Digital Identity Management in Formal Education

Implications for Policy and Decision-Making

Alan Moran

Digital Identity Management in Formal Education

Digital Identity Management in Formal Education offers a broad analysis of the online self considered from educational policy, technological, legal and social perspectives. This book introduces the reader to the notion that digital identity is a multifaceted topic which requires a broad and systematic approach that is rooted in risk-based policy. It provides educational technologists, leaders and decision-makers with an accessible, jargon-free guide to their responsibilities towards students and instructors in today's digitally networked schools and universities. Real-life examples illustrate how digital identities impact management and delivery, privacy and transactions, governance and accountability, and other interconnected choices in the use of technology-enabled services in formal learning.

Alan Moran is a senior education technology manager specialising in risk management, information security and data protection. He is currently working in the Swiss public and non-profit sectors, where he frequently engages in matters relating to the socially acceptable and legitimate use of technology in society.

Digital Identity Management in Formal Education

Implications for Policy and Decision-Making

Alan Moran

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

First published 2022
by Routledge
605 Third Avenue, New York, NY 10158

and by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

Library of Congress Cataloging-in-Publication Data
A catalog record for this title has been requested

ISBN: 978-0-367-64798-8 (hbk)
ISBN: 978-0-367-67839-5 (pbk)
ISBN: 978-1-003-13307-0 (ebk)

DOI: 10.4324/9781003133070

Typeset in Baskerville
by Taylor & Francis Books

To my lovely wife Helen for this, that, and the other

Contents

Illustrations

Figures

Tables

Preface

This book focuses on identity management within formal education, by which is implied the teaching and learning that takes place within an organisation and which is structured around a syllabus or curriculum, whose primary function is to facilitate the acquisition of knowledge, behaviours, skills, attitudes and practices typically leading to some form of certification. Accordingly, this context provides the organisational setting in which identity management is situated and which in turn lends legitimacy to the online delivery of education.

Identity management is a complex topic that is considered from multiple perspectives, each of which adds something to reveal a more complete picture of what digital identity means in education today and what developments decision makers need to be aware of in the near future. Each perspective is intended to highlight the primary issues faced by decision makers in that domain and to draw attention to their relationship with the other perspectives. This ought to encourage the reader to consider identity in systematic terms, rather than focus purely on the organisational or technical aspects. Accordingly, this book is not an operations manual detailing specific courses of action or providing templates and checklists. Rather, the implications are formulated broadly in terms of risk-based policy, the belief that technology is never as neutral as may first appear and a concern for the legal liabilities faced by educational establishments as well as the need to be cognisant of the social implications of identity. For those wishing to delve deeper into any of these topics there are copious references at the end of each chapter which may offer further information.

The reflections in this book are a product of my last couple of decades working in the fields of (educational) technology, risk management, information security and data protection tackling a broad spectrum of issues at all levels of formal education. Such work was carried out in consultation with a wide range of stakeholders spanning front line professionals responsible for the delivery of education, data protection officers and information security specialists working in the public sector, service providers in the private sector developing innovative digital solutions for education through to politicians tasked with devising a coherent vision for education in a digital society. Secure

in the knowledge that I do not have all the answers, my humble hope is that the following chapters will nonetheless illuminate some of the issues surrounding digital identity and provide insights for those seeking to achieve a more humane and social digital transformation of their organisations.

Alan Moran

1 Introduction

Identity is a fluid concept that is highly context sensitive and any attempt to define it must be situated within a specific frame of reference. Indeed, a glance at the history of the development of identity management infrastructures reveals the assumption that access to computing networks and online services was somehow premised on one's membership of an organisation that was in a position to attest to that fact. As access broadened, market imperatives took hold and architectures emerged that favoured the growth of subscribers, by widening the base of permissible users to anyone who registered an interest in a service. At the same time new data-driven models emerged that saw profit in the exchange of personal data and the monitoring of online behaviours. These developments in turn prompted responses aimed at regaining control over identity, by circumventing the need to appeal to centralised public or private authorities in order to affirm one's presence online. In tracing this path, it becomes apparent that traditional educational identity management remains situated within an organisational context, from which it derives its legitimacy as the deliverer of education. At the same time, however, there is a dein-stitutionalisation and a "democratisation" of identity at play that seeks to establish new norms and power relations. An individual therefore finds herself faced with the task of determining the appropriate balance between asserting her own identity and the need for certain characteristics to be affirmed by others (e.g. schools, universities).

Education finds itself in the vortex of these developments as it attempts to leverage its own organisational legitimacy through identity infrastructures (by means described in Chapter 5) and yet redefine itself along as yet unchartered lines. All this must occur in support of the wider goals of education based on lifelong learning (that extend beyond organisational boundaries) while appro-priately recognising the different roles of identity in non-formal, informal and formal education. In attempting to understand this state of affairs a brief initial survey of how technology is used in education is undertaken. Thereafter, a tech-nical definition of digital identity is presented that finds wide application both within and outside of education and which is found to be appropriate, at least in an initial assessment of the identity needs of education. This definition is then explored further in terms of those infrastructural components, atop of which

DOI: 10.4324/9781003133070-1

classical digital identity lifecycle management is usually realised. This enables a more detailed examination of the identity infrastructures that are commonly found in education some of which may be linked to other public sector infrastructures such as national electronic identity schemes. It will, however, be noted that a standardised technocratic definition of digital identity falls short of the mark by failing to acknowledge the social and psychological dimensions of identity (that will be explored further in Chapter 5). Finally, an exploration of the potential implications of reframing identity in education along user-centric lines (as alluded to above) is deferred in order that a baseline can be established first (though this will be returned to again in Chapters 3 and 4).

Educational Technology

There are considerable expectations placed on the shoulders of educational technology that it will simultaneously reinvigorate learning in the classroom, while also easing the administrative burdens associated with teaching (Selwood, 2005). These two facets of education are linked by the belief that the data generated by digital services can be automatically processed to render information of use to educators. This mindset is spurned on by the notion that data-driven technologies such as learning analytics support the behavioural analysis of student interaction with materials in ways that are "scientific" and "evidence-driven" suggesting the neutrality and inevitability of their conclusions. It is rather telling that this claim that technology will "revolutionise" teaching practices, leading to significant efficiency gains which ultimately "democratise" access to education has a long history in education dating back to similar claims made of film, radio, television, microcomputers (Cuban, 1986) and more recently the Internet. These promises of reinvigorated learning and administrative efficiency gain have resulted in technology being positioned as a "production factor" within education. Over time, this has extended beyond its use in "lab" environments where computers were considered an important but ancillary support function, to become a central instrument in curriculum design and delivery. Nowadays entire teaching, learning and administrative processes can be found that are built entirely around digitalised models. These include online enrolment (including fee payment using Bitcoin), procurement of services through smart contracts, collaborative e-learning in groups, behaviour analysis and intervention (using artificial intelligence), assessment using remote proctoring, through to the award of digital achievement badges. Such developments have not only become commonplace, but are now defining new norms in terms of how education is understood. This is particularly important since it is the emergence of socio-technical infrastructures and the values they represent, that underpins the legitimacy of digital education. These include enablers in the efforts to create more student-centric models of education, wherein the role of the educator is reinterpreted as that of a coach or mentor assisting the learner in her efforts to make sense of the world.

This phenomenon is particular true of Massive Open Online Courses (MOOC) that garnished credibility through the efforts of leading US universities (e.g. MIT), who during the early 2000s made available vast amounts of resources via online platforms. These efforts culminating in the various e-Learning standards that included not only Open Educational Resources (OER)[1] but also content delivery (e.g. SCORM) and assessment (e.g. QTI) formats. Keenly supported by transnational organisations (e.g. UNESCO, OECD) and technology providers, MOOCs have found widespread popularity by prompting more modular forms of education that challenge the "monolithic" approaches offered by traditional learning institutions (e.g. qualifications based on the Bologna and its associated ECTS scheme). Despite the aspirations of (sometimes) free access to online resources, the availability of expert teachers and a relaxation of temporal and geographical constraints on learning evidence, suggests that they have mostly stood to benefit those who were already well educated[2] (Matthias & Mario, 2015). Furthermore, there remain reservations concerning completion rates, the recognition of awards and how data being gathered on learners might be sold on and used in other contexts. Consistent with this research on MOOCs it has been found that a wide range of factors (e.g. gender, race, socioeconomic status, educational level, etc.) are associated with disparities in access to and use of technology (Goode, 2010). Indeed, there appears to be evidence that technology is used differently based on socio-economic status, reflecting pre-existent divides within society that technology was only serving to widen (Warschauer, 2000). This suggests that one's relationship with technology may well be shaped by culturally situated experiences.

Of central importance to the deployment of such technologies is how online activities are bound to individual persons and how the data gathered from such sources can be aggregated with other sources of data to build more complete profiles of online behaviours. In this context digital identity constitutes not only a cornerstone of the digital learning ecosystem, but also an arena in which new power structures are defined. Specifically, the ownership of digital identity management infrastructures confers considerable power in terms of the insights that can be gleaned and the ability to draw inferences from them. This is particularly concerning when applied to education as in the case of learning analytics that are often protected by trade secrets thereby releasing service providers from the scrutiny that would otherwise be afforded to education. For example, despite the popularity of digital learning tools in schools (estimated to be supported in the US by over 90% of learners and nearly as much by staff), studies involving OECD countries (incl. the US) have indicated that those who heavily use computers, perform a lot worse in most learning outcomes (Wexler, 2020). A similar study conducted within the US by the National Education Policy Center at the University of Colorado, concluded that the self-interests of the technology industry ensured a lack of transparency prevailed in relation to learning products and their algorithms (which were protected as trade secrets), exposing students to serious privacy threats. Moreover, the privatisation of educational decision-making has been

cited as a specific criticism that resulted in a distortion of pedagogy in a manner that stifled student learning and inhibited their ability to participate in the democratic system (Boninger et al., 2019).

One's own individual command of digital identity technology may also act as a technical or social barrier to involvement in certain activities (a notion referred to as technological identity in Chapter 5). For example, while determining appropriate password strength may perhaps be considered mundane enough for most users, securing other forms of digital identity such as private keys can be technically quite challenging. This may entail using a key store (e.g. PKCS#12[3]), deciding on an appropriate level of confidence in the identities of others within a web of trust model[4] or applying a digital signature using the appropriate cryptographic standards (e.g. PKCS#7,[5] PKCS#11[6]). While these are entirely learnable skills, they are likely to appear daunting to the uninitiated and therefore constitute a limiting factor to their access to services. This is particularly so in education where the first steps to socialisation in a digital world occur and where early identity habits are being formed. These and other barriers may drive social divisions surrounding education later in life, particularly when it comes to online study. For example, based on estimates by UNICEF nearly half a billion children worldwide lacked basic access to remote schooling (e.g. Internet, television or radio) during the COVID-19 pandemic (UNICEF, 2020), thereby putting them at a (further) disadvantage when compared to their wealthier peers. This is particularly pronounced in low-income economies where despite few households having a home computer, remote learning delivery models of service providers continually make overly generous assumptions concerning access to technology that appear to exclusively benefit more affluent households (Chauvin & Faiola, 2020).

While technology determines what is within the realm of the possible, it falls to society to consider what is morally and ethically legitimate. This requires an open public discourse surrounding the needs and values of society including how best to deliver a socially acceptable digital transformation of education as discussed throughout this book. Only then can legislators govern these social arrangements through appropriate legislative and regulatory measures (as described in Chapter 4). That identity is a political issue is evident in the United Nations' Sustainable Development Goals where "provid[ing] legal identity for all including free birth registrations" (UN, 2015) has loomed large. Though this goal emphasises the political dimension of identity in terms of citizenship and the accountable of a state towards its members, it also sets a clear trajectory for how digital identity must be framed in an increasingly digitalised world, for which a definition must first be formulated.

What is Digital Identity?

Ultimately the purpose of digital identity is to ensure access to services and accountability in respect of their use. When engaging with a digital realm (e.g. an online learning community, a student administration system, an e-portfolio

platform or interacting virtually with robots or other devices, etc.), a physical person requires a digital presence. This binding of an individual to a set of characteristics that uniquely identify her within that digital realm and makes her involvement in transactions therein traceable, is broadly what is understood by digital identity (a more precise definition of which will follow later). Similarly, other actors within the digital realm may also be made identifiable as in the case when "device identity" is used for hardware entities (e.g. robots, sensors, etc.) though this form of identity will not feature prominently in what follows. While the former requires technical mechanisms that will be described in greater detail in Chapter 3, the latter is moderated by the social norms and values that govern acceptable behaviour as described in Chapter 5.

Within any digital realm, a *subject* refers to an entity that actively engages with a (passive) object. A subject generally refers to a human being (acting through her digital manifestation), though the term could equally apply to a physical mechanical component (operating via an interface) or a software process that acts directly on an object of interest (e.g. a document, process or other digitalised entity). On closer inspection it becomes apparent that this subject–object relationship is relative in that a given entity may act both as a subject or as an object depending on the context. For example, a faculty member may edit student records located in a student administration system which in turn places those records in persistent long-term storage (e.g. a database). In the initial stage of this interaction the faculty member (subject) edits records in the system (object), however, it is the system (subject) that later stores these records in records (objects) located in a database (see Figure 1.1).

Clearly there is a need to clarify the operational context of a subject in order to reliably associate that subject with its actions if there is to be any trust in the validity and integrity of the notion of digital identity. It therefore becomes evident that this situation is best understood if one separates these two contexts into distinct domains of control. Each domain will thus be regulated by different

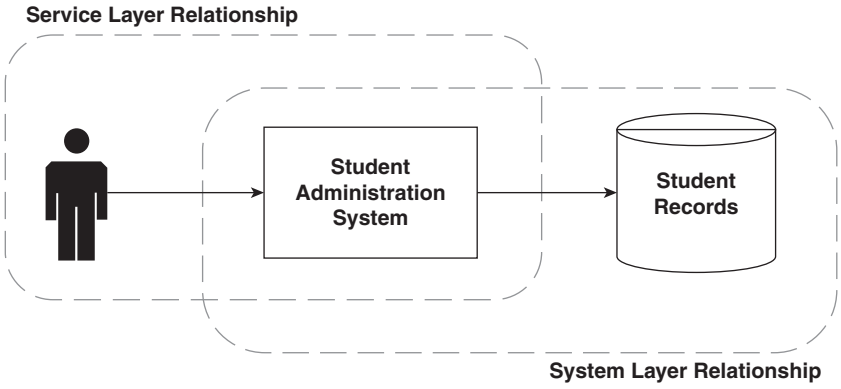

Figure 1.1 Subject–Object Relationships

digital identity systems wherein the relevant user assumes an identity that is valid only for that domain. For example, initially the faculty member will probably be assigned an account with the appropriate rights to edit records (e.g. create, update, delete) in the student administration system. Separately, the administrative system itself will likely be operating as a technical process that uses a technical account whose rights (issued by an operating system) permit it access to specified databases and their directories and files. When discussing the technical details of identities (see Chapter 3), a distinction will therefore be drawn between these two types of identities, where the former will be referred to as a service layer identity (that is high-level, people-focused and session-oriented) and the latter as a system layer identity (that is low-level, machine-focused and connection-oriented) whose domains are indicated by the dashed boxes in Figure 1.1. This suggests that any definition of digital identity that seeks to bind a subject to its actions is constrained by the operational context in which such actions are undertaken, as defined by a domain of control. Hereafter when referring to digital identities the emphasis will be on persons operating in "service layer" relationships who shall simply be referred to as users.[7]

There is no consistent definition for digital identity within information security owing to the very fact that context plays such a big role in understanding what precisely is to be meant by that term. Indeed, even in respect of "user identity" the circumstances of specific operational environments make clear the multifaceted nature of "online personas" by which a user can represent herself (NIST, 2017a) which of course complicates efforts to arrive at a single consistent definition (as explored further in Chapter 5). Generally speaking, however, a digital identity must at least be capable of uniquely identifying an actor within the domain of control in which it is applied and must in addition be capable of ascribing "a set of attributes" to that actor (ISO/IEC, 2019). Additional facets may also be associated with digital identity, not least the ability to assert an identity (e.g. credentials) though these more often than not pertain to how a digital identity is implemented in practice. For the purposes of this book a *digital identity* is defined as the unique "representation of an entity (or group of entities) in the form of one or more information elements which allow the entity(s) to be uniquely recognised within a context to the extent that is necessary (for the relevant applications)" (ITU-T, 2009). In other words, a unique representation of a user in a manner that encapsulates those attributes of that user, that are deemed relevant within a specific operational or transactional context and which enables claims relating to that identity to be verified. In practice the information to be found in a digital identity typically consists of (ITU-T, 2009):

- an *identifier* that uniquely identifies the user within a specific domain[8] (e.g. an email address, social security number or other technical identifier such as a username);

- one or more *credentials* with which the binding of the user to the identity may be verified (e.g. a password, digital certificate or biometric information);
- one or more *attributes* which describe specific characteristics of the user[9] and from which claims relating to the user may be asserted (e.g. role(s), first name, surname, institutional association, etc.).

The inclusion of credentials in the definition of digital identity is a pragmatic reflection of the fact that digital identities only have meaning for their users when they can be used a means asserting themselves. A more nuanced approach might suggest that a digital identity must first be established and validated before it can be issued a credential attesting to its validity (as described later in this chapter). In keeping with the notion that a digital identity represents the key to the economic, social and cultural engagement of an individual in a globalised digital world (World Bank, 2016; EU, 2014a),[10] the above definition also affirms the manner in which a user (e.g. a person) can be uniquely identified within any of these environments. Beyond this technocratic description of digital identity lies the question of what meaning can be ascribed to an identity and how it can be said to represent the totality of online experience of an individual participating in a digital society which will be returned to again in Chapter 5. This notwithstanding the meaning attributed to an identity within a specific domain, determines what attributes or characteristics that identity must possess, irrespective of whether or not these are recognised by other domains. For example, a passport document reflects an (analogue) identity that bears political significance (e.g. by asserting one's citizenship) and accordingly mandates specific attributes[11] (e.g. first name, surname, date of birth). That precisely these attributes may also be used in other scenarios not initially envisaged by that context (e.g. the use of a passport to purchase alcohol) is less important than the assurance that that form of identity provides (e.g. the acceptance of "official documentation"). Indeed, it is the context specificity of identity that precludes the notion that there will ever be a single identity scheme that could service all contexts equally, owing to the fact that each context has its own "vertical" concerns that are not shared by other contexts. This, of course, prompts the question of what is meant by the notion of a *domain*, which will be taken here to refer to the formal sphere of trust in a set of identities issued by a specific authority. This refers to those computing systems that are willing to accept the assertions made on behalf of an identity as embedded in their formal organisational and technical arrangements with the issuer of that identity. Policy-makers and their security managers need to be rather punctilious about defining what constitutes a domain within their organisations, as this underpins the most important aspect of digital identities, namely their role in access control and accountability. Domains are considered in more detail in light of the specific identity infrastructures found in education that are described in Chapter 3.

Once a service has chosen to accept an identity, it may independently continue to gather or profile data concerning that individual, that is of