4 Measuring the digitalisation of higher education in Hungary

This chapter focuses on approaches to measure the digitalisation of higher education in Hungary. It first introduces different data collection approaches and indicators used internationally to measure the digital transformation in higher education, assessing the benefits and drawbacks of different approaches. It then provides an overview of higher education data collection in Hungary and discusses potential future data collection and priority indicators to assess progress in the digitalisation of Hungarian higher education.

4.1. Challenges in measuring the digitalisation of higher education

The coronavirus (COVID-19) pandemic has accelerated the pace of digitalisation in higher education worldwide. However, data that could help obtain a nuanced understanding of how much and what type of digitalisation is taking place in higher education are rarely collected in a consistent manner at an institutional or national level.

Several factors may contribute to the lack of data in this area:

- **Low policy priority until recently**: Digital higher education, while it has existed for decades, has long represented a small share of total higher education enrolments across OECD countries. It has often been a means to reach students not able to attend higher education institutions (HEIs) in person due to geographic, time or personal constraints. In the **United States**, one country that has data on online enrolments, it has often developed in non-selective, often private, HEIs (see for instance (Xu and Xu, 2019[1]). Digital higher education has therefore not been a priority of governments requiring data to be collected and reported by publicly funded HEIs.
- **Difficulty developing definitions of digital higher education**: The institutional autonomy and academic freedom characteristic of higher education systems in most OECD countries – in contrast to public school systems – means that digitalisation takes a wide variety of forms in higher education. Identifying what course, programme or learning module counts as "e-learning" or "digitally enhanced teaching and learning", for instance, varies considerably across and within both countries and HEIs. The concept of digital enhancement is not binary: a course can use digital tools as an add-on to traditional in-person delivery (for instance, making materials and videos of lectures available on line); or as an essential component of delivery (such as making some parts of the delivery and assessment available only on line); or making all aspects of the course available on line (so there is no in-person component at all). These differences are important to take into account in the design of a measurement system (Guiney, 2016[2]; Ifenthaler, 2021[3]), but entail costs to ensure rigorous definition development, data collection, categorisation and reporting.
- **Need for data collection tools that help understand user practices**: Understanding digitalisation in higher education involves not just measuring inputs that characterise digital readiness (e.g. access to hardware and software) but also the uses of digital technologies by students and staff. As discussed later in this chapter, these tools are often surveys that can involve significant development costs, data quality issues and a high compliance burden for respondents.

These factors make it difficult to quantify in a comparable manner at an institutional or system level the *amount* of e-learning or digitally enhanced teaching and learning that takes place. Furthermore, this diversity makes it challenging to measure the *efficiency, quality and equity* of digitally enhanced teaching and learning, which require adapting higher education data collection to this diverse and fast-changing type of teaching and learning provision.

At the same time, digital higher education involves new measurement opportunities. This is because digital teaching and learning practices generate a large amount of detailed data that, coupled with student outcomes data, can generate rich insights into student weaknesses and strengths and support student success. In particular, if instructors design their courses to make central use of a learning management system (LMS) or virtual learning environment (VLE), the system will generate a record of the transactions of each student with the course components (Ifenthaler, 2012[4]). The data generated in the LMS/VLE creates an opportunity for *learning analytics*, which is the use of that data – often in conjunction with other sources of student data – to track a student's engagement with learning.

Despite the challenges to measure the digitalisation of higher education, some governments and HEIs across OECD countries have developed methods to monitor the provision of digitally enhanced teaching

and learning. Three key methods for measuring the digitalisation of higher education include administrative data collection, surveys of higher education students and staff, and the use of learning analytics.

The following sections look at these three methods in turn, discussing the data collection approach and indicators they generate and their benefits and drawbacks. While not discussed in this chapter, it should be noted that other methods, such as interviews and focus groups with users of digital technologies in HEIs, also offer rich qualitative data that are important to the understanding of the level of digitalisation at the institutional and system levels.

A final section provides a summary table of the three methods, discusses the benefits of combining the three methods to obtain a deeper understanding of digitalisation in higher education and discusses common issues, such as data privacy and use.

4.2. National administrative data systems

Data collection approach and indicators

Administrative data on higher education is the data an institution collects to manage its processes (for instance, of enrolment, assessment and completion), students, staff, academic programmes, research, finances and physical assets. Administrative data is housed in the institution's databases and is processed by its systems – such as its student management system, finance system and asset management system.

Most of an institution's data on students, staff and academic processes will be held at unit-record level; each individual student is assigned an identifier, with the databases holding data that enable the identification of and communication with the student, his or her demographic characteristics, academic history, as well as what classes he or she is taking and the results of assessments in those classes. This sort of data is used for critical administrative functions, such as generating class lists, recording grades, producing result notices and academic transcripts, and establishing entitlements to graduate. Likewise, staff data are held at the unit-record level and are used to populate the payroll system, etc.

To manage a higher education system – to run its funding system, inform policy, and monitor the system's performance and quality – governments require institutions to submit extracts or summaries of each institution's administrative data. Governments typically specify the form of the data it collects and the fields on which it requires data, and institutions will be obliged to ensure that the data they collect from students is sufficient to enable them to complete the government's data collection.

As a result, administrative data collected at the institutional level is consolidated to create a national administrative dataset. The collection of HEI administrative data by the government is usually done in one of two possible ways – by uploading an extract from the *unit-record* data or by collections of *aggregated* or summary data (Box 4.1).

Box 4.1. Two approaches to collecting national administrative data: Unit-level and aggregated

Public authorities in OECD countries take different approaches to collecting higher education administrative data at the national level. Government unit-record data on enrolments and completions, courses and programmes are routine in **Australia** (Australian Government, 2021[5]), **New Zealand** (New Zealand Ministry of Education, 2021[6]) and the **United Kingdom** (HESA, 2021[7]), as well as **Hungary** (DSN/DHECC, 2020[8]). In the **United States**, aggregated data is collected by the federal government agency, the National Center for Education Statistics (NCES), in the Integrated Postsecondary Education Data System (IPEDS) collection (NCES, 2021[9]).

A national unit-record data collection requires all institutions to manage their data in ways that fit the government's data model. This can impose high costs on smaller institutions. It can also create concerns about data privacy. The **United States** has a long-standing ban on federal collection of unit-level higher education data motivated primarily by privacy concerns (Miller and Shedd, 2019[10]). However, many US states do collect unit-record data on students enrolled in public higher education institutions (SHEEO, 2021[11]).

However, a unit-record dataset allows analysts the flexibility to look at any combination of variables. It also allows for the use of more advanced statistical modelling techniques to establish the statistical significance of the relationship between variables – for instance, which factors are significantly related to the probability of completing a higher education programme.

Given the challenges of defining digital higher education, collecting administrative data on the digitalisation of higher education is difficult, particularly at the national level. However, some countries do require HEIs to report administrative data that shed light on the provision of digital higher education at a system-wide level. The indicators collected typically help provide a picture of the scale of digital provision, in terms of institutional provision (number of courses, programmes, fields and levels of study) and student participation (enrolment, completion, student demographics).

In the **United States**, the National Center for Education Statistics manages the Integrated Postsecondary Data System (IPEDS), a national database that collects data on a wide array of indicators enabling a detailed understanding of the US higher education system. IPEDS collects data through institutional surveys covering the following topics: institutional characteristics, completions, 12-month enrolment, student financial aid, graduation rates, 200% graduation rates, admissions, outcome measures, fall enrolment, finance, human resources and academic libraries (NCES, 2021[12]). IPEDS also collects data on distance education, defined as:

> *education that uses one or more types of technology to deliver instruction to students who are separated from the instructor and to support regular and substantive interaction between the students and the instructor synchronously or asynchronously. The following types of technology may be used for distance instruction: Internet; satellite or wireless communication; and audio and video conferencing. (NCES, 2021[13])*

Courses and programmes offered by HEIs are only considered distance education if all instructional components can be completed remotely. As a result, degree programmes that offer a blend of in-person and online instruction are not classified as distance education by IPEDS. However, IPEDS does track whether distance education programmes have non-instructional onsite components, e.g. orientation for new students or testing. Table 4.1, adapted from the IPEDS website, outlines the distance education data indicators collected by IPEDS.

Table 4.1. Overview of distance education data collection in IPEDS

Survey component	Collection period	Data coverage period	Distance education (DE) courses	Distance education programmes	Data collected
Institutional characteristics (IC)	Fall	Current academic year	Yes	Yes	Captures whether institutions offer DE courses and/or programmes for undergraduate and graduate students and whether all programmes are offered exclusively via DE
12-month enrolment (E12)	Fall	1 July-30 June (prior year)	Yes	No	Captures the number of students enrolled in DE courses over 12-month period
Fall enrolment (FE)	Spring	Institutions' official fall	Yes	No	Captures the number of students enrolled in DE courses in the fall term and, of the students enrolled exclusively via DE, the number located in various geographic categories

Survey component	Collection period	Data coverage period	Distance education (DE) courses	Distance education programmes	Data collected
		reporting period			
Completions	Fall	1 July-30 June (prior year)	No	Yes	Captures whether all, some or none of the programmes within each CIP code and award level can be completed entirely via DE and whether certain DE programmes have onsite components (non-instructional components such as orientation or testing, do not exclude a course from being classified as DE).

Note: CIP code refers to the Classification of Instructional Programs and is the coding system for different programme types an institution offers.
Source: NCES (2021[13]), *Distance Education in IPEDS*, https://nces.ed.gov/ipeds/use-the-data/distance-education-in-ipeds.

In addition to distance education (DE) data, IPEDS collects data on institutions' digital/electronic library resources, including the number of digital/online books, databases, media and serials. While IPEDS distance education does not track digital practices (use of digital technology by students and staff), library resource data is one measure of digital readiness, as it covers the availability of digital resources. The 12-month enrolment in distance education has only been added in the most recent academic year, however, and therefore longitudinal data series are not yet available. The other three elements are presented from 2012 in individual, institutional data profiles.

The IPEDS collection does not count the number of students enrolled in DE who complete or pass courses or programmes, meaning there is no comparative data on completions in DE and other delivery modes.

In **Australia** and the **United Kingdom**, where institutions supply unit-record data to government agencies (Australian Government, 2021[5]; HESA, 2021[7]), data on delivery mode make it possible to report on pass rates in distance education.

In **New Zealand**, the Single Data Return (SDR) system collects data from HEIs on their programmes and courses and on each student's enrolment and completion of courses and programmes to generate a detailed view of the system. Courses are categorised by the academic department responsible for delivery of the course on whether they have elements of e-learning and, if so, the extent of the e-learning – whether the online components are optional add-ons, or essential and significant, or if the course is delivered wholly on line (New Zealand Ministry of Education, 2021[6]). However, there is some uncertainty as to whether the definitions in the SDR manual of those categories are precise and detailed enough to ensure that institutions apply them in a uniform and consistent manner.

This approach enables analysis of the degree to which institutions are making use of online teaching, differences between different programmes in their uptake of online delivery, the proportion and characteristics of students studying on line and, significantly, how the pass rates of students differ between fully online, partly online and fully in-person delivery, controlling for student demographic characteristics, level and field of study and other variables (Guiney, 2016[2]).

The examples from the United States and New Zealand show that the indicators produced through administrative data systems enable a country to estimate the extent of the take-up of digitalisation, with more granularity in the cases of systems enabling unit-level data collection.

Advantages and limitations of national administrative data systems

In most cases, administrative data systems cover the whole of the higher education system, counting every enrolment and every completion in every HEI. Therefore, unlike survey data, there is no margin of error (or sample error) and no sample bias in administrative data. Because it is an essential component of the management of the system, administrative data is collected to a consistent standard, with the core

variables unchanging from year to year. This provides continuity. It also creates the opportunity for government to link administrative data on higher education to data from other sources (such as labour market data) to provide for deeper analysis of the system's performance.

The purpose of administrative data is efficient management of an entity or system. It is factual data that derive from administrative transactions, also called events. In a higher education institution, the administrative data on transactions such as enrolment, fees payment and passing of courses is critical for the running of the institution. It also has considerable analytical value, for instance, enabling the institution to look at how different groups of its student population perform. However, such analysis is necessarily limited because while it can show that one group of students performs at a lower level than other groups, its ability to explain why is limited. The variables in the administrative dataset will not normally include items that give insight into attitudes, experiences and judgements.

The varying capacity levels of HEIs providing data is also an important consideration. In some instances, the task of reporting data to public authorities may fall to institutional staff members undertaking a wide range of tasks who do not have prior training or experience with administrative data collection and reporting, operating within small organisations with limited resources. Therefore, balancing and reducing as much as possible the reporting burden for institutions is an important lesson learnt from countries with complex administrative data systems, such as the United States.

Furthermore, the quality of administrative data may vary between fields; fields perceived by the institutions as important for their own institutional management purposes and those used by the government for its system management are likely to be best maintained.

The long-standing problem of defining and categorising courses according to their online content, as discussed previously, may have been exacerbated by the pandemic. The switch to online learning during the pandemic has given rise to growing interest in hybrid learning – combining online and in-person elements – even after in-person instruction becomes possible again. If most courses incorporate online components, distinguishing courses delivered fully or partly on line is likely to become more difficult to do. Near universal adoption of online delivery makes the question of how well digitalisation has occurred more significant than the question of whether digitalisation has occurred.

4.3. Surveys

Data collection approach and indicators

In some higher education systems, nationwide surveys are used as an important data collection instrument to complement other data sources on higher education. Surveys are often conducted to obtain a nuanced understanding of the experience of higher education students, graduates and employers, which helps calibrate public policies and institutional strategies according to the feedback of the key "users" of the higher education system.

Some systems regularly conduct student and graduate surveys to examine their higher education experience and satisfaction, as in **Denmark** and **Hungary** (Danish Ministry of Higher Education and Science, 2020[14]; Educational Authority, 2020[15]). Other systems, such as **Australia** and the **United Kingdom**, use surveys to collect feedback from employers and local stakeholders on the relevance of higher education (Australian Government, 2020[16]; UK Department for Education, 2020[17]).

These surveys may be census style, where every student and staff member is invited to participate, or may have representative samples of the targeted groups. While administrative data are collected and managed by HEIs and public authorities, survey data may be collected by other higher education stakeholders (e.g. student and teacher unions) and private companies, in addition to HEIs and public authorities.

Some countries opt for using surveys to collect evidence on the digital transformation in higher education. For example, in **Ireland**, the National Forum for the Enhancement of Teaching and Learning in Higher Education (National Forum) conducted the Irish National Digital Experience (INDEx) survey in 2019 (National Forum, 2020[18]). INDEx was a system-wide survey conducted in 32 higher education institutions (including 7 Irish universities, 12 institutes of technology and 13 private colleges/other higher education institutions), representing 96% of the entire higher education sector in Ireland. According to experts involved in the survey interviewed by the OECD team, the results have been used at the institutional and policy level to consider new approaches to support the effective use of digitalisation in higher education. A second round of the survey is currently under discussion.

The INDEx survey covered a broad range of questions about digital readiness, practices, and performance, from student and staff activities and experience in using technologies to digital infrastructure. In addition, it dealt with attitudes and preferences regarding digital learning and assessment.

Most of the INDEx survey questions were adapted from an existing survey – the Digital Experience Insights (DEI) survey used in higher education institutions in Australia, New Zealand and the United Kingdom (Beetham, Newman and Knight, 2019[19]; Jisc, 2020[20]), with the responses from students and teachers in those countries presented alongside the Irish results. Table 4.2 includes examples of internationally comparable indicators for students and "staff who teach" that have been highlighted in INDEx summary communications material.

Table 4.2. Selected internationally comparable INDEx and DEI findings

Theme	Students	Staff who teach
1. Digital teaching and learning practices	• Percentage that regularly accessed VLE • Percentage that agreed online assessments were delivered and managed well	• Percentage that regularly used VLE as part or all of their delivery • Percentage that agreed the institutional online system for marking and feedback was easy to use
2. Digital infrastructure	• Percentage that had access to recorded lectures • Percentage that had access to reliable Wi-Fi	• Percentage that had access to lecture capture whenever needed • Percentage that had access to reliable Wi-Fi whenever needed
3. Digital skills development and support	• Percentage that had access to Internet-based skills training • Percentage that agreed they had regular opportunities to review and update digital skills	• Percentage that had access to Internet-based skills training • Percentage that agreed they had regular opportunities to review and update digital skills
4. Digital environment and culture	• Percentage that agreed their institution protected their data privacy • Percentage that agreed they had the opportunity to be involved in decisions about digital services	• Percentage that agreed they were informed about their responsibilities on how to manage learner data securely • Percentage that agreed they had the opportunity to be involved in decisions about digital services
5. Attitudes to digital	• Percentage that rated the quality of institutional digital provision as above average • Percentage that rated the quality of digital teaching and learning on their course as above average	• Percentage that rated the quality of institutional digital provision as above average • Percentage that rated the institutional support to develop digital aspects of the role as above average

Note: International benchmarking is enabled by the use of the DEI survey, which is run in the United Kingdom (for students and staff) and in several universities in Australia and New Zealand (for students only) (Beetham, Newman and Knight, 2019[19]; Jisc, 2020[20]).
Source: National Forum (2020[18]), *Irish National Digital Experience (INDEx) Survey: Findings from students and staff who teach in higher education*, https://hub.teachingandlearning.ie/resource/irish-national-digital-experience-index-survey-findings-from-students-and-staff-who-teach-in-higher-education/.

Advantages and limitations of survey data

While rich administrative data gives a clear and comprehensive view of a higher education institution (or, in the case of the national administrative data collection, of the whole of the higher education system), it cannot provide a nuanced understanding of the practices and experiences of key higher education stakeholders – students and staff. For instance, administrative data can provide information about whether a student passed a course, but it is blind to a range of possible explanatory information – for instance, the person's experience of, or satisfaction with, the programme; whether the person is from a family where higher education is the norm and an expectation; or whether the person was in employment concurrently with study. That deeper exploration of students' backgrounds, attitudes and motivations and their experiences of and responses to the study environment is best managed through a survey.

At the same time, survey data is self-reported, and some questions require the respondent to make evaluative judgements. Responses may not fully reflect respondents' behaviours or experiences – they are impacted by memory and social context, meaning that there is a risk that two individuals with identical experiences of digitalisation and similar attitudes may respond differently to the same question (OECD, 2019[21]). In addition, surveys do not capture every member of the survey population; as a result, they will have sample error and the risk of non-response bias. Furthermore, there is the possibility of sample bias, where the response is more likely to occur from some groups in the survey population whose experience is different from the norm. There are means of mitigating the risk of sample bias, especially if the survey population is created using a robust sample frame (Statistics Canada, 2021[22]).

4.4. Learning analytics

Data collection approach and indicators

"Learning analytics" – or "educational analytics" – are defined as:

> *the use, assessment, elicitation and analysis of static and dynamic information about learners and learning environments, for the near real-time modelling, prediction and optimisation of learning processes, and learning environments, as well as for educational decision making. (Ifenthaler, 2015, p. 447[23])*

Learning analytics is receiving much attention as a promising tool to support student success, and a number of HEIs have used these systems to reduce failure rates, especially among disadvantaged groups. For instance, at Georgia State University (**United States**), predictive analytics have been used since 2012 to follow student performance. Over 40 000 students are assessed for a wide range of risk factors every day, and alerts are sent to both students and faculty when risks are identified, followed by one-on-one meetings to help the student improve. The results demonstrate both a decrease of more than a semester in average time to degree and an improvement in attainment for disadvantaged students (Georgia State University, 2018[24]; Georgia State University, 2021[25]). Similarly, at Purdue University, predictive analytics, plus the provision of support for those identified as at risk of failing, led to measurable improvement in pass rates. The same approach has been used in many universities in the **United Kingdom** and **Australia** (Sclater, Peasgood and Mullan, 2016[26]).

Meta-analyses show that using learning analytics can be successful in improving student pass rates, in particular among disadvantaged students, although with differences in extent according to the field of study, institution and other contextual factors (Ifenthaler and Widanapathirana, 2014[27]; Sclater, Peasgood and Mullan, 2016[26]; Wise and Cui, 2018[28]; Ifenthaler and Yau, 2020[29]).

Learning analytics can also be used for other purposes in HEIs – for instance, to compare courses and cohorts of learners and analyse attrition and track enrolments. But, most importantly, learning analytics is a tool that can be used to evaluate (and improve) pedagogical models (Wise and Jung, 2019[30]).

Data used in learning analytics are often derived from the use of learning management systems or virtual learning environments by students and staff. While LMS/VLE data are usually focused on a particular course, it is possible to link an individual student's LMS/VLE data from all of his/her courses to get a view of the student's engagement and progress across the whole of his/her programme of study. Furthermore, if used widely across an HEI, a LMS/VLE system can provide measures of how engaged students are in their learning and can be used by teachers to identify student difficulties or shape pedagogical decisions.

LMS and VLE systems provide data on the use of digital technologies by students and teachers and on their types of engagement with the digital technologies. The types of indicators that can be derived from learning analytics are diverse and include:

- Student scores, pass rates, retention.
- Student activity (also called transactions or events), such as student engagement measured through a login or the opening of a document/viewing of a video, the use of a chat room, the time spent in viewing or reading, including times at which student attention drops, the taking of a quiz and the submission of an assessment. These are examples of *non-reactive* data that can be mined from an LMS/VLE and that are available in near-real-time.
- Students' opinions, for example, through satisfaction surveys embedded in LMS/VLE systems (these types of data are also referred to as *reactive data*).

LMS/VLE data can be linked to administrative data held by the HEI, such as data on students' demographics, prior educational achievement and entitlements, and/or to other institutional data, such as data drawn from responses to student surveys. Linking such data can help with targeted student support interventions, as discussed above. Furthermore, such data-linking processes can be automated so that deep analysis can be performed and results communicated to students and instructors promptly, at a relatively low cost.

While primarily focused on supporting student success, learning analytics data can provide a wealth of information on student behaviours, engagement and satisfaction with digitally enhanced higher education, which, combined with data on student success, holds significant potential to shed light on both the digital practices of students (and of staff who use LMS/VLE systems) and on the digital performance of higher education.

Advantages and limitations of learning analytics data

Learning analytics has some of the characteristics of administrative data in that it records transactions and events during a student's study of a course. The data can be used to create measures of student engagement with digitally enhanced teaching and learning. It can also incorporate surveys (allowing for the creation of measures of learners' experience and satisfaction with online learning). In addition, it creates the opportunity to develop proxy measures of the effectiveness of digitalisation through comparing student achievement – such as completion rates or assessment outcomes – across different study modes (controlling for factors like prior educational achievement).

While learning analytics data contains elements of the two other forms of data – administrative and survey data – it differs from those two other forms in that it draws from LMS/VLE data specific to a course. It can, however, be aggregated in some circumstances. An institution that wants to exploit the potential of learning analytics for the purpose of improving learner success needs to ensure high levels of take-up of the LMS/VLE across the HEI so as to generate comprehensive learning analytics data. While students may be required to use the LMS/VLE by teachers, some teachers are reluctant to use all of the functionality of these systems and, in some cases, to use the LMS/VLE at all (Weaver, Spratt and Nair, 2008[31]). In addition, the richness, complexity and volume of the data generated in an LMS/VLE may make it a challenge to analyse and use for decision-making purposes (DSN/DHECC, 2020[8]).

A further issue with learning analytics data, which is recorded at the course level, is that it depends on each instructor's requirements regarding LMS/VLE use. This is determined in part by the teacher's confidence, interest and capability in using these systems and in part by the nature of the field of study and course material. This may make it difficult for an institution to establish a base configuration for its LMS/VLE rich enough to enable LMS/VLE data to be aggregated to create meaningful indicators of engagement, experience and effectiveness. However, the Georgia State University example (referred to earlier) is just one example where an institution has been able to demonstrate what can be done (Georgia State University, 2018[24]).

If a country wants to use data analytics nationally to create national measures, it needs wide take-up in all institutions, and it needs to create a consensus among institutions on the configuration of diverse systems so that each institution produces data sufficiently comparable to allow for aggregation.

4.5. Summary, complementarity and common issues

The section that follows provides two summary tables of the preceding discussion:

- Table 4.3 provides a summary of the types of indicators on the digitalisation of higher education that may be generated by national administrative data systems (using IEPDS as an example), surveys (using INDEx as an example) and learning analytics.
- Table 4.4 compares the strengths and weaknesses of administrative, survey and learning analytics data and incorporates some comments on potential costs, ease of implementation and repeatability of data.

It then discusses the value of using the three approaches in a complementary manner to obtain a rich understanding of the digitalisation of higher education. Finally, it closes with a discussion of data privacy and use concerns, which are relevant for all data collection approaches.

Summary of administrative, survey and learning analytics data

Table 4.3. Digitalisation indicators generated by administrative, survey and learning analytics data

Digitalisation aspect	Existing indicator type	Areas for growth
Administrative data (Example: The United States' IPEDS)		
Digital readiness	Digital library holdings may be seen as a form of digital readiness indicator (availability of technology).	Two areas for growth are: • A move to a unit-record student data system would allow for greater detail in the analysis of online education and allow for online education performance to be analysed, controlling for demographic and study-related factors. • A move to collect additional data on digital transformation that can be captured through existing administrative data processes, e.g. asking institutions to report on the usage of an LMS/VLE would be useful. These growth areas must be measured against the additional administrative burden and cost that it would require from HEIs.
Digital practices	Data on the enrolment of students in online delivery (proxied by distance education) and the course types that an institution provides through online education.	
Performance in digital higher education	IPEDS does not directly capture data on the performance of distance education or other digital indicators, though some correlational insights can be drawn for institutions substantially engaged in distance education delivery.	
Survey data (Example: Ireland's INDEx)		
Digital readiness	Data are captured on the current use of technology by staff and students and on potential barriers (such as access to devices, levels of skills and knowledge, as well as provision of essential infrastructure).	While future fielding of the INDEx survey is not confirmed, interest in the Irish higher education sector is high. The impact of COVID-19 will have a significant impact on the results from any future survey. Continued stakeholder discussion of findings will likely lead to the identification of opportunities for further development. One limitation to the current survey is that institutional results are not published, which limits some of the potential insights, e.g. it is not possible to compare survey outcomes to institutional resources or directly compare institutional approaches and outcomes. However, the fact that data are available to HEIs privately is already resulting in policy changes.
Digital practices	Data are captured on the use of technologies by staff and students, including the use of a LMS/VLE, the design of course content for online delivery, and the use of tools such as mobile applications and "gamification" of learning.	
Performance in digital higher education	Quality and equity can be evaluated by: • assessing equity in areas such as device and service access by a demographic analysis • responses to queries regarding the quality of digital services (e.g. ease of use of the LMS/VLE). However, INDEx data is not released at an institutional level, so there are limitations (e.g. comparing HEI spending against responses on the quality of service would help to gauge efficiency).	

Learning analytics		
Digital readiness	Adoption and use of LMS/VLE systems is an indicator of digital readiness. LMS/VLEs provide data on student and staff engagement (e.g. number of courses using LMS/VLEs, students active on LMS/VLE).	Learning analytics is a primary area for growth in higher education private-sector data. However, the utilisation of the full potential of learning analytics among institutions using LMS/VLEs and SIS appears limited, particularly in Europe. Learning analytics (and in particular predictive factors) is an opportunity to improve student outcomes, the design of programmes and pedagogical approaches, as well as the overall quality of higher education. Data privacy, ownership, and appropriate-use regulations and concerns need to be addressed. A further area for growth is the possibility of aggregating learning analytics data from multiple institutions to inform system-wide higher education policy. This will require significant investment both in the development of system-wide data infrastructure, agreement on and specification of common measures, interoperability between proprietary systems, and the engagement and buy-in of higher education stakeholders.
Digital practices	LMS/VLEs and student information system (SIS) systems together track a range of digital behaviours, such as the use of digital resources, e.g: • uploads and downloads • interactions with online discussions • staff practices and usage of the LMS/VLE.	
Performance in digital higher education	LMS/VLEs can capture student experiences on the quality of education delivered by embedding surveys as part of the LMS/VLE. Linking LMS/VLE data to SIS data allows for data mining and statistical analysis to: • identify risk factors for academic failure • evaluate the performance of pedagogical approaches • analyse aspects of equity and efficiency.	

Table 4.4. Comparing the strengths and weaknesses of administrative, survey and learning analytics data

	Administrative data	Survey data	Learning analytics data
Description	Administrative data reports on transactions or events (such as enrolment and completion), so it can be used to report on participation and success in online learning. It can give a view of how much online learning occurs, how many students and teachers participate, pass rates in online learning as opposed to other forms of delivery. It cannot report on how engaged students are with online learning or on their experience of online learning.	Survey data can report on how students (and teachers) experience online learning and their satisfaction with it.	Learning analytics reports on transactions, such as logins, use of chat rooms, time spent reading/viewing material, assessment, etc. It can incorporate surveys. It can look at participation in online learning, engagement with it, success in it, and its effectiveness.
Coverage	Comprehensive: every student is counted. Administrative data on staff is separate from data on students. Administrative data is collected at the institutional level and nationally.	Student and staff surveys can be complementary – asking how each used/experienced the same tools. There can be census-style surveys, where every student/staff member is invited to participate, or the creation of a sample frame leading to a survey of representative samples of students and teachers. Surveys can be used at the institutional level or across groups of institutions (for instance, institutions in the United Kingdom, Australia and New Zealand can opt to take the DEI survey), or nationally, like the INDEx survey in Ireland.	The data originates at course level, where it is comprehensive. However, in some HEIs, teachers are free not to use the LMS/VLE or not to use some of its functions. This means that the learning analytics data may not be comprehensive at an institutional level in some institutions. To aggregate LMS/VLE data at the institutional level, the LMS/VLE must be configured to produce common data elements, with teachers free to use additional features of the LMS/VLE if they wish. It may be difficult to develop a national view of engagement with online learning through learning analytics, given HEI autonomy in most jurisdictions.
Frequency and timing	The frequency of collection of administrative data depends on the business cycle of HEIs (and, in the case of national data, of government requirements). It is usually repeated annually (or more frequently).	Surveys can be repeated at regular cycles, annually or less frequently.	LMS/VLE data can be analysed annually, each semester or more frequently. There is the opportunity to automate the reports that the LMS/VLE produces at the course and institutional levels.
Compliance burden	Any change in the specification of administrative data requires a long lead-in time as institutions need to arrange for its collection at enrolment time and may need to modify their student information management systems. This imposes a high compliance burden on HEIs. This applies whether the change is initiated at the institutional or national level. The burden on students is relatively slight.	Surveys are likely to impose relatively low compliance burdens on HEIs. If the survey is repeated frequently and is a census-style survey, then there is a compliance burden on respondents that could lead to lower response rates in the medium term.	Setting up, configuring and using LMS/VLE systems imposes a relatively slight burden on HEIs, staff and students. However, generating high-quality learning analytics data – aggregating data on a student across all of his/her courses, linking LMS/VLE data to the institution's core administrative data, developing and programming the production of indicators and dashboards and aggregation of data across an institution's courses – can be burdensome. Reports can be automated, however, so the data generation burden can lessen once the setup is complete.

	Administrative data	Survey data	Learning analytics data
Cost and flexibility	The cost of managing administrative data – institutionally and nationally – is high. The cost is especially high if all HEIs have to supply unit-record data. The cost of change is high if the change involves adding a new variable, especially if the programming of the student management system changes. A change in the national administrative data collection, however, involves a low cost if the collection is of aggregated data and if there are no new variables. The cost of change makes administrative data relatively inflexible.	The cost of designing surveys, setting them up in survey administration software, piloting, launching a survey, monitoring responses, promoting responses and analysis of data is high. Surveys can be modified for subsequent iterations relatively easily and at relatively low cost.	The cost of a LMS/VLE is high. There is a high cost in configuring and developing the reporting needed for the development of learning analytics. The ongoing cost of learning analytics is relatively low. The complexity of the development of rich learning analytics data means that modifying the system may be costly.
Data quality and reliability	Administrative data give high levels of certainty as long as data quality is managed well. Data have high quality and are recorded in a consistent manner between individuals but cover only events/transactions (such as enrolments, completions and fees payment), not judgements, assessments or evaluations. Questions about response rate and sample bias do not arise.	Surveys always have sample error and non-response bias. However, if the survey is large enough, well designed, and implemented well, these are relatively slight. There is a risk of sample bias. Also, survey data rely on self-reporting and the evaluative judgement of respondents. So, two individuals with identical experiences could describe the experience differently.	As long as the measures and indicators are rich and well-designed and if the take-up of the LMS/VLE by teachers is high, the data is likely to have high levels of certainty.
Relevance and impact	Administrative data collections are designed to capture what occurs, but the important qualitative variables are likely to not be present – attitudes, motivations, experiences.	Surveys can explore qualitative background variables and can look at attitudes to digital tools. In addition, they can take behavioural and evaluative perspectives, such as looking at what tools were used, how they were used, and how well they worked. They can look at how well students and teachers were supported in their use of digital tools.	As long as the configuration and data mining behind the reporting are well designed, the learning analytics data can provide highly relevant and timely insights into students' learning behaviours and outcomes.
Longitudinal potential, i.e. the potential to track change at an individual student level over time	Institutional administrative data on students and staff are necessarily longitudinal, given the business needs of institutions. National administrative data cannot be made longitudinal if collected as aggregated data but can be used to create time series and cross-sectional analyses. National administrative data collected at the unit-record level can be used for longitudinal analysis if there is a personal identifier (such as a national student number) or through probabilistic matching of students between different years' datasets.	Survey data are usually used for cross-sectional analysis. It is possible to create a longitudinal view by returning to the same survey population (or to a sample drawn from that population) over several years.	These data, like institutional administrative data, are longitudinal.

	Administrative data	Survey data	Learning analytics data
Overall challenges and weaknesses	Cannot report on matters like the experience of or satisfaction with digitalised learning. The pandemic has reduced the value of administrative data measures (because most courses now have a substantial online component, reducing the ability to differentiate between courses). A national unit-record collection imposes high costs on smaller HEIs. The cost of change – especially if it involves adding a new variable – is high. This reduces the flexibility of administrative data.	Survey data has a level of uncertainty deriving from sample error, especially if the number of responses is low. Online surveys implemented through standard, commercially available survey software may be less expensive than administrative data. However, survey design, piloting, monitoring responses, promoting responses and analysis of data carries high costs.	The main challenges are: • to ensure that most, if not all, instructors in an institution use the LMS/VLE • to develop a LMS/VLE configuration that ensures the collection of a common core of data that is uniform across all courses in an institution, without denying early adopters the opportunity to exploit the advanced features of the LMS/VLE • to develop the links between the LMS/VLE data from all the institution's courses and to the institutional administrative dataset that allows for rich data analytics. It is very complex to move to a national collection based on LMS/VLE data.

Complementarity of administrative, survey and learning analytics data

Administrative data, survey data and learning analytics data can be viewed as complementary, rather than alternative sources of data to shed light on the digitalisation of higher education. For example, administrative data can provide information about higher education activity but cannot shed light on the way students experience their courses and programmes – one of the important dimensions of the quality of higher education, including in a digital environment. On the other hand, survey data can provide rich data on experiences and satisfaction, while learning analytics create an opportunity to observe student learning practices.

Administrative data are comprehensive – there is no sample error or sample bias. On the other hand, survey data require students to make interpretations and judgements (which may mean there are differences between individuals in the way they describe identical experiences), involves sample error, and may contain sample bias. Learning analytics require a significant breadth and depth of LMS/VLE usage in order to generate useful data for analysis.

The differing strengths and weaknesses of the three types of data mean that they can be used to complement each other. For example, there would be value in data reporting that gives a profile of the use of online learning across an institution (or nationally) drawn from administrative data, alongside the results of a survey that explores student and teachers' experience of online learning and learning analytics that provide insight on students' learning practices.

Another way administrative, survey and learning analytics data can be complementary is in dealing with additional information requirements. For example, adding a new variable to a national administrative data collection can be very costly, as all HEIs must undertake expensive programming of their student information management systems. That means that it is worth adding a new variable only if the government is certain that the new variable is needed for the long term. If not, and if there is a national survey, it is relatively less expensive to use that survey to explore the additional data. If learning analytics are widely used, these can also incorporate surveys.

Data privacy and use

The issue of privacy has been a concern for any kind of data collection in the last decades. In the case of administrative data, HEIs are responsible for managing data within national and supranational data protection standards – such as the European Union's General Data Protection Regulation (European Commission, 2021[32]) or the privacy and data protection legislation in the relevant jurisdiction. If an institution contracts out the housing and management of its administrative data, it must ensure that the contractor maintains those standards. Institutions have an obligation to keep data secure and to control access to the data, releasing only what is necessary to administer their work and to comply with national and legal reporting obligations. Institutions need to be explicit in disclosing to students what they will use the data for and who it could be shared with, and if so, which variables, for what purpose and in what form. Governments must also comply with such standards if they collect unit-record administrative data.

As with administrative data, survey owners are also responsible for managing the data they gather within national data protection standards. They, too, have an obligation to keep data secure, and if they contract out the management of the survey, they need to require the contractor to comply with those standards. Again, as with administrative data, those conducting a survey need to be explicit in disclosing to respondents the purpose of the data collection and how the data will be used and to whom it will be disclosed. Surveys that allow respondents to decline to respond to a given question while completing the rest of the questionnaire offer a greater level of privacy.

Learning analytics data represent a subset of institutional data and need to be subject to the same privacy and management standards as other HEI data. Data security and privacy may require further investment to ensure that legal and ethical standards are met (Jones, 2019[33]; Ochoa, Knight and Wise, 2020[34]).

Further, learning analytics relies on algorithmic processing that builds on the choices and judgements of the designer and on statistical generalisations that may "lose sight of" context and unmeasured variables (such as traits, attitudes and motivations) (Wise and Cui, 2018[28]). These concerns mean that those who develop the learning analytics systems need to be aware of the limitations, while follow-up interventions such as those used at Georgia State University need to be designed in a way that prompts thought and discussion by the student and advisor (Georgia State University, 2018[24]; Wise and Jung, 2019[30]).

4.6. Higher education data collection in Hungary and considerations for the development of indicators

Data collection in Hungarian higher education

Higher education data

In Hungary, system-level higher education data collection has increased in recent years, in particular through the expansion of the administrative data system and the use of surveys of students and HEIs. However, while most HEIs use a learning management system, surveys and studies indicate that learning analytics are infrequently used (DSN/DHECC, 2020[8]).

With regard to administrative data, several public databases store extensive higher education data. The **Higher Education Database and Information System** (FIR) is a national registry containing the majority of administrative data on Hungarian higher education. It is managed by the Educational Authority (OH) and includes, for example:

- data on HEIs and their programme offerings, such as the number of HEIs by type and the number of study programmes by level
- unit-record data on students and their mode of enrolment and staff by types of contract, such as their characteristics (gender, nationality, etc.) and registration status (full-time/part-time)
- data on digital infrastructure at HEIs, such as the number of computers and access to the Internet.

The FIR was established following the implementation of the 2011 CCIV Higher Education Act as a national database for higher education. Before the development of FIR, HEIs submitted administrative data to several data collectors, such as the National Health Insurance Fund, the Hungarian State Treasury and the Central Administration of National Pension Insurance, while students and staff participated in ad hoc surveys (Educational Authority, 2018[35]). Hungarian HEIs are obliged to provide data to the FIR system, which public authorities use to manage the higher education system. For example, state funding to higher education relies to a large extent on FIR data (DSN/DHECC, 2021[36]).

The **Database on Student Stipends** (HÖSZ) holds financial data on students whose studies are fully or partially covered by state support. The **Online Library of Hungarian Academic Works** (MTMT) also stores information on academic publications and is connected to a global citation database, Scopus (DSN/DHECC, 2021[36]). In addition, the **Adult Education Reporting System** (FAR) keeps a list of short, non-degree education programmes (Adult Education Reporting System, 2021[37]).

Furthermore, OH administers the **Graduate Career Tracking System** (DPR), which combines survey data on graduate labour market outcomes with administrative data (from FIR, HÖSZ, the National Tax and Customs Administration, the National Health Insurance Fund, and the Ministry for Innovation and Technology [MIT]) (Educational Authority, 2020[15]).

Public HEIs use the **NEPTUN** student information system (SIS), while private institutions are free to select a SIS of their choice. In addition, HEIs use **information management systems** to collect and store data

concerning institutional management, such as financial and human resources management. HEIs also submit an **institutional development plan** to MIT, in which they set goals for the next five years.

Hungarian HEIs use the **LMS/VLE** of their choice – many of them using Moodle or Blackboard (both widely used systems internationally) or the Hungarian system CourseGarden (DSN/DHECC, 2021[36]).

Data on the digitalisation of higher education

While data systems described above offer comprehensive information to support higher education policy making and institutional planning and management, data collection concerning the digitalisation of higher education appears limited (DSN/DHECC, 2020[8]). FIR data on study programmes, for instance, do not refer to modes of delivery, i.e. whether instruction is on line, hybrid, or in person.

Current evidence on digital transformation in higher education is mainly collected through ad hoc surveys in Hungary. The National Union of Students in Hungary conducted **a student survey** shortly after the transition to emergency remote learning in spring 2020. More than 17 000 students participated (12 000 student responses were used in the analysis), with a majority of undergraduate students responding. Students were asked to provide their views on their online education experiences, including their level of satisfaction with online learning and preference between online and in-person settings (HÖOK, 2020[38]).

The Ministry for Innovation and Technology commissioned **two surveys on digital higher education** in the fall of 2020, administered by the Digital Higher Education Competence Centre. The first survey was carried out in September 2020 and sought institutional leaders' views on factors determining HEIs' level of digitalisation, including external factors (e.g. students' digital skills) and internal factors (e.g. access to digital infrastructure at an HEI, teachers' digital skills, etc.), with a view to identifying ways to monitor digitalisation in Hungarian education. The participating institutions were also asked to share their digitalisation practices (e.g. creation of digital content, e-learning support services, updating of pedagogical methodologies, digital dissemination of research outputs). The second survey was conducted in November 2020 to collect data on access to digital infrastructure at Hungarian HEIs, such as high-speed Internet access and the availability of digital tools. For both surveys, responses were collected from over 85% of all accredited institutions (DSN/DHECC, 2021[36]).

In addition, the OECD conducted a **higher education stakeholder consultation survey** in February-March 2021 as part of the present project to obtain information about digital practices from higher education students and staff. Completed responses were submitted by over 1 000 higher education stakeholders (629 students, 354 teachers, 38 leaders, 3 policy makers, 5 staff from non-governmental organisations and private companies, and 10 others). The survey asked about the access to and use of digital infrastructure and data systems and about students' and teachers' experiences of digitally enhanced teaching and learning. It also collected stakeholders' views on public policies and institutional practices supporting the digital transformation of higher education (Annex B).

Those surveys shed important light on the current digital readiness and digital practices of Hungarian higher education. However, data on digital performance remains limited. For example, the OECD survey asked about students' and teachers' satisfaction with online teaching and learning. However, its data are provided based on the experience of "emergency" remote learning and may not accurately present the performance of digital higher education in Hungary.

The use of learning analytics, while taking place in some institutions, does not seem to be widespread in Hungarian higher education. However, the wide use of LMS/VLE creates a source of data, which, in conjunction with the SIS data, provide the opportunity to create rich information on digital practices and performances. With the pandemic having led to greater use of LMS/VLE in Hungary, the potential value of learning analytics in the Hungarian higher education system has grown. According to the OECD survey, while around 40% of student respondents reported having access to LMS or VLE before the pandemic, an

additional 40% reported getting access to these tools since the start of the pandemic. The survey also shows that two-thirds of student respondents reported having used a LMS/VLE at least weekly (44% daily and 25% weekly) at the time of the survey (February-March 2021) (see Annex B).

As noted in the previous section of this chapter, administrative data, self-reported data from surveys and trace data from LMS/VLEs have different advantages and drawbacks. While administrative data presents the advantage of reliability and broad coverage, it is not as timely or as rich as learning analytics data. Administrative data covers mainly transactions or "events" and does not give information on students' or teachers' experience of digitalisation or on the quality and effectiveness of digitalisation. Survey data helps understand the behaviours and motivations of students and teachers but is self-reported and comes with sample error. Trace data from LMS/VLE is data that is generated by the real-time use of digital technologies, such as the opening of a document or time spent on a webpage. This data offers reliable accounts of digital technology use, but it can only be analysed when students and teachers regularly use LMS/VLE. Combining different methods is thus the most promising approach to assess the digital transformation of Hungarian higher education.

Options for further data development

Defining the purpose of data collection

Evidence on digital readiness – infrastructure and policies that maximise the take-up of digital technologies in higher education – and on the digital practices of students and staff in HEIs is important to understand the scale, pace and effectiveness of digitalisation in the Hungarian system.

Evidence on digital performance – on the equity, quality and efficiency of digital higher education – is needed to monitor whether digital higher education is designed and delivered in a way that maximises the benefits of digital technologies in higher education while mitigating its risks. The benefits of digitalisation can be considerable – from greater access to diverse and flexible learning options to the individualisation of learning and the development of more effective data-informed teaching methods. But there are also important risks: in particular, disadvantaged students are at risk of falling further behind because they may lack adequate equipment and learning attitudes to do well in an online environment.

As Hungary considers new data development to monitor the digital transformation of higher education, it needs to clearly identify the:

- purpose of new data (examples of potential goals and the types of indicators that might be most relevant are illustrated in Table 4.5)
- level at which data is needed, be this at a system, institution, course or student level
- data collection methods most suited for the purpose, given different advantages and drawbacks of each method
- possibility of collecting the new data as an add-on to the existing extensive data collections
- trade-offs between the benefits of new data collection and the burden of establishing data specifications and developing collection and reporting processes
- ways in which HEIs are incentivised (or required) to collect and report data
- capacity in both HEIs and government to develop adequate data systems
- capacity in both HEIs and government to utilise the data for the purpose they have identified.

Table 4.5. Potential purposes of data collection and potential types of indicators

Potential purposes	Potential types of indicators
Inform public policies and institutional strategies	
Inform digital infrastructure policies • *At national level*: Consideration of broadband infrastructure, capital expenses for digital infrastructure and equipment, investments in the creation of digital content, the purchase of digital equipment, research and development (R&D) policies regarding the use of advanced digital technologies in higher education • *At HEI level*: Consideration of budget for digital infrastructure and equipment, governance structure to manage the identification of needs and purchasing of equipment	• Quantitative indicators on infrastructure, student and teacher access to suitable digital infrastructure and content, and student and teacher use of digital technologies in their studies/work • Qualitative information about HEI leaders' perceived challenges and opportunities related to digital infrastructure
Inform policies on digitally enhanced higher education offerings • *At national level*: Consideration of legislation/regulation/quality assurance requirements on the structure/content of courses and programmes, funding policies promoting the development of digitally enhanced courses/programmes and other learning opportunities • *At HEI level*: Consideration of institutional strategies and investments regarding the offerings of digitally enhanced courses/programmes and other learning opportunities	• Qualitative information about the existence of public policies and institutional strategies supporting digitally enhanced higher education offerings • Quantitative indicators on the supply of digitally enhanced courses, programmes and other learning opportunities
Inform policies related to digitally enhanced teaching, research and engagement • *At national level*: Consideration of legislation/regulations/quality assurance requirements related to the profiles, careers and work of teaching staff, policies regarding professional development, institutional funding promoting the development of pedagogical practices adapted to the digital environment, provision of financial supports to teachers to participate in digital skills training programmes • *At HEI level*: Consideration of institutional strategies related to the profiles, careers and work of teaching staff, policies regarding professional development	• Qualitative information about the existence of public policies and institutional strategies supporting digitally enhanced teaching, research and engagement • Quantitative indicators on the use of digital technology and reported satisfaction by higher education staff • Quantitative and qualitative indicators on the quality and effectiveness of digital courses (including how students perform in digitalised courses, relative to others, and related to background variables, and measures of student engagement) • Quantitative and qualitative information about the barriers to technology take-up among higher education staff and their views on opportunities for improvement
Inform policies related to digitally enhanced learning • *At national level*: Consideration of institutional funding to HEIs supporting investments in student supports, provision of student financial supports to participate in digitally enhanced courses/programmes and other learning opportunities (e.g. alternative credentials) • *At HEI level*: Consideration of the provision of academic/career/personal advising provided in a digital environment	• Qualitative information about the existence of public policies and institutional strategies supporting digitally enhanced learning • Quantitative indicators on the use of digital technology and reported satisfaction by higher education students • Quantitative and qualitative indicators on the quality and effectiveness of digital courses • Quantitative and qualitative information about the barriers to technology take-up among higher education students and their views on opportunities for improvement
Inform teacher and student practices	
Inform teachers of students' practices and outcomes to provide targeted student support and/or adjust their own pedagogical practices	• Quantitative data on students' learning practices, learning outcomes (e.g. retention, pass rates, grades) and teachers' pedagogical practices • Quantitative and qualitative indicators on the quality and effectiveness of digital courses • Qualitative information about barriers to teaching and learning performance in a digital environment according to both teachers and students
Inform students about their own learning practices and outcomes to foster self-awareness and improvement	

Strengthening data use

One key challenge facing Hungary as it considers collecting data on the digitalisation of higher education is its capacity to use this data.

While Hungary's national data on higher education is already very rich, the use of data in policy evaluation and policy research is limited. Even at the institutional level, the use of data to support decision making appears "rare and undeveloped". Hungary is taking steps to manage and derive value from the large datasets it holds – specifically through the creation in 2020 of the National Data Asset Agency (DSN/DHECC, 2020[8]). However, plans for new higher education data collection should specifically outline how data use could be extended, identifying current gaps limiting the use of data, and the support (including human and financial resources) needed both at the national and institutional levels to make better use of data.

Building on existing data

Hungary's comprehensive approach to higher education data collection is based on the FIR, which is set up in legislation. Links also exist between the collection and reporting of data by HEIs and public funding through the HEIs' institutional development plans that draw on FIR data and other data provided by the HEI.

Adding digitalisation-related indicators to the current administrative data system could offer rich evidence on digitalisation of higher education at a national level, noting, however, that with the increase in the uptake of online learning resulting from the pandemic, some of which may continue in future, particular attention will need to be paid to providing clear definitions of what constitutes digitally enhanced teaching and learning.

The benefits of this approach would need to be considered in light of the feasibility of introducing new variables into a complex data collection system, possibly requiring changes in all HEIs' student management systems. The technical feasibility and the human and financial resources implications of such an approach should be considered carefully. Immediate costs should also be assessed against the long-term benefits of regular administrative data collection. The policy levers that the government intends to employ to incentivise HEIs to collect and report this data must also be identified.

Regular system-wide surveys of higher education students and staff would be important tools to collect qualitative information on the perspectives of students and teachers on online teaching and learning experiences and monitor change over time. Here too, the costs and benefits should be carefully weighed. The option to build upon existing, regular surveys of current or recent students (e.g. annual survey of graduates' labour market outcomes) could be explored to minimise the costs of creating new survey tools. International experience in the area of student and staff surveys should also be considered (e.g. Ireland, Denmark and Australia). In addition, the experience of the National Union of Students and the Digital Higher Education Competence Centre, which implemented surveys focused on digitalisation in 2020, should provide insights into approaches to surveying HEI leaders as well as students to monitor progress in digital teaching and learning. It would also be important to gather views from higher education staff, who are key actors in the digitalisation – its scale and depth – of higher education in Hungary, as discussed in Chapters 2 and 3 of this report.

Learning analytics may be a rich source of data to complement system-level administrative and survey data by providing data on the use of digital tools and student learning outcomes. The wide variation in the use of learning analytics between and within HEIs suggests, however, that learning analytics may be primarily a source of information for individuals and departments/faculties within HEIs who use these systems, and at the institutional level for HEIs that use them broadly. Thus, obtaining a system-level picture would require broad usage of LMS/VLE systems within and across Hungarian HEIs. It would also require consensus on the types of data to be collected and an agreement by all HEIs to configure their LMS/VLE

to collect that information (without constraining the ability of expert users of the LMS/VLE to extract deeper, richer data of value for their [and their institution's] practices).

Several approaches would need to be pursued to encourage the use of learning analytics in Hungary. This includes clear standards that HEIs can use as they work with providers of LMS/VLE (whether external or in-house) to protect student data and clarify its uses. It also includes ensuring that academic and professional staff have the skills to make use of learning analytics and identifying the incentives that drive individuals and HEIs in using learning analytics. Finally, insights from HEIs and systems where learning analytics have developed the most internationally would be important for Hungary to consider.

Research may also be commissioned to better understand the use of learning analytics at present in Hungary, to understand the current state of learning analytics use, barriers to their further take-up, and opportunities to increase use. For example, **Australia** and **Germany** have been successful over the past decade in supporting the digitalisation of HEIs through research and development grants, which produced empirical evidence and helped change pedagogical practices using digital technologies at individual institutions.

Combining data sources may also offer important insights. Taking Hungary's Graduate Career Tracking System (DPR) as a model, the combination of administrative and survey data may offer a solid evidence base for Hungarian digital higher education.

It would also be important to consider how data collected could support several levels of analysis. For example, indicators developed to provide a national view of digital readiness, practices and performance in Hungarian higher education may be designed to permit the reporting of data nationwide, and per HEI, to inform national-level policy making. HEI-specific indicators may also be envisioned by HEI themselves, based on their areas of interest.

Potential indicators for Hungary

Given the broad scope of digitalisation in higher education discussed in this report, a number of indicators could be relevant to measure the digitalisation of higher education in Hungary.

To assist the Hungarian government and higher education stakeholders in monitoring the digitalisation of higher education, a preliminary list of 30 potential indicators that can be used to measure progress over time at the institutional and national level have been compiled. The list is presented in three tables:

- Table 4.6 contains digital readiness indicators.
- Table 4.7 contains indicators on digital practices.
- Table 4.8 contains indicators of digital performance.

The possible indicators were developed: 1) based on the analytical framework developed for the project that considers digital readiness as well as digital practices and digital performance; 2) building on international experience; and 3) taking Hungary's current data systems into account.

The indicators have been designed:

- to establish a baseline index of the state of digitalisation at a national level and then to measure progress over time
- to provide a measure of each institution's situation in a way that can be aggregated to provide a national view
- to compare progress in digitalisation in different parts of the higher education sector (either between HEIs or HEI groups or in types of programmes) in Hungary
- by making a link, where possible, to indicators used internationally, to provide a basis for comparison with other countries' state of digitalisation.

Indicators have only been proposed where it is likely that they can be populated at relatively low cost. However, some of the measures will depend on a national survey of higher education students and teachers that can explore the state of digitalisation. That survey would need to be developed and run to establish the baseline and then administered at regular intervals to measure change over time. Other indicators – for instance, those that look into outcomes for graduates – would need a detailed analysis of existing national administrative data. Some indicators would require the use of learning management system data.

Publication of the results of the indicator set should be accompanied by a clear, descriptive summary of the state of the alignment between Hungary's higher education policy framework and the needs of a digitalised higher education system.

Such a summary needs to address some of the most important issues identified in Chapters 2 and 3 as hindering the adoption of digitalisation. This would mean:

- ensuring the funding system is neutral between online and in-person delivery and that it supports the development of the capabilities of students and staff
- ensuring that the funding system provides support for digital equipment, teaching, research, and engagement and learning in a digital environment
- ensuring the accreditation and quality assurance practices and requirements are neutral between online and in-person delivery
- identifying the criteria for assessing teacher performance to respond to the need for teachers to master digitally enhanced teaching
- setting employment conditions for higher education teachers that allow and encourage them to take on professional development that provides the skills needed for delivering and assessing online learning
- ensuring transfer of credit arrangements are neutral between prior learning obtained via online learning and in-person
- providing information about government support for innovations, such as micro-credentials, open educational resources and open science.

In addition to listing the possible indicators, the tables contain comments, based on information available to the OECD team, on potential data sources that may be considered in Hungary to collect data on these indicators.

The list is deliberately extensive and aims to be a starting point as Hungary's public authorities and higher education stakeholders begin the development of a system to monitor the digital transformation in the nation's higher education system.

Notes on indicators tables

1. The possible indicators are designed to be recorded at an institutional level in a way that allows aggregation to give an indicator of the progress towards digitalisation of higher education across Hungary.
2. Where appropriate, the indicators have been designed to align with the Irish INDEx survey. Wherever possible, the questions have been phrased in a way that means that they could be answered in other OECD countries.
3. Indicators that could be populated only through complex interrogation of systems (for instance, questions about the percentage of operational expenditure devoted to supporting online delivery) have been avoided.

4. Indicators that look at the enrolment and completion of students should be reported disaggregated by student characteristics (e.g. gender and regional or socio-economic grouping) and also by study characteristics (level and field of study) to ensure that differences are not misattributed to online status, when the driving factor may be student-linked or course-linked.
5. As noted earlier, categorising courses and programmes will pose particular challenges in a post-pandemic context where "fully on line", "partially on line" (blended or hybrid) or "fully in person" may no longer be granular enough to understand the types of courses and programmes provided, as online learning is becoming an increasingly prevalent component of most programmes, and possibly courses. A proposed approach in the following indicators is to use four categories, rather than three, as follows:
    a. **Courses**: The variable considered is time spent on line as part of a student's "total theoretical study time", which could include both synchronous and asynchronous course-related activities. Such an approach would require departments or individual faculty members to make determinations of the course online status and for these to be recorded in the HEI's data systems. The four categories could be: a) 50% or more of the student's total theoretical study time is to be spent on line; b) 26-49% on line; c) 1-25% on line; or d) fully in person. Indicators A4, C1 and C2 use this proposed categorisation.
    b. **Programmes**: The variable considered is the share of courses a student takes according to the course online status, as discussed above. Because students may have the option to complete the same programme using a different mix of online, blended and in-person courses, the extent to which a programme is on line or in person is a characteristic of the student's enrolment rather than of the programme itself (i.e. two students in the same programme could select different courses and hence, have a different online profile).

       A categorisation of student enrolment in a programme could follow a similar logic as for courses, such as: a) 50% or more of the student's courses were either blended or fully on line; b) 26-49% blended or fully on line; c) 1-25% blended or fully on line; or d) fully in person. This requires the HEI's data systems to have clear definitions to record each course in these delivery mode categories, as proposed above. Indicators C3 and C4, which relate to time to completion and to attrition (both of which are programme measures), use this categorisation.
6. Three indicators (C5, C6, C8) relate to labour market outcomes. They will require analysis of microdata held in the DPR database. Given that labour market outcomes are dependent on programme characteristics (especially level and field of study) and, possibly on student characteristics (e.g. gender and regional or socio-economic grouping), data on C5 (employment rates), C6 (earnings premium for graduates in employment) and C8 (graduates reporting trust in the credibility of their credential) also need to be reported by student characteristics, and by study characteristics (level and field of study), as well as by the categorisation of enrolment (as in Point 5 above) according to the extent the student has taken online courses as part of his/her programme.

Table 4.6. Potential indicators to measure the digital readiness of Hungarian higher education

Sub-domains	Indicators	Potential data collection approaches				Comment
		Admin data	Survey data	Learning analytics	Other	
Access to digital tools and content	A1. Percentage of students and staff that have access to reliable Wi-Fi whenever it is needed: a) at the HEI b) outside the HEI		X			INDEx Q13. Administrative data can supply information about the availability of Wi-Fi but not about reliability.
	A2. Percentage of students and staff that have access to hardware and software essential for online teaching and learning and research: a) provided by the HEI b) personally owned or provided by a third party		X			
	A3. Percentage of students and staff that have access whenever needed to digital learning and teaching content (e.g. recorded lectures, online course materials, e-journals and e-books)		X			This indicator summarises content from INDEx Q13. There may be an argument for breaking this indicator down by type of content, as in INDEx.
	A4. Percentage of courses available fully on line and in a blended format (using categories outlined in indicator C1)	X				While this comes from administrative data in some OECD countries, it may need to be collected via a survey of HEIs, unless the NEPTUN and FIR systems are modified to collect it. If not, for aggregation, the actual numbers will need to be supplied, not simply the percentages.
Financial and human resources dedicated to digital infrastructure and systems	A5. Institutional capital expenditure on digital infrastructure as a percentage of total capital expenditure	X	X		X	While these indicators ask only for ratios/percentages, if the government wants to get an aggregated view, it will be necessary for HEIs to supply to the government the actual numbers (e.g. the value of all capital expenditure and the value of digital infrastructure capital expenditure), not just percentages, because percentages/ratios cannot be added.
	A6. Ratio of full-time equivalent (FTE) professional staff supporting users of information and communication technology (ICT) systems to total FTE staff	X				Estimates of time spent on ICT support may be complex for HEIs to do.

Table 4.7. Potential indicators to measure the digital practices of Hungarian higher education

Sub-domains	Indicators	Potential data collection approaches				Comment
		Admin data	Survey data	Learning analytics	Other	
Use of digital tools by students and teachers	B1. Percentage of students/staff reporting using digital technologies for teaching/learning, by type of technology (computer, mobile, etc.)		X	X		INDEx Q11
	B2. Percentage of students accessing LMS/VLE weekly or more frequently, in their own learning time, by purpose of accessing the LMS/VLE		X	X		INDEx Q12, Q18
	B3. Percentage of students/staff regularly accessing non-educational technologies to support teaching and learning (e.g. videoconferencing software, social media), by type of technology		X	X		
	B4. Percentage of students/staff reporting regularly accessing digital educational resources, in their own learning time, by type of resource		X	X		INDEx Q17
	B5. Percentage of students/staff reporting regular use of academic and other supports to facilitate online teaching and learning, research and engagement, by type of support		X	X		
	B6. Percentage of staff participating in an international professional development opportunity on line (e.g. sabbatical, etc.)	X	X			
Institutional digital environment and culture	B7. Percentage of students who agreed their institution protected their data privacy		X			INDEx Q14.5, Q19.4
	B8. Percentage of staff who agreed they were informed about their responsibilities on how to manage learner data securely		X			
	B9. Percentage of students/staff who agreed they had the opportunity to be involved in decisions about digital services		X			INDEx Q20.5
	B10. Percentage of students reporting satisfaction with the learning experience in online, blended and in-person formats		X			INDEx Q21, Q20.1-20.4, Q24
	B11. Percentage of staff reporting job satisfaction, in online, blended and in-person formats		X			
Digital skills	B12. Self-reported level of digital skills by students and staff		X			
	B13. Students' assessment of the digital skills of staff / staff's assessment of students' digital skills		X			
	B14. Percentage of students/staff reporting regular use of opportunities to review and update digital skills		X			INDEx Q20.2

Table 4.8. Potential indicators to measure the digital performance of Hungarian higher education

Sub-domains	Indicators	Potential data collection approaches				Comment
		Admin data	Survey data	Learning analytics	Other	
Access and equity	C1. Percentage of students enrolled in courses that involve a share of study time on line, by student characteristics and course characteristics (level/field of study). Course categories could be categorised as: a) 50%+ of study time on line b) 26-49% of study time on line c) 1-25% of study time on line d) Fully in person	X				This would require each course's status to be recoded in NEPTUN.
	C2. Rates of course completion according to course categories (see Indicator C1), by student characteristics and course characteristics (level/field of study)	X		X		
	C3. Ratio of time-to-completion/minimum time-to-completion, by student enrolment type a) 50%+ blended or fully on line b) 26-49% blended or fully on line c) 1-25% blended or fully on line d) Fully in person Present results by student characteristics and by programme characteristics (e.g. level/field of study)	X		X		This is a programme-level question. Because being in a fully online, blended or in-person course is a characteristic of the student's enrolment in most cases (rather than of the programme), answering a programme-level question involves categorising students/graduates by the extent to which their enrolment over the whole programme is on line, blended or in person. To create this variable, the HEI's student information system (SIS) will need to have recorded the extent to which each course is on line (see indicator C1).
	C4. Rates of attrition after one year by enrolment type (see Indicator C3), student characteristics and programme characteristics	X		X		
Quality	C5. Employment rates of graduates by enrolment type (see Indicator C3), student characteristics and programme characteristics	X	X			Employment outcomes data requires data mining of the DPR graduate tracking system and applying an Enrolment Type variable (see Indicator C3). Graduates need to be categorised by their enrolment type and the results shown by student characteristics (e.g. demographic characteristics, regional characteristics, etc.) as well as by programme characteristics (especially level and field of study).
	C6. Earnings premium of graduates (compared to upper secondary graduates) by enrolment type (see Indicator C3), student characteristics and programme characteristics	X				

Sub-domains	Indicators	Potential data collection approaches				Comment
		Admin data	Survey data	Learning analytics	Other	
	C7. Percentage of students reporting satisfaction with the knowledge and skills they obtained through online, blended or in-person instruction		X			INDEx Q19, 21
Quality	C8, Percentage of graduates reporting trust in the credibility of their credential obtained by enrolment type (see Indicator C3) student characteristics and programme characteristics	X	X			Uses DPR survey data
	C9. Percentage of academic staff reporting trust with assessment, in online, blended and in-person formats		X			
	C10. Percentage of employers reporting satisfaction with the quality of graduates from online, blended or in-person programmes		X			

References

Adult Education Reporting System (2021), *Felnőttképzési Adatszolgáltatási Rendszer [Adult Education Reporting System]*, https://tudasbazis.ekreta.hu/pages/viewpage.action?pageId=46760376 (accessed on 1 July 2021). [37]

Australian Government (2021), *Higher Education Data Collection: Element Dictionary*, https://heimshelp.dese.gov.au/ (accessed on 1 September 2021). [5]

Australian Government (2020), *Upholding Quality - Quality Indicators for Learning and Teaching*, https://www.dese.gov.au/higher-education-statistics/upholding-quality-quality-indicators-learning-and-teaching (accessed on 1 August 2021). [16]

Beetham, H., T. Newman and S. Knight (2019), *Digital Experience Insights Survey 2018: Findings from Australian and New Zealand University Students*, Jisc, Bristol, https://www.jisc.ac.uk/reports/digital-experience-insights-survey-2018-students-anz (accessed on 1 September 2021). [19]

Danish Ministry of Higher Education and Science (2020), *Information about the survey*, https://ufm.dk/en/education/OLDfocus-areas/laeringsbarometer/information-about-the-survey (accessed on 1 August 2021). [14]

DSN/DHECC (2021), *An Analysis of Current Higher Education Data Collected in Hungary and the Value of This Data to Assess Digital Readiness and Digital Practices*, Digital Success Nonprofit Ltd. (DSN)/Digital Higher Education Competence Centre (DHECC), Budapest, document provided to OECD for the project "Supporting the Digital Transformation of Higher Education in Hungary". [36]

DSN/DHECC (2020), *Position Paper on Digitalisation of Hungarian Higher Education*, Digital Success Nonprofit Ltd. (DSN)/Digital Higher Education Competence Centre (DHECC), Budapest, document provided to OECD for the project "Supporting the Digital Transformation of Higher Education in Hungary". [8]

Educational Authority (2020), *Graduate Career Tracking System*, https://www.diplomantul.hu/ (accessed on 1 July 2021). [15]

Educational Authority (2018), *Felsőoktatási Információs Rendszer (FIR) [Higher Education Information System (FIR)]*, https://www.oktatas.hu/felsooktatas/fir/fir_mukodes_alkalmazas (accessed on 1 July 2021). [35]

European Commission (2021), *Data protection in the EU*, https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en (accessed on 26 July 2021). [32]

Georgia State University (2021), *Student Success Programs*, https://success.gsu.edu/ (accessed on 1 May 2021). [25]

Georgia State University (2018), *2018 Status Report - Complete College Georgia*, Georgia State University, Atlanta, https://success.gsu.edu/download/2018-status-report-georgia-state-university-complete-college-georgia/ (accessed on 30 August 2021). [24]

Guiney, P. (2016), *E-learning Provision, Participation and Performance*, New Zealand Ministry of Education, Wellington, https://www.educationcounts.govt.nz/publications/tertiary_education/e-learning/e-learning-provision,-participation-and-performance (accessed on 1 September 2021). [2]

HESA (2021), *HESA Collections*, Higher Education Statistics Agency UK, https://www.hesa.ac.uk/collection/c20051/index (accessed on 1 September 2021). [7]

HÖOK (2020), *Távoktatási jelentés [E-learning report]*, National Union of Students' in Hungary (HÖOK), Budapest, https://hook.hu/hu/felsooktatas/tavoktatas-jelentes-2851 (accessed on 30 August 2021). [38]

Ifenthaler, D. (2021), *Student-centred Perspective in the Digitalisation of Higher Education*, paper prepared for the European Commission-Hungary-OECD project "Supporting the Digital Transformation of Hungarian Higher Education". [3]

Ifenthaler, D. (2015), "Learning Analytics", in Spector, J. (ed.), *The SAGE Encyclopedia of Educational Technology*, SAGE Publications, Thousand Oaks, http://www.doi.org/10.4135/9781483346397.n187. [23]

Ifenthaler, D. (2012), "Learning Management System", in Seel, N. (ed.), *Encyclopedia of the Sciences of Learning*, Springer, Boston, https://doi.org/10.1007/978-1-4419-1428-6_187. [4]

Ifenthaler, D. and C. Widanapathirana (2014), "Development and Validation of a Learning Analytics Framework: Two Case Studies Using Support Vector Machines", *Technology, Knowledge and Learning*, Vol. 19, pp. 221–240, https://doi.org/10.1007/s10758-014-9226-4. [27]

Ifenthaler, D. and J. Yau (2020), "Utilising Learning Analytics to Support Study Success in Higher Education: A Systematic Review", *Educational Technology Research and Development*, Vol. 68, pp. 1961–1990, https://doi.org/10.1007/s11423-020-09788-z. [29]

Jisc (2020), *Student Digital Experience Insights Survey 2020: Question by Question Analysis of Findings from Students in UK Further and Higher Education*, Jisc, Bristol, https://www.jisc.ac.uk/sites/default/files/dei-2020-student-survey-question-by-question-analysis.pdf (accessed on 1 September 2021). [20]

Jones, K. (2019), "Learning analytics and higher education: a proposed model for establishing informed consent mechanisms to promote student privacy and autonomy", *International Journal of Educational Technology in Higher Education*, Vol. 16/24, https://doi.org/10.1186/s41239-019-0155-0. [33]

Miller, E. and J. Shedd (2019), "The History and Evolution of IPEDS", *New Directions for Institutional Research*, Vol. 2019/181, pp. 47-58, https://doi.org/10.1002/ir.20297. [10]

National Forum (2020), *Irish National Digital Experience (INDEx) Survey: Findings from students and staff who teach in higher education*, National Forum for the Enhancement of Teaching and Learning in Higher Education (National Forum), Dublin, https://hub.teachingandlearning.ie/resource/irish-national-digital-experience-index-survey-findings-from-students-and-staff-who-teach-in-higher-education/ (accessed on 1 September 2021). [18]

NCES (2021), *2020-2021 Data Collection System*, https://surveys.nces.ed.gov/IPEDS/ (accessed on 1 April 2021). [12]

NCES (2021), *Distance Education in IPEDS*, https://nces.ed.gov/ipeds/use-the-data/distance-education-in-ipeds (accessed on 1 April 2021). [13]

NCES (2021), *Integrated Postsecondary Education Data System: Overview of IPEDS Data*, https://nces.ed.gov/ipeds/use-the-data/overview-of-ipeds-data (accessed on 1 September 2021). [9]

New Zealand Ministry of Education (2021), *National Student Index (NSI) Web Application*, https://applications.education.govt.nz/national-student-index-nsi-web-application (accessed on 1 September 2021). [6]

Ochoa, X., S. Knight and A. Wise (2020), "Learning analytics impact: Critical conversations on relevance and social responsibility", *Journal of Learning Analytics*, Vol. 7/3, pp. 1-5, https://doi.org/10.18608/JLA.2020.73.1. [34]

OECD (2019), *Benchmarking Higher Education System Performance*, Higher Education, OECD Publishing, Paris, https://dx.doi.org/10.1787/be5514d7-en. [21]

Sclater, N., A. Peasgood and J. Mullan (2016), *Learning Analytics in Higher Education: A Review of UK and International Practice*, Jisc, Bristol, https://www.jisc.ac.uk/sites/default/files/learning-analytics-in-he-v2_0.pdf (accessed on 1 September 2021). [26]

SHEEO (2021), *New Data from SHEEO Strong Foundations 2020 Survey Shows Growth of State Postsecondary Data Systems Over 10 Years*, https://sheeo.org/new-data-from-sheeo-strong-foundations-2020-survey-shows-growth-of-state-postsecondary-data-systems-over-10-years/ (accessed on 6 September 2021). [11]

Statistics Canada (2021), *Data gathering and processing: Estimation*, https://www150.statcan.gc.ca/n1/edu/power-pouvoir/ch13/estimation/5214893-eng.htm (accessed on 1 September 2021). [22]

UK Department for Education (2020), *Employer skills survey 2019*, https://www.gov.uk/government/collections/employer-skills-survey-2019 (accessed on 1 August 2021). [17]

Weaver, D., C. Spratt and C. Nair (2008), "Academic and student use of a learning management system: Implications for quality", *Australasian Journal of Educational Technology*, Vol. 24/1, pp. 30-41, https://doi.org/10.14742/ajet.1228. [31]

Wise, A. and Y. Cui (2018), *Envisioning a Learning Analytics for the Learning Sciences*, International Society of the Learning Sciences, https://nyuscholars.nyu.edu/ws/files/39109305/Wise_Cui_LAforLS_ICLS18.pdf (accessed on 1 September 2021). [28]

Wise, A. and Y. Jung (2019), "Teaching with Analytics: Towards a Situated Model of Instructional Decision-Making", *Journal of Learning Analytics*, Vol. 6/2, pp. 53-69, http://dx.doi.org/10.18608/jla.2019.62.4. [30]

Xu, D. and Y. Xu (2019), *The Promises and Limits of Online Higher Education - Understanding How Distance Education Affects Access, Cost and Quality*, American Enterprise Institute (AEI), Washington, DC, https://www.aei.org/wp-content/uploads/2019/03/The-Promises-and-Limits-of-Online-Higher-Education.pdf (accessed on 8 April 2020). [1]