, R.A., Koedinger, K.R.: Instruction Based on Adaptive Learning Technologies. *Handbook of Research on Learning and Instruction*, pp. 522–560 (2016)
11. Alfaro, E., Bressan, M., Girardin, F., Murillo, J., Someh, I., Wixom, B.H.: BBVA's data monetization journey. *MIS Q.* **18**(2), 66 (2019)
12. Alkhalil, A., Ramadan, R.A.: IoT data provenance implementation challenges. *Procedia Comput. Sci.* **109**, 1134–1139 (2017)
13. Asatiani, A., Penttinen, E., Rinta-Kahila, T., Salovaara, A.: Implementation of automation as distributed cognition in knowledge work organizations: six recommendations for managers. In: *ICIS*, pp. 1–16 (2019)
14. Associates, H.Z.: Allegheny County Predictive Risk Modeling Tool Implementation: Process Evaluation (2018)
15. Azuan, N.A., Embury, S.M., Paton, N.W.: Observing the data scientist: using manual corrections as implicit feedback. In: *2nd Workshop on Human-in-the-Loop Data Analytics*, pp. 1–6 (2017)
16. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate. *ArXiv Preprint* (2014)
17. Baird, A., Maruping, L.M.: The next generation of research on its use: a theoretical framework of delegation to and from agentic artifacts. *MIS Q.* **45**(1), 66 (2021)
18. Batini, C., Cappiello, C., Francalanci, C., Maurino, A.: Methodologies for data quality assessment and improvement. *ACM Comput. Surv.* **41**(3), 1–52 (2009)
19. Belkin, R., Patil, D.: Everything we wish we'd known about building data products (2018)
20. Breiman, L.: Random forests. *Mach. Learn.* **45**(1), 5–32 (2001)
21. Brown, A., Chouldechova, A., Putnam-Hornstein, E., Tobin, A., Vaithianathan, R.: Toward algorithmic accountability in public services: a qualitative study of affected community perspectives on algorithmic decision-making in child welfare services. In: *CHI*, pp. 1–12 (2019)
22. Brynjolfsson, E., Mitchell, T.: What can machine learning do? Workforce implications. *Science* **358**(6370), 1530–1534 (2017)
23. Bull, S.: There are open learner models about! *TLT* **13**, 425–448 (2020)
24. Buneman, P., Tan, W.C.: Data provenance: What next? *ACM SIGMOD Rec.* **47**(3), 5–16 (2019)
25. Bunt, A., Lount, M., Lauzon, C.: Are explanations always important? A study of deployed, low-cost intelligent interactive systems. In: *IUI*, pp. 169–178 (2012)
26. Burton-Jones, A., Grange, C.: From use to effective use: A representation theory perspective. *Inf. Syst. Res.* **24**(3), 632–658 (2013)
27. Chen, T., Han, L., Demartini, G., Indulska, M., Sadiq, S.: Building data curation processes with crowd intelligence. In: *CAiSE*, pp. 29–42 (2020)
28. Chen, W., Wang, S., Long, G., Yao, L., Sheng, Q.Z., Li, X.: Dynamic illness severity prediction via multi-task RNNs for intensive care unit. In: *ICDM*, pp. 917–922 (2018)
29. Chen, W., Yin, H., Wang, W., Zhao, L., Hua, W., Zhou, X.: Exploiting spatio-temporal user behaviors for user linkage. In: *CIKM*, pp. 517–526 (2017)
30. Chouldechova, A., Benavides-Prado, D., Fialko, O., Vaithianathan, R.: A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. In: *FACCT*, pp. 134–148 (2018)
31. Christophides, V., Efthymiou, V., Palpanas, T., Papadakis, G., Stefanidis, K.: An overview of end-to-end entity resolution for big data. *ACM Comput. Surv.* **53**(6), 1–42 (2020)
32. Chung, Y., Krishnan, S., Kraska, T.: A data quality metric (DQM): how to estimate the number of undetected errors in data sets. *VLDB* **10**(10), 1094–1105 (2017)
33. Clarke, R.: Big data, big risks. *Inf. Syst. J.* **26**(1), 77–90 (2016)
34. Couronné, R., Probst, P., Boulesteix, A.L.: Random forest versus logistic regression: a large-scale benchmark experiment. *BMC Bioinform.* **19**(1), 1–14 (2018)
35. Cronin, M.A., Weingart, L.R.: Representational gaps, information processing, and conflict in functionally diverse teams. *Acad. Manag. Rev.* **32**(3), 761–773 (2007)
36. Cunha, J., Carugati, A.: Transfiguration work and the system of transfiguration: how employees represent and misrepresent their work. *MIS Q.* **42**(3), 66 (2018)
37. DAMA. <https://www.dama.org/cpages/home>
38. Data markets: The emerging data economy (2012). <https://techcrunch.com/2012/09/30/data-markets-the-emerging-data-economy/>
39. Darvishi, A., Khosravi, H., Sadiq, S.: Utilising learnersourcing to inform design loop adaptivity. In: *Addressing Global Challenges and Quality Education*, pp. 332–346 (2020)
40. Davenport, T.H., Bean, R.: Big companies are embracing analytics, but most still don't have a data-driven culture. *Harv. Bus. Rev.* **6**, 1–4 (2018)
41. Davenport, T.H., et al.: Competing on analytics. *Harv. Bus. Rev.* **84**(1), 98 (2006)
42. Demartini, G., Difallah, D.E., Cudré-Mauroux, P.: ZenCrowd: leveraging probabilistic reasoning and crowdsourcing techniques for large-scale entity linking. In: *WWW*, pp. 469–478 (2012)
43. Demartini, G., Difallah, D.E., Gadiraju, U., Catasta, M.: An introduction to hybrid human-machine information systems. *Found. Trends Web Sci.* **7**(1), 1–87 (2017)
44. Deng, L., Hinton, G., Kingsbury, B.: New types of deep neural network learning for speech recognition and related applications: an overview. In: *ICASSP*, pp. 8599–8603 (2013)
45. Desmarais, M.C., d Baker, R.S.: A review of recent advances in learner and skill modeling in intelligent learning environments. *User Model. User Adapt. Interact.* **22**(1), 9–38 (2012)
46. Dong, X.L., Gabrilovich, E., Murphy, K., Dang, V., Horn, W., Lugaresi, C., Sun, S., Zhang, W.: Knowledge-based trust: estimating the trustworthiness of web sources. *VLDB* **8**(9), 66 (2015)
47. Dong, X.L., Rekatsinas, T.: Data integration and machine learning: a natural synergy. *VLDB* **11**(12), 66 (2018)
48. Dumitrache, A., Aroyo, L., Welty, C.: Capturing ambiguity in crowdsourcing frame disambiguation. In: *AAAI*, vol. 6 (2018)
49. Dutton, T.: An overview of national AI strategies (2018)
50. Dwork, C.: Differential privacy: a survey of results. In: *TAMC*, pp. 1–19 (2008)
51. Economics, Deloitte Access: ACS Australia's digital pulse 2019 booming today, but how can we sustain digital workforce growth (2019)

52. Eubanks, V.: Automating inequality: How high-tech tools profile, police, and punish the poor (2018)
53. Forrester Consulting: Unlock the power of data to transform your business (2018)
54. Farrell, P.: Government halting key part of robodebt scheme, will freeze debts for some welfare recipients (2019)
55. Fernandez, R.C., Abedjan, Z., Koko, F., Yuan, G., Madden, S., Stonebraker, M.: Aurum: a data discovery system. In: ICDE, pp. 1001–1012 (2018)
56. Fernandez, R.C., Abedjan, Z., Madden, S., Stonebraker, M.: Towards large-scale data discovery: position paper. In: ExploreDB, pp. 3–5 (2016)
57. Fisher, A., Rudin, C., Dominici, F.: Model class reliance: Variable importance measures for any machine learning model class, from the rashomon perspective. ArXiv Preprint 68 (2018)
58. Floyd, S.W., Cornelissen, J.P., Wright, M., Delios, A.: Processes and practices of strategizing and organizing: review, development, and the role of bridging and umbrella constructs. *J. Manag. Stud.* **48**(5), 933–952 (2011)
59. Franklin, M.J., Kossmann, D., Kraska, T., Ramesh, S., Xin, R.: CrowdDB: answering queries with crowdsourcing. In: SIGMOD, pp. 61–72 (2011)
60. Freedberg, S.J., Jr.: Big bad data: Achilles' heel of artificial intelligence (2018)
61. Freitas, A., Curry, E.: Big data curation. In: *New Horizons for a Data-Driven Economy*, pp. 87–118 (2016)
62. Friedman, B., Nissenbaum, H.: Bias in computer systems. *TOIS* **14**(3), 330–347 (1996)
63. Garud, R., Jain, S., Tuertscher, P.: Incomplete by design and designing for incompleteness. *Organ. Stud.* **29**(3), 351–371 (2008)
64. Goldhaber-Fiebert, J., Prince, L.: Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County's Child Welfare Office. Allegheny County, Pittsburgh (2019)
65. Gregory, R.W., Henfridsson, O., Kaganer, E., Kyriakou, H.: The role of artificial intelligence and data network effects for creating user value. *Acad. Manag. Rev.* **6**, 66 (2020)
66. Gregory, R.W., Kaganer, E., Henfridsson, O., Ruch, T.J.: IT consumerization and the transformation of IT governance. *MIS Q.* **42**(4), 1225–1253 (2018)
67. Gross, T., Inkpen, K., Lim, B.Y., Veale, M.: The human(s) in the loop-bringing AI and HCI together. In: *HCI*, pp. 731–734 (2019)
68. Grover, A., Leskovec, J.: node2vec: scalable feature learning for networks. In: *SIGKDD*, pp. 855–864 (2016)
69. Guan, Y., Asudeh, A., Mayuram, P., Jagadish, H., Stoyanovich, J., Miklau, G., Das, G.: Mithrarranking: a system for responsible ranking design. In: *SIGMOD*, pp. 1913–1916 (2019)
70. Hagi, A., Wright, J.: Data-enabled learning, network effects and competitive advantage (2020)
71. Hajian, S., Bonchi, F., Castillo, C.: Algorithmic bias: from discrimination discovery to fairness-aware data mining. In: *SIGKDD*, pp. 2125–2126 (2016)
72. Han, L., Checco, A., Difallah, D., Demartini, G., Sadiq, S.: Modelling user behavior dynamics with embeddings. In: *CIKM*, pp. 445–454 (2020)
73. Han, L., Chen, T., Demartini, G., Indulska, M., Sadiq, S.: On understanding data worker interaction behaviors. In: *SIGIR*, pp. 269–278 (2020)
74. He, X., Cormode, G., Machanavajjhala, A., Procopiu, C.M., Srivastava, D.: DPT: differentially private trajectory synthesis using hierarchical reference systems. *VLDB* **8**(11), 1154–1165 (2015)
75. Herschel, M., Diestelkämper, R., Lahmar, H.B.: A survey on provenance: What for? What form? What from? *VLDBJ* **26**(6), 881–906 (2017)
76. Hey, T., Trefethen, A.: The data deluge: an e-science perspective. In *Grid Computing: Making the Global Infrastructure a Reality*, pp. 809–824 (2003)
77. Hillis, S., Mercy, J., Amobi, A., Kress, H.: Global prevalence of past-year violence against children: a systematic review and minimum estimates. *Pediatrics* **1**, 66 (2016)
78. Hu, Y.H., Lo, C.L., Shih, S.P.: Developing early warning systems to predict students' online learning performance. *Comput. Hum. Behav.* **36**, 469–478 (2014)
79. Hua, W., Wang, Z., Wang, H., Zheng, K., Zhou, X.: Short text understanding through lexical-semantic analysis. In: *ICDE*, pp. 495–506 (2015)
80. ISO/TS 8000-1:2011 data quality: Overview (2011)
81. Jin, F., Hua, W., Xu, J., Zhou, X.: Moving object linking based on historical trace. In: *ICDE*, pp. 1058–1069 (2019)
82. Jin, F., Hua, W., Zhou, T., Xu, J., Francia, M., Orowska, M., Zhou, X.: Trajectory-based spatiotemporal entity linking. In: *ICDE* (2020)
83. Khosravi, H., Cooper, K., Kitto, K.: RiPLE: recommendation in peer-learning environments based on knowledge gaps and interests. *JEDM* **9**(1), 42–67 (2017)
84. Khosravi, H., Demartini, H., Sadiq, S., Gasevic, D.: Charting the design and analytics agenda of learnersourcing systems. In: *LAK* (2021)
85. Khosravi, H., Kitto, K., Williams, J.J.: RiPPLE: a crowdsourced adaptive platform for recommendation of learning activities. *J. Learn. Anal.* **6**(3), 91–105 (2019)
86. Khosravi, H., Sadiq, S., Gasevic, D.: Development and adoption of an adaptive learning system: reflections and lessons learned. In: *SIGCSE*, pp. 58–64 (2020)
87. Korolov, M.: AI's biggest risk factor: data gone wrong (2018)
88. Laney, D.B.: Infonomics: how to monetize, manage, and measure information as an asset for competitive advantage (2017)
89. Lasecki, W.S., Miller, C.D., Naim, I., Kushalnagar, R., Sadilek, A., Gildea, D., Bigham, J.P.: Scribe: deep integration of human and machine intelligence to caption speech in real time. *ACM Commun.* **60**(9), 93–100 (2017)
90. Lee, Y.W., Strong, D.M., Kahn, B.K., Wang, R.Y.: AIMQ: a methodology for information quality assessment. *Inf. Manag.* **40**(2), 133–146 (2002)
91. Leonard, P.G.: Social licence and digital trust in data-driven applications and ai: a problem statement and possible solutions. In: *SSRN* (2018)
92. Li, N., Li, T., Venkatasubramanian, S.: t-closeness: privacy beyond k-anonymity and l-diversity. In: *ICDE*, pp. 106–115 (2007)
93. Li, X.H., Cao, C.C., Shi, Y., Bai, W., Gao, H., Qiu, L., Wang, C., Gao, Y., Zhang, S., Xue, X., et al.: A survey of data-driven and knowledge-aware explainable AI. In: *TKDE* (2020)
94. Lu, J., Xiong, C., Parikh, D., Socher, R.: Knowing when to look: adaptive attention via a visual sentinel for image captioning. In: *CVPR*, pp. 375–383 (2017)
95. Lundberg, S., Lee, S.I.: A unified approach to interpreting model predictions. ArXiv Preprint (2017)
96. Machanavajjhala, A., Gehrke, J., Kifer, D., Venkatasubramanian, M.: L-diversity: privacy beyond k-anonymity. In: *ICDE*, p. 24 (2006)
97. Marbouti, F., Diefes-Dux, H.A., Madhavan, K.: Models for early prediction of at-risk students in a course using standards-based grading. *Comput. Educ.* **66**, 1–15 (2016)
98. Marcus, A., Parameswaran, A.: Crowdsourced data management: industry and academic perspectives. *Found. Trends Databases* **6**(1–2), 1–161 (2015)
99. Marcus, A., Wu, E., Karger, D.R., Madden, S., Miller, R.C.: *Crowdsourced Databases: Query Processing with People* (2011)

100. Marr, B.: *Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results*. Wiley (2016)
101. Matthews, K.E., Dwyer, A., Hine, L., Turner, J.: Conceptions of students as partners. *Higher Educ.* **76**(6), 957–971 (2018)
102. Melville, N., Kraemer, K., Gurbaxani, V.: Information technology and organizational performance: an integrative model of IT business value. *MIS Q.* **66**, 283–322 (2004)
103. Mikalef, P., Krogstie, J., Pappas, I.O., Pavlou, P.: Exploring the relationship between big data analytics capability and competitive performance: the mediating roles of dynamic and operational capabilities. *Inf. Manag.* **57**(2), 103169 (2020)
104. Muller, M., Lange, I., Wang, D., Piorowski, D., Tsay, J., Liao, Q.V., Dugan, C., Erickson, T.: How data science workers work with data: discovery, capture, curation, design, creation. In: CHI, p. 126 (2019)
105. Narayanan, A.: *How to Recognize AI Snake Oil*, p. 20. Princeton University (2019)
106. Nielsen, J.: *Participation Inequality: Encouraging More Users to Contribute* (2006)
107. Open data initiative: Government open data portals. <https://opendatainitiative.github.io/transparency>
108. O’Neil, C.: *Weapons of Math Destruction: How Big Data Increases Inequality and* (2008)
109. Østerlie, T., Monteiro, E.: Digital sand: The becoming of digital representations. *Inf. Organ.* **30**(1), 100275 (2020)
110. Özsu, M.T., Valduriez, P.: *Principles of Distributed Database Systems*, vol. 2 (1999)
111. O’Loughlin, T., Bukowitz, R.: A new approach toward social licensing of data analytics in the public sector. *Aust. J. Soc. Issues* **6**, 66 (2021)
112. Pang, M.S., Lee, G., DeLone, W.H.: IT resources, organizational capabilities, and value creation in public-sector organizations: a public-value management perspective. *J. Inf. Technol.* **29**(3), 187–205 (2014)
113. Parameswaran, A.G.: Letter from the 2017 IEEE TCDE early career award winner. *IEEE Data Eng. Bull.* **40**(3), 4 (2017)
114. Passi, S., Jackson, S.: Data vision: learning to see through algorithmic abstraction. In: CSCW, pp. 2436–2447 (2017)
115. Peppard, J.: Rethinking the concept of the IS organization. *Inf. Syst. J.* **28**(1), 76–103 (2018)
116. Peppard, J.: *The metamorphosis of the IT unit*. MIT CISR Research Briefing, Vol. XIX, p. 7 (2019)
117. Porter, M.E., Heppelmann, J.E.: How smart, connected products are transforming competition. *Harv. Bus. Rev.* **92**(11), 64–88 (2014)
118. Purushotham, S., Meng, C., Che, Z., Liu, Y.: Benchmark of deep learning models on large healthcare mimic datasets. *ArXiv Preprint* (2017)
119. Qu, J., Hua, W., Ouyang, D., Zhou, X., Li, X.: A fine-grained and noise-aware method for neural relation extraction. In: CIKM, pp. 659–668 (2019)
120. Rahm, E., Do, H.H.: Data cleaning: problems and current approaches. *IEEE Data Eng. Bull.* **23**(4), 3–13 (2000)
121. Rai, A., Constantinides, P., Sarker, S.: Editor’s comments: next-generation digital platforms: toward human-AI hybrids. *MIS Quarterly* **43**(1), iii–x (2019)
122. Rao, A.S., Verweij, G.: *Sizing the Prize: What’s the Real Value of AI for Your Business and How Can You Capitalise*. PwC Publication (2017)
123. Rattenbury, T., Hellerstein, J.M., Heer, J., Kandel, S., Carreras, C.: *Principles of Data Wrangling: Practical Techniques for Data Preparation* (2017)
124. Raufflet, E., Baba, S., Perras, C., Delannon, N.: Social License, pp. 2223–2230 (2013)
125. Ribeiro, M.T., Singh, S., Guestrin, C.: Model-agnostic interpretability of machine learning. *ArXiv Preprint* (2016)
126. Ribeiro, M.T., Singh, S., Guestrin, C.: Why should I trust you? explaining the predictions of any classifier. In: SIGKDD, pp. 1135–1144 (2016)
127. Ruan, P., Chen, G., Dinh, T.T.A., Lin, Q., Ooi, B.C., Zhang, M.: Fine-grained, secure and efficient data provenance on blockchain systems. *VLDB* **12**(9), 975–988 (2019)
128. Sadiq, S.: *Handbook of Data Quality: Research and Practice* (2013)
129. Sadiq, S., Dasu, T., Dong, X.L., Freire, J., Ilyas, I.F., Link, S., Miller, M.J., Naumann, F., Zhou, X., Srivastava, D.: Data quality: the role of empiricism. *ACM SIGMOD Rec.* **46**(4), 35–43 (2018)
130. Samarati, P., Sweeney, L.: Generalizing data to provide anonymity when disclosing information. In: PODS, vol. 98, pp. 275487–275508 (1998)
131. Samek, W., Wiegand, T., Müller, K.R.: Explainable artificial intelligence: understanding, visualizing and interpreting deep learning models. *ITU J.* **66**, 39–48 (2017)
132. Sarawagi, S.: *Information Extraction*. Now Publishers Inc (2008)
133. Shapley, L.S.: A value for n-person games. *Contrib. Theory Games* **2**(28), 307–317 (1953)
134. Shickel, B., Tighe, P.J., Bihorac, A., Rashidi, P.: Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *J-BHI* **22**(5), 1589–1604 (2017)
135. Sigwart, M., Borkowski, M., Peise, M., Schulte, S., Tai, S.: Blockchain-based data provenance for the internet of things. In: IoT, pp. 1–8 (2019)
136. Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al.: Mastering the game of go with deep neural networks and tree search. *Nature* **529**(7587), 484–489 (2016)
137. Simmhan, Y.L., Plale, B., Gannon, D.: A survey of data provenance in e-science. *ACM SIGMOD Rec.* **34**(3), 31–36 (2005)
138. Someh, I., Davern, M., Breidbach, C.F., Shanks, G.: Ethical issues in big data analytics: a stakeholder perspective. *CAIS* **44**(1), 34 (2019)
139. Someh, I., Shanks, G., Davern, M.: Reconceptualizing synergy to explain the value of business analytics systems. *J. Inf. Technol.* **34**(4), 371–391 (2019)
140. Song, C., Shmatikov, V.: Auditing data provenance in text-generation models. In: SIGKDD, pp. 196–206 (2019)
141. Stonebraker, M., Bruckner, D., Ilyas, I.F., Beskales, G., Cherniack, M., Zdonik, S.B., Pagan, A., Xu, S.: Data curation at scale: the data tamer system. In: CIDR, vol. 2013 (2013)
142. Stoyanovich, J., Howe, B., Jagadish, H.: Responsible data management. *VLDB* **13**(12), 3474–3488 (2020)
143. Surbakti, F.P.S., Wang, W., Indulska, M., Sadiq, S.: Factors influencing effective use of big data: a research framework. *Inf. Manag.* **57**(1), 103146 (2020)
144. Susha, I., Janssen, M., Verhulst, S.: Data collaboratives as bazaars? *Transf. Govern. People Process Policy* **11**(1), 157–172 (2017)
145. Sutton, C., Hobson, T., Geddes, J., Caruana, R.: Data diff: interpretable, executable summaries of changes in distributions for data wrangling. In: SIGKDD, pp. 2279–2288 (2018)
146. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: CVPR, pp. 2818–2826 (2016)
147. The Ethics Centre—ethics explainer: Social license to operate (2018). <https://ethics.org.au>
148. The British Standards Institution: *Cybersecurity And Information Resilience* (2019)
149. The Guardian—global prevalence of past-year violence against children: A systematic review and minimum estimates (2019)

150. Tibshirani, R.: Regression shrinkage and selection via the LASSO. *J. R. Stat. Soc.* **58**(1), 267–288 (1996)
151. Tucker, C.: Digital data, platforms and the usual antitrust suspects: network effects, switching costs, essential facility. *Rev. Ind. Organ.* **54**(4), 683–694 (2019)
152. UNESCO: Beijing consensus on artificial intelligence and education (2019)
153. Vaithianathan, R., Putnam-Hornstein, E., Chouldechova, A., Benavides-Prado, D., Berger, R.: Hospital injury encounters of children identified by a predictive risk model for screening child maltreatment referrals: evidence from the allegheny family screening tool. *JAMA Pediatr.* **8**, e202770 (2020)
154. Vaithianathan, R., Putnam-Hornstein, E., Jiang, N., Nand, P., Maloney, T.: Developing predictive models to support child maltreatment hotline screening decisions: allegheny county methodology and implementation (2017)
155. Wade, M., Hulland, J.: The resource-based view and information systems research: review, extension, and suggestions for future research. *MIS Q.* **66**, 107–142 (2004)
156. Walsh, T., Levy, N., Bell, G., Elliott, A., Maclaurin, J., Mareels, I., Wood, F.: The Effective and Ethical Development of Artificial Intelligence: an Opportunity to Improve Our Wellbeing (2019)
157. Wang, D., Yang, Q., Abdul, A., Lim, B.Y.: Designing theory-driven user-centric explainable AI. In: CHI, pp. 1–15 (2019)
158. Wang, H., Wang, N., Yeung, D.Y.: Collaborative deep learning for recommender systems. In: SIGKDD, pp. 1235–1244 (2015)
159. Wang, S., Li, X., Chang*, X., Yao, L., Sheng, Q.Z., Long, G.: Learning multiple diagnosis codes for ICU patients with local disease correlation mining. *TKDD* **11**(3), 1–21 (2017)
160. Wang, W., An, B., Jiang, Y.: Optimal spot-checking for improving evaluation accuracy of peer grading systems. In: AAAI (2018)
161. Wang, X., Dong, X.L., Li, Y., Meliou, A.: MIDAS: Finding the right web sources to fill knowledge gaps. In: ICDE, pp. 578–589 (2019)
162. Wang, Y.R., Madnick, S.E.: A polygen model for heterogeneous database systems: the source tagging perspective. In: VLDB, pp. 519–538 (1990)
163. Weld, D.S., Adar, E., Chilton, L.B., Hoffmann, R., Horvitz, E., Koch, M., Landay, J.A., Lin, C.H., Mausam, M.: Personalized online education—a crowdsourcing challenge. In: AAAI (2012)
164. Whiteford, P.: Robodebt was a Policy Fiasco with a Human Cost We Have Yet to Fully Appreciate (2020)
165. Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., et al.: The FAIR guiding principles for scientific data management and stewardship. *Sci. Data* **3**(1), 1–9 (2016)
166. Winkler, W.E.: The state of record linkage and current research problems. In: Statistical Research Division, US Census Bureau (1999)
167. Wixom, B.H., Ross, J.W.: How to monetize your data. *MIT Sloan Manag. Rev.* **58**(3), 66 (2017)
168. Wu, T.T., Chen, Y.F., Hastie, T., Sobel, E., Lange, K.: Genome-wide association analysis by lasso penalized logistic regression. *Bioinformatics* **25**(6), 714–721 (2009)
169. Yoo, Y., Henfridsson, O., Lyytinen, K.: Research commentary—the new organizing logic of digital innovation: an agenda for information systems research. *Inf. Syst. Res.* **21**(4), 724–735 (2010)
170. Zadeh, A., Liang, P.P., Mazumder, N., Poria, S., Cambria, E., Morency, L.P.: Memory fusion network for multi-view sequential learning. In: AAAI (2018)
171. Zanzotto, F.M.: Human-in-the-loop artificial intelligence. *J. Artif. Intell. Res.* **64**, 243–252 (2019)
172. Zapletal, D., Herout, A.: Vehicle re-identification for automatic video traffic surveillance. In: CVPR, pp. 25–31 (2016)
173. Zhang, Q., Cao, R., Shi, F., Wu, Y.N., Zhu, S.C.: Interpreting CNN knowledge via an explanatory graph. In: AAAI, vol. 32 (2018)
174. Zhang, Q., Yang, Y., Ma, H., Wu, Y.N.: Interpreting CNNs via decision trees. In: CVPR, pp. 6261–6270 (2019)
175. Zhang, R., Indulska, M., Sadiq, S.: Discovering data quality problems. *Bus. Inf. Syst. Eng.* **61**(5), 575–593 (2019)
176. Zheng, L., Zhang, H., Sun, S., Chandraker, M., Yang, Y., Tian, Q.: Person re-identification in the wild. In: CVPR, pp. 1367–1376 (2017)
177. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A.: Object detectors emerge in deep scene CNNs. *ArXiv Preprint* (2014)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.