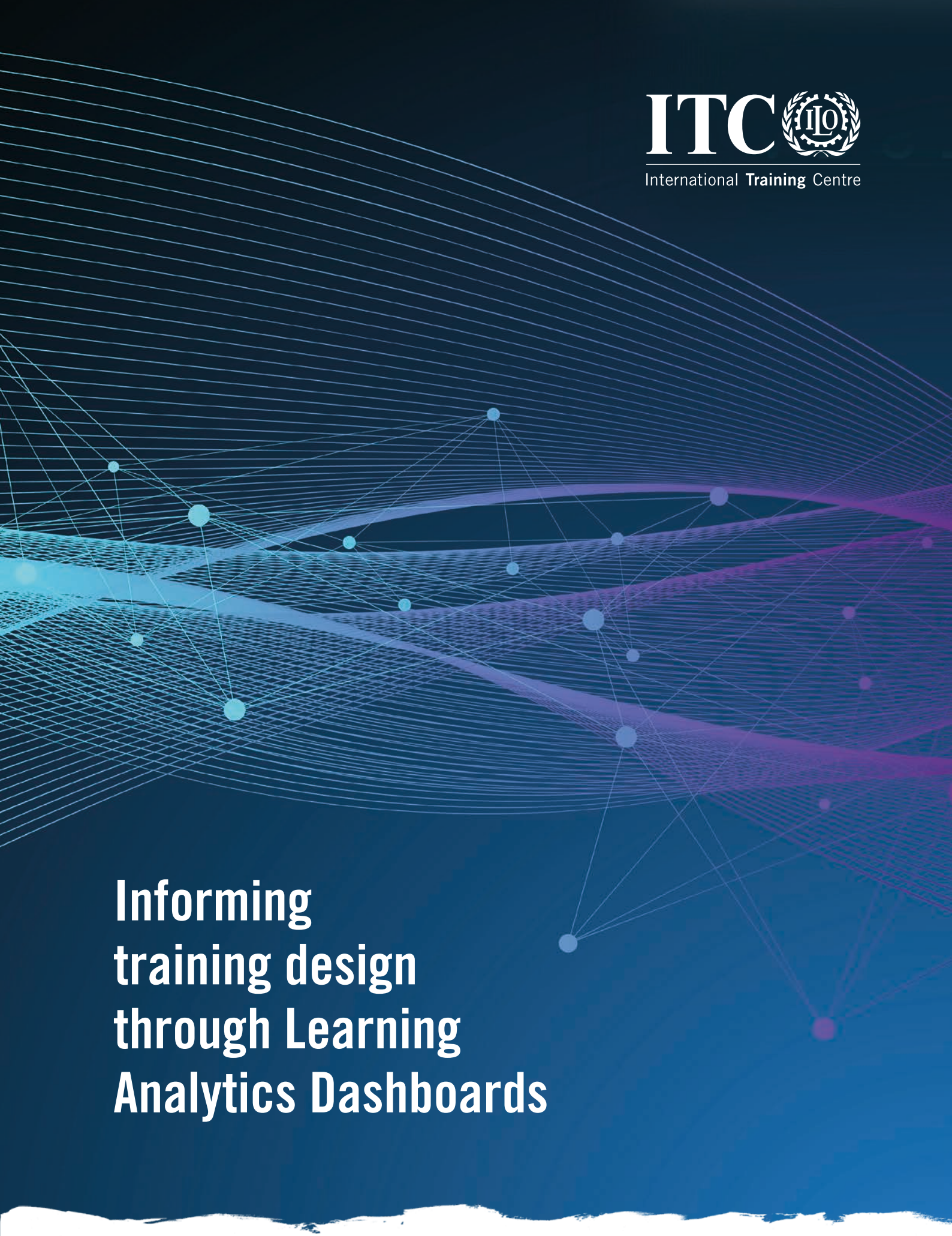


ITC ILO
International Training Centre

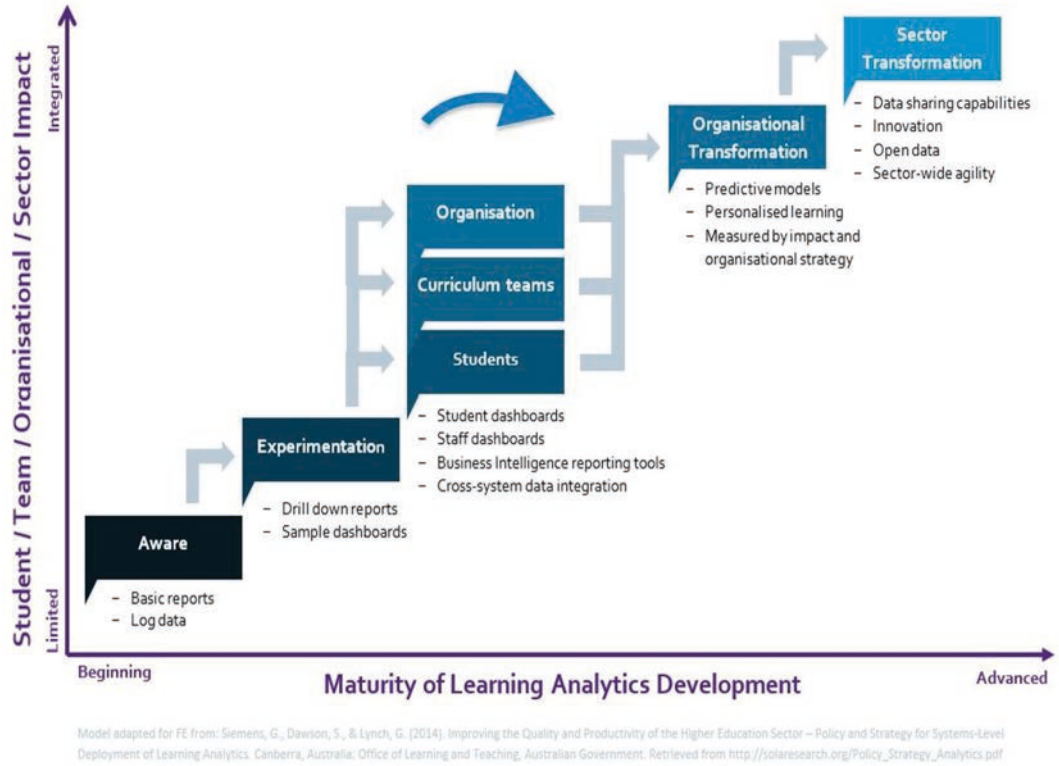
Informing training design through Learning Analytics Dashboards

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The aim of this report is to present a roadmap for reaching Step 3 of the learning analytics sophistication model. Based on Drachsler and Kalz's report from 2020, we analyse the current learning analytics status of the ITCILO in terms of the dashboard design and its implementation in organizational processes

.



Source: Siemens, G., Dawson, S., & Lynch, G. (2014). Improving the Quality and Productivity of the Higher Education Sector – Policy and Strategy for Systems-Level Deployment of Learning Analytics. Canberra, Australia: Office of Learning and Teaching, Australian Government.

Reference dashboard fragments are developed and a tangible roadmap towards implementing the necessary technical and organizational steps is outlined. These provide what we believe to be the best usage of available information according to the interviews conducted, and are meant to be the stepping stone and process outline for a long-lasting improvement process that involves continuous re-invention of learning analytics at the Centre. By viewing this report as the first iteration in that process, we show how continuous improvement can be achieved in moving to and beyond Level 3, and towards being a pioneer in digital transformation as a training organization.

These are specific actionables to be taken, ordered by appearance in this document. They are to be understood in addition to the prior report, designed to detail prior recommendations and give further insights.

#	Actionable	Link
R1	Encourage **A/B tests** to track progress	Stakeholder Analysis
R2	Establish **wording** around Data Dimensions	**Data Analysis**, **Participant Dimensions**
R3	Rethink professional context	Participant Dimensions
R4	Find human resources for learning analytics design (compounding)	Indicator Analysis
R5	Introduce **course similarity** as a measure to lay the groundwork for relative indicators	**Relative Dimensions**, **Course Similarity**
R6	Introduce **data availability** metric on course level (meso and macro) to track and improve tracking capabilities	Data Availability
R7	Implement first version of dashboard with access to high level **compound data** and drill-down options for transversal roles (mindset: don't display data, answer questions)	Prototype
R8	Put PDCA in place officially by means of innovation rituals, specific workflow proposed; **transversal meeting** between all units, director of training and ICTS	PDCA workflow
R9	Establish and use new **learning quality KPIs**	Learning Quality KPIs
R10	Implement better tracking and processing capabilities in some areas of the **infrastructure**	Technical Roadmap
R11	Establish a knowledge base with **Standard Operating Procedure**	**Standardization**, **SOP**
R12	Professional development for staff to speed up adoption of new processes	**Professional Development**
R13	Already be in the mindset for Level 4; some ideas are provided	**Future Possible Objectives**

INTRODUCTION

This report is the result of collaboration between the ITCILO and the authors. The ITCILO is "dedicated to achieving decent work while exploring the frontiers of the future of work". It delivers advisory services and training to workers' organizations, employers' organizations and governments, and has been at the cutting edge of training for many years.

In 2020, a report was commissioned to develop high-level advice on the consolidation of the Centre's learning analytics system in light of the overall digital transformation of the Centre. It developed two major visions and five challenges, including a short-, mid- and long-term roadmap to provide the Office of the Director of Training with the ability to systematically analyse and interpret learning data. The authors, Drachsler and Kalz, used the learning analytics sophistication model to evaluate the current state of LA at the Centre and assessed it as being situated between Step 2 and Step 3.

This report stands in a long line of dedicated examinations of the learning process in terms of digitally collected data and objective measures to capture quality and progress. In 2015, the Learning Innovation Programme produced a report on Quality Management in Learning Design, which accelerated the discourse on this topic. Drawing on the financial and governance data regularly collected and structured by the Quality Assurance Unit of the Office of the Director of Training and ICTS, many high-level insights are already possible through the use of advanced business intelligence dashboards.

Due to the pandemic situation, 2020 and 2021 have seen great change in the Centre's distribution of face-to-face vs distance learning. Many steps have already been taken to adjust focus towards online learning. Notably, pre- and post-knowledge acquisition tests have been used for most of the 700 courses offered to measure their effectiveness, in addition to the long-established satisfaction surveys. In this regard, the eCampus-LMS has been reworked in 2021 to lay the foundations for further standardization and centralization of the organically grown IT landscape.

The present cooperation between the authors and the Office of the Director of Training of the ITCILO is designed to follow up on the prior report and reify the specifics of building the necessary user interfaces so that all stakeholders can benefit from the said data in an effort to move the ITCILO to Step 3 of the sophistication model. It draws on the current move towards unified IT systems and suggests a roadmap for implementing the necessary learning analytics dashboards and ongoing developments thereof, as well as their integration into existing processes. It also makes suggestions for a shared vocabulary when referencing LA terms.

The structure of this document is as follows:

Takeaways: The prior report is analysed with particular emphasis on the relevant parts concerning the building of dashboards.

Methodology gives a quick overview of the timeline for the creation of this report and a list of the interviews conducted.

Current infrastructure examines existing IT systems and processes in light of building and integrating learning analytics dashboards.

Modelling: in this critical section, we work through the process of modelling existing learning processes and collecting data so that the result satisfies every objective set by any stakeholder who needs to be considered. The 2020 report is a structural framework for this analysis.

Prototype: we will illustrate several options for a prototype dashboard that can be applied at different places in the existing structure. This is a natural consequence of the modelling process as explained above.

Way forward: lastly, we will propose a roadmap for how to proceed with technical implementation, taking into account current processes within the organization.

LIST OF TABLES AND ABBREVIATIONS

Abbreviation	Meaning
LA	Learning analytics
LAD	Learning analytics dashboard
TinCan / xAPI	De facto standard learning experience API for exchanging learning data with the LMS
LMS	Learning Management System, in this case Moodle, which is used by eCampus
PDCA	Plan-Do-Check-Act, described in detail later
SSO	Single Sign On, which allows unified access to all systems

CONTENT

Takeaways from the earlier learning analytics consultancy service

This section looks back to the report by Drachsler and Kalz and attempts a short summary of the key takeaways, focusing on relevant information concerning the creation of LADs that has been helpful in carrying out this assignment. The results are closely integrated into the structure of this report. In this document, we attempt to give an understanding of the next steps in tackling every challenge and problem posed in the 2020 report. Definitions and background information on LA terms will not be repeated here.

Four chapters of the earlier report are relevant to this assignment. First, the authors provide an overview of the current state of LA as a research field and within the ITCILO, including descriptions of the six dimensions of LA, which will be picked up in the analyses in this report. The chapter on best practices for infrastructure gives an overview of common challenges and issues to be taken into account at the design stage; these will be referenced in detail in the following chapters. Drachsler and Kalz's overview of the currently used information systems will not be repeated in this report, but rather slightly updated and interpreted from a purely architectural point of view.

Most importantly, the authors describe two possible visions and five challenges in building a solid LA ecosystem. All challenges they describe will be picked up on in this document, with a reference to the respective challenge identifier (e.g. "R1.1").

Current state of LA

The Centre is described as situated between Step 2 and Step 3 of the learning analytics sophistication model. This implies that, while several dashboards are offered to different stakeholders in the organization, there is little focus on actual learning data. Currently, the dashboards mainly serve the purpose of identifying existing KPIs based around financial data and outreach statistics. For example, there is no dropout data on the current macro dashboard.

Moving to Step 3 as an organization, includes providing all stakeholders with a dashboard that gives insight into a fully connected set of data at the micro, meso and macro levels, and which is current and relevant to the everyday challenges faced by the stakeholder. Step 3 does not involve organizational transformation towards predictive models and personalized learning. The difference between Step 3 and Step 4 manifests in the proactive component of dashboards and in actively supporting the educational process. The goal of this report should therefore be to illustrate dashboards that provide conclusive insight into the learning and training processes as measured by a fully integrated system. It should provide data and analysis that lead to incremental improvements in training, though not yet wired into these improvement cycles. In the remainder

of the report, there will be references to proactive measures to be taken in anticipation of what is to come after Step 3 has been achieved.

In the analysis and roadmap section of this report, the six dimensions of Learning Analytics are considered and dissected in detail. Please refer to the 2020 report for a detailed description of the dimensions and their relevance to Learning Analytics dashboards (LADs).

- Stakeholder
- Objectives
- Data
- Instruments
- external constraints
- internal limitations

Best practices for LA infrastructure

In their chapter on best practices for LA infrastructure, the authors touch on a number of key factors when designing technical architecture for LA systems and give real-world examples. The questions raised in the report include the trade-off between the advantages of data lake architecture vs the relative simplicity of setting up a data warehouse, especially given the current ecosystem at ITCILO. There are given examples of pieces of technology, which each support possible architectural choices. Three universities are cited as running lighthouse projects in the sector. These include the dropout-focused dashboard at the Universitat Oberta de Catalunya, the at-risk analysis dashboard of the British Open University, and the move from guided self-study towards an activated learning model at the Open University of the Netherlands.

The authors also detail several mental models as food for thought when designing dashboards. LADs should be based on a clear educational concept, rather than showing all available data; use well-defined reference frameworks (social, achievement and / or progress) and focus on constructivist learning. Data sources should not only include activity data but also take into account the content of completed assignments. More specifically, several groups of indicators, relating to learners, actions, content, results, context and social factors, require different kinds of data sources; these will later be analyzed in depth. The objectives of a micro dashboard might include metacognitive, as well as the cognitive, behavioral and emotional, progress of the learner. Meso dashboards typically focus on courses and students at a high level of risk and allow drill-down to find out the reason and take appropriate action. Last but not least, the factors differentiating micro, meso and macro dashboards are discussed.

Information Systems used

The point is made that the ITCILO currently operates in four fields: strategy, business development, clients and customers. Importantly, learning data is not one of them. The authors give an overview of existing technical systems. They mention the lack of SSO support, which increases the difficulty in collecting user-related data on third-party tools, as used, for example, in eCampus courses.

Visions and Challenges

Rather than repeating the challenges here, they are summarized. Please refer to the tables at the end of every subsection in Chapter 5 of the prior report for a complete list.

The authors identified a need to shift the Centre's **paradigm** towards a data-driven course-provider platform, possibly focusing more on lifelong learning and continuous professional development. Repeat participants should be taken into consideration and more complex indicators based on learning data should be used as KPIs. The so-called Product Owners should be enabled to iterate on the existing dashboards and modify them over time, adjusting KPIs, reports and data collection to incorporate them to the insights gained along the way.

In terms of **infrastructure**, new technology was recommended for more granular tracking of user activity, for example using existing logging tools available on Moodle. Data collection should be applied to all customer services, including learning activities realized using third-party software that is not currently integrated with data collection. A data-lake solution might be considered as a way of improving the infrastructure.

Explicit examples were given as to which **data and indicators** should be worked on. These included a recommendation to increase the focus on collecting knowledge-acquisition data using the KAT concept, which was already in place. Additionally, existing indicators that map collected data should be reframed and pulled together into more meaningful and applicable compound indicators. On the meso level, the collection of satisfaction and engagement data was not yet made available to dashboard users.

The PDCA cycle – already in application – should be used to support future **evidence-centered learning design**. To achieve this goal, the recommendation was to use the data and insights gathered as an integral part of learning design for new courses or the overhaul of existing ones.

In terms of the **user-centricity** of the LA tools provided, future dashboards such as the ones designed in this report should emphasize prioritization and avoid over-complication. A clean LAD needs to take into consideration the user experience, which benefits from clear communication and the reduction of visual messages to a minimum in getting the necessary points across. An important aspect of a well-functioning dashboard is uptake and adoption by the users. Training opportunities to achieve higher data literacy should therefore be provided, and a Code of Practice adopted to raise awareness of how to usefully apply dashboards to everyday work.

Methodology

As a first step, the current situation was analysed and the terms of reference put into perspective in the wider frame of the current state of Learning Analytics. Introductory meetings were held with the authors of the prior report to exchange information and to pick up the baton where the prior report has left off.

Secondly, five interviews were conducted in order to understand the different perspectives of executives and thought leaders within the organization. These interviews served to get an overview of the current situation in terms of data movement, analysis and distribution. A summary of these introductory interviews is set out in the table below.

Name	Topic	Date
Eiman Elmasry	Introduction to the situation, how data is captured and evaluated, organizational structure	8.11.2021
Andreas Klemmer	Communication of high-level goals for this assignment and the state of data from the executive perspective	12.11.2021
Gael Lams	Overview of existing systems and data flows, technical perspective	16.11.2021
Tom Wambeke	Forward-looking discussion of the state of tools and innovation, first look into eCampus	17.11.2021
Alessia Messuti	Deep dive into eCampus and existing basic dashboards, real-life application and challenges	19.11.2021

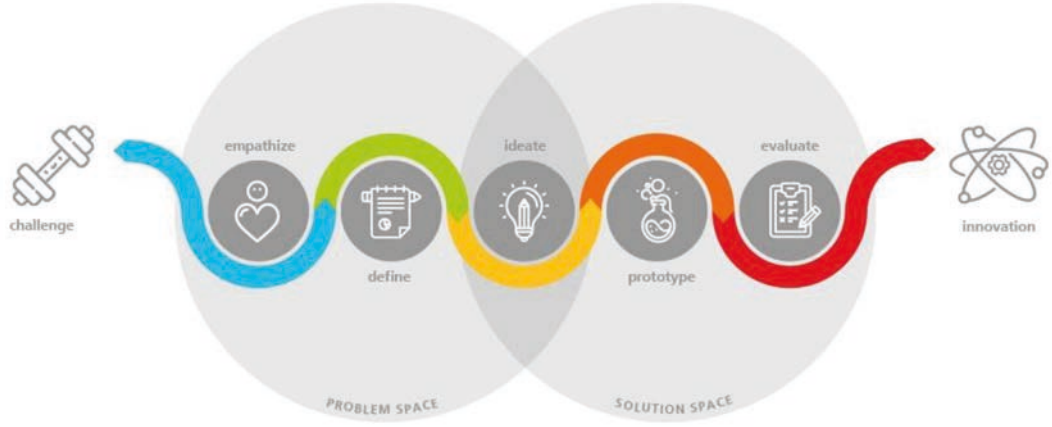
In a second round of interviews, we aggregated an overview of the most common didactic approaches within the Centre and how data processing is applied to these. The main purpose of these interviews was to figure out which dimensions and perspectives we need to consider in designing the dashboards. A plethora of information was also gathered on what is already planned and how the next steps can be integrated in order to avoid unnecessary duplication of work.

Onur Karademir	Discussion of the current state of research in the field, cutting-edge options that make sense to apply here	2.12.2021
Lars Dawidowsky	IT expert at Adornis, about the current architecture and possible improvements	6.12.2021
Eiman Elmasry	More information about the data collected and the structure of the report	10.12.2021
Hendrik Drachsler	Discussion of the integration of prior findings, feedback from research perspective	21.12.2021
Dino Ferrero	Technical details of data-warehouse architecture and manual breaks in the existing process	22.12.2021
Charles Crevier	Perspective of a programme and activities manager, needing to gain an overview of dozens of courses at once	05.01.2022
Guillaume Mercier	Perspective of a programme and activities manager: definition of learning innovation, qualitative vs quantitative data	07.01.2022
Tom Wambeke	Possible implementation strategies within the organization	10.01.2022
and several more	In between and spontaneously	

The suggested dashboards are created in an iterative process that borrows from the design-thinking approach. At its core, the design-thinking methodology is a close neighbor to the PDCA cycle, which is also deployed by the Learning Innovation Programme at the ITCILO. Design thinking is more tailored towards building tech solutions and approaches, creating new interfaces and user experiences in five consecutive steps:

- Empathize: Researching and collecting data about the current state of both the Centre and Learning Analytics as a rapidly evolving field of research;
- Define: Tightly define the problem at hand to find a well-defined solution at the end of the process;
- Ideate: Collect possible approaches to creating the dashboards, including possible consequences and side effects;
- Prototype: Create a rough prototype of the most promising ideas, considering the desired outcome but also the given constraints;
- Test: Evaluate the generated prototype with all groups of stakeholders in order to understand its potential and advantages, as well as its limits and challenges, in order to improve in the next iteration.

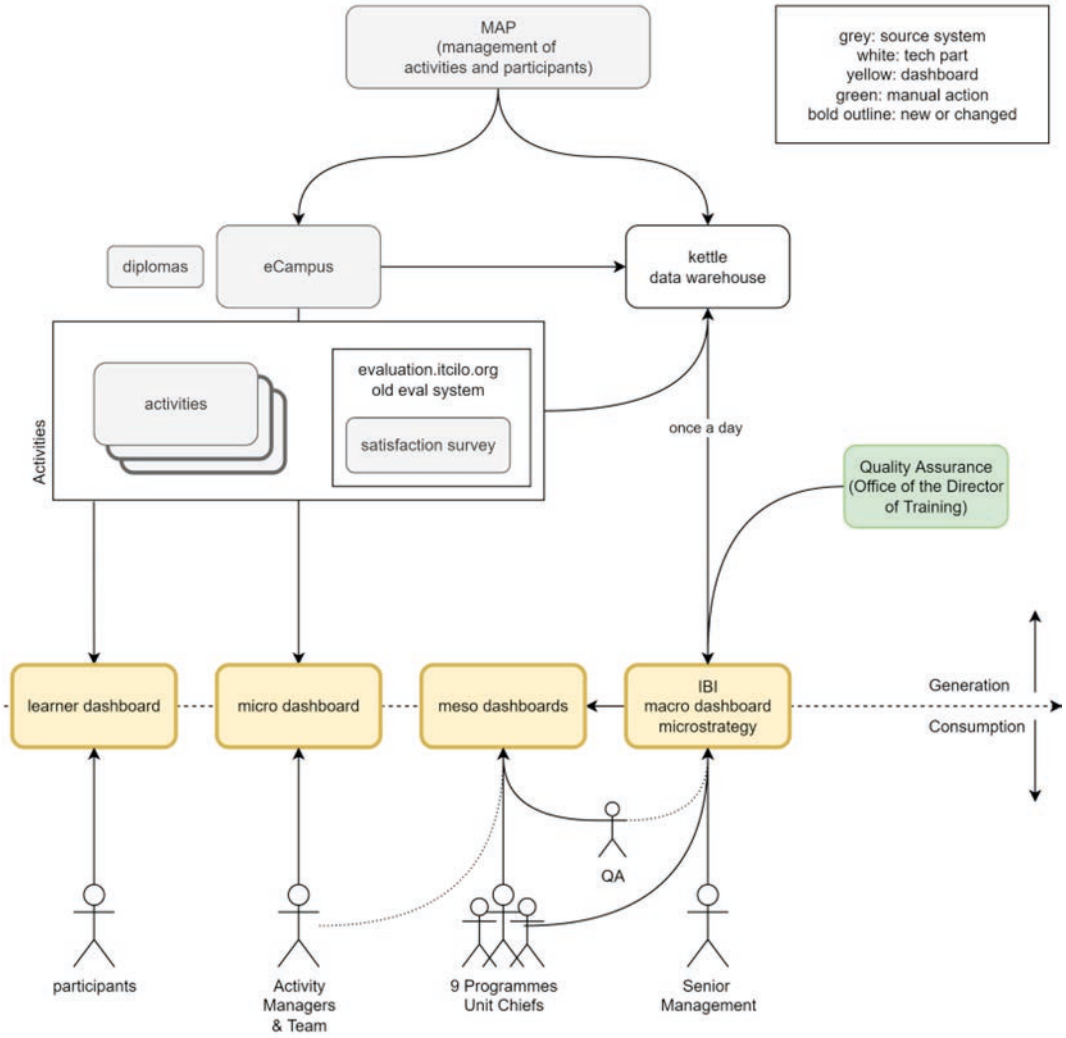
Roughly until the beginning of December, we focused on the "empathize", "define" and "ideate" phases, collecting data, information and ideas from internal resources and the current state of research with the help of DIPF. Afterwards, the dashboards were prototyped, and then tested. The final prototypes can be found below in the respective sections. User testing took place at the end of the year and will represent both the end of the report-writing process and the beginning of the next continuous improvement cycle, to be carried out internally. The results of the first design-thinking iteration are described in the remainder of this report.



Source: https://www.design-thinking-factory.fit.fraunhofer.de/en/design-thinking.html

Conceptualization of ITCILO's learning analytics infrastructure

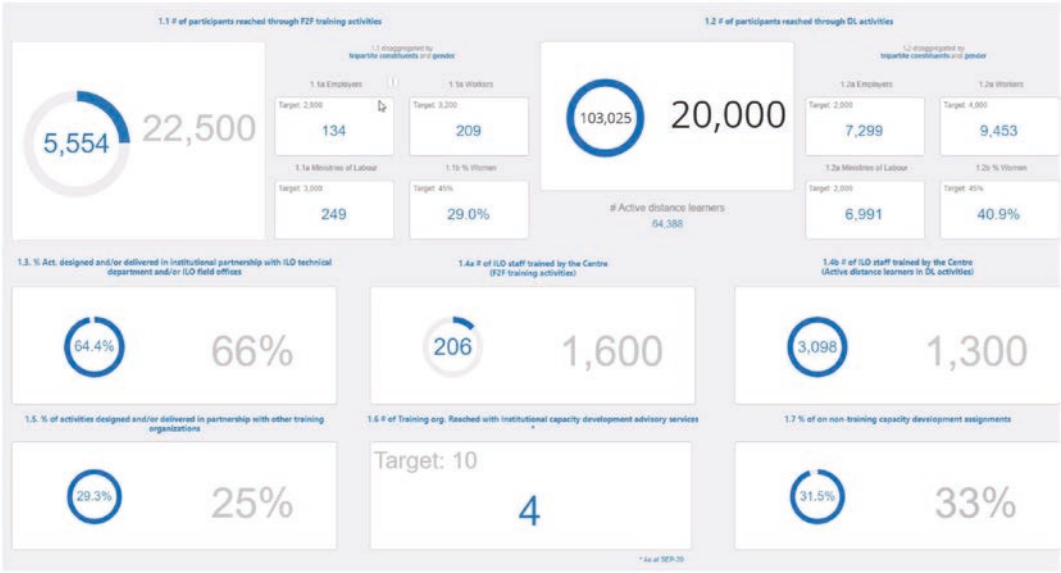
In cooperation with the ICTS, the following diagram has been created to illustrate the technical infrastructure for handling LA data. This is explicitly not meant to represent only the current state of affairs, but also includes modifications and clean-up planned for the near future. For simplicity's sake, SOLICOMM has been omitted from the following graphic. In the mid-term, SOLICOMM and the eCampus should be unified under one roof to avoid blocking the progress of Learning Analytics to the next steps in its evolution. Also, the existing evaluation system has been assumed to be a part of eCampus, which technically speaking is not presently the case.



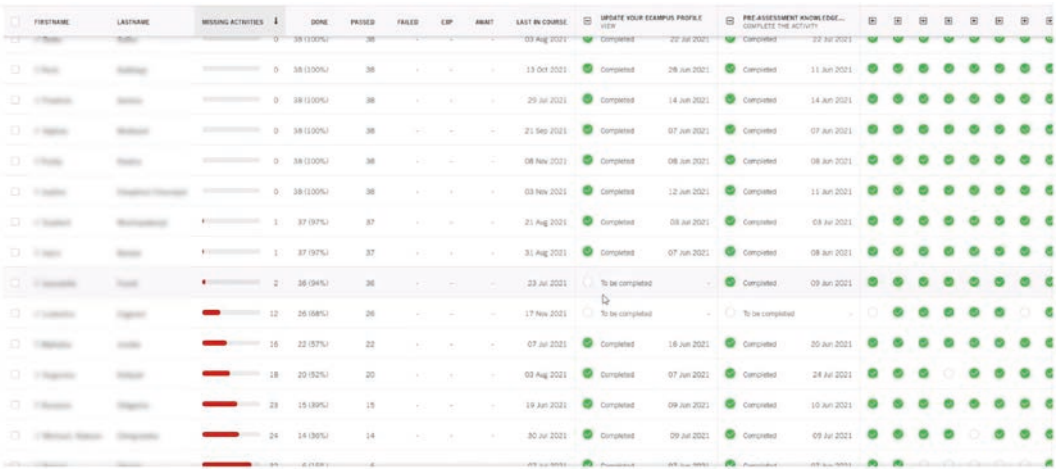
The following section provides an overview of what the current ITCILO infrastructure is capable of in terms of Learning Analytics. The focus is on systems that are relevant for LA data; nonetheless, several IT systems omitted are an essential part of the Centre's operations in other areas.

In terms of dashboards, there are two systems already in place, which offer some form of LA indicators.

1) While not focused exclusively on learning, the IBI as described below already focuses on the macro perspective.



2) On the eCampus, there is a micro-level activity manager dashboard, which gives access to course and activity completion states and grades by user.



In this proposal for a plan on how to build future dashboards, we will attempt to integrate existing solutions as best as possible to avoid duplication and unnecessary effort. In general, the external cloud-architecture experts we interviewed agree that in principle the set-up is capable of reaching Step 3 of the LA sophistication model. In the roadmap section at the end of the report, a possible alternative to the current data-warehouse set-up is discussed as a way of making further progress and allowing for predictive and proactive learning analytics.

Online Learning on eCampus

The eCampus system is based on Moodle but modified to suit the Centre's needs. A major relaunch was conducted in 2021, mainly under the leadership of LIP in cooperation with ICTS, to modernize its functionality and appearance.

Existing Micro-Dashboards

On the eCampus, the ICTS has built a micro-dashboard, mainly to track course completion rates. It gives the activity manager the ability to drill down into the performance of every participant and see which activities have been completed and with which grade (if applicable). This data is also made available to the participants by means of a progress bar on their dashboard and on the course landing page.

Utilized learning-activity types

Due to the Covid-19 pandemic, several activity types were added or used more often in 2021 to accommodate the increase in distance-learning, as compared with the decreasing demand for face-to-face courses. These include webinars (via Zoom and Webex), increased use and evaluation of KATs (knowledge acquisition tests) both before and after the course (so called pre- and post-KATs), and other TinCan-compatible tools like Padlet and Articulate. With the exception of the KATs, user interaction during use of these tools cannot be measured at present. A comprehensive list of the subset of tools that generate usable data can be found in the analysis of available data below.

Data warehouse - Kettle

Kettle is at the heart of data processing. It serves the purpose of a warehouse that collects data from the MAP and eCampus systems, maps and pushes it to the IBI, where it can be displayed. Kettle snapshots the data once every 24 hours (at night) and processes it into its own structured SQL database. An architectural choice has been made not to take the time dimension of the data into consideration, due to the increased complexity and computational power needed.

Business Intelligence - IBI

The IBI is based on micro-strategy. It provides a macro overview of the Centre's activities, historically focused on financial and governance data and KPIs, including their achievement rates. The user can drill down to the precise participant or activity that generated the data and find out exactly how the aggregated data is constructed. There are several additional dashboards that show only high-level KPIs, in accordance with the biennial planning cycle.

Some manual work is done to clean up data that cannot be generated reliably using the current systems. For example, participants enter some data in their eCampus profile, such as employer details, which might not always be accurate and needs to be supplemented and corrected by Quality Management. We will not focus on these processes in the current report since, from what we have gathered in the interviews, they do not affect LA. That being said, it would be beneficial to think about preventing the occurrence of unclean data by means of process design in the software itself. For example, where the free-text field for entering one's employer is concerned, a smart suggestion drop-down could be added that also allows for new values based on the user's email address and what they enter into the field. Participants who pick one of the options would not need manual clean up later on.

Evaluation System

The current evaluation system is due for replacement in the near future. It serves the purpose of providing participants with an interface to input their satisfaction survey responses. It will soon be integrated into the eCampus as a Moodle module. In this way, data will no longer need manual extraction via Excel sheets, but can be collected by Kettle, like all other quiz results.

Online certificates

A new system based on Accredible has been put in place to automatically create digital certificates for participants who have fulfilled the respective requirements. It allows insights into many aspects of digital credentials, such as sharing activity on social media, automating credential creation and linking to the correct person.

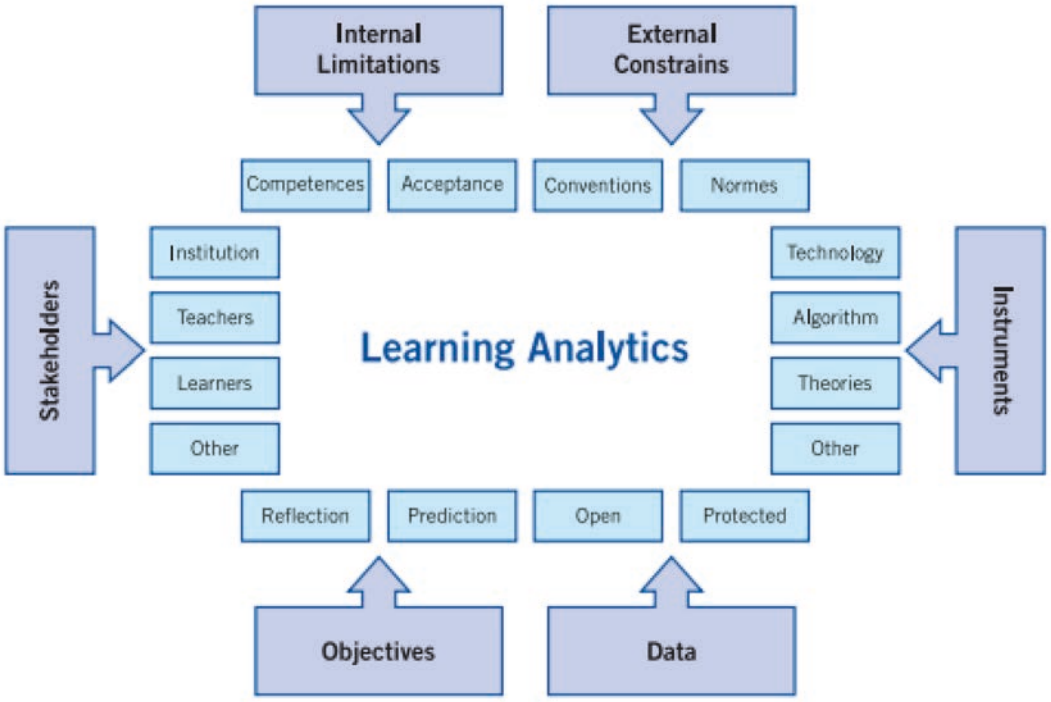
Metadata - MAP

The MAP platform (Management of Activities and Participants) holds all relevant metadata about Centre activities and participants. It is one of the two primary sources of data alongside the eCampus but is not visible to end users. The MAP was originally built to track financial and governance metrics and provide a way to store all necessary entities. Free MOOC courses are therefore not always tracked on this system, though this is the future intention. For the intents and purposes of this report, the MAP data hides behind the data warehousing software, is collected automatically, and can be made available on existing and future dashboards.

The Modeling Process

The section will not be restricted to the data currently collected in the data warehouse, but rather to the metaphysical data that is generated in the process of everyday work, which often is yet to be captured. In the final subsection, we will discuss what additional data needs to be collected to actually represent what is discussed in the data warehouse.

This process will be based on the dimensions of learning analytics proposed by Drachsler and Kalz in the previous report. Internal limitations and external constraints will be discussed in the section "The way forward". Stakeholders, objectives, data and instruments will be discussed in the following subsections of this section on dashboard modeling.



Source: Taken from the 2021 report by Drachsler and Kalz

Stakeholder Analysis

Within the Centre, there are several different positions relevant to this assignment, as well as different perspectives on each of these positions. We will conflate some of the nuanced roles within the organization, given the similarities we expect them to exhibit in their use of the respective dashboards. Because of the different needs, different dashboards need to be created. A one-size-fits-all solution would not only violate several objectives but would also go against the principles of the user-experience design outlined above. In the following paragraphs, we will differentiate between the terms "stakeholder" and "perspective". For the sake of this analysis, "stakeholder" is the term used internally within ITCILO for the job description of any given person. A "perspective", on the other hand, can be shared by different stakeholders but will always involve a similar user

story. Two people with the same perspective ask the same questions and share similar concerns when using the indicators and dashboards.

When considering different stakeholder groups, it is important to note that some are interested in a well-defined information space, while others play a transversal role.

Stakeholders

Participants: everyone who is partaking in any course covered by this analysis.

Coaches / Tutors / Activity Managers and their teams design and lead courses and therefore require a micro perspective on the data. Their role in learning design is to understand how well the learning process is working for different groups of participants and either support them specifically or take their learning experiences into consideration when designing the next course.

Cluster managers / Unit managers / Programme Managers are responsible for several courses or a whole unit. They require a meso level overview of the data and are not concerned with individual participants. Rather, they need to have an overview of courses and be able to measure irregularities and the overall development of learning-process KPIs, such as the rates for drop-out, knowledge acquisition and participant engagement and satisfaction.

The **Learning Innovation Programme** functions as two stakeholders in this analysis. Firstly, they are a programme, in common with other Centre programmes. Secondly, due to the nature of their particular programme, they are especially involved in learning technologies and innovation as part of their mission to mainstream an e-learning culture of high quality.[1] Much like the Quality Assurance team, they therefore deploy a transversal approach which requires access to in-depth data across programme boundaries.

The Director of Training and Senior Management need to oversee the KPIs and track progress on a macro level. Unit chiefs may also need to adopt this perspective from time to time. Most helpful in this scenario is aggregated information on the different indicators displayed at the meso level, though they also need to monitor variables at other levels. Their dashboard should be more KPI-oriented and give a good summary of how the learning process is developing, allowing access to the state of learning in one quick glance. At this level, the statistical composition of the participants is also relevant, for example details of country of origin or number of participants by cluster. These topics are already covered on the current IBI dashboard, so not much time will be devoted to them in this report.

Quality Assurance, as part of the Office of the Director of Training, is a transversal role, requiring access to all data at a detailed level. However, due to the high data literacy of this department, direct access to raw data is often more helpful. The dashboard should then function as a report-generation tool, allowing for further processing into custom reports.

[1] Learning Innovation Action Plan

The **ICTS** in its function as an internal IT service provider plays a central part in training quality. Without its diligent and detailed work, no innovation is possible. Its members generate new tooling, supervise existing systems and support their users. Prepared for constant change, they are able to flexibly adapt to developing requirements. Since they have direct access to the data sources, and the skills to interface with them, not much thought will be devoted to how to accommodate them within dashboards.

Actions

In our perception, the most common activities include the consumption of training, training itself, evaluation of current or past training activities, reporting on past activities, and innovating the training process itself. Below, we attempt an overview of which actions are usually executed by which stakeholders, roughly sorted by data literacy.

	consume	train	evaluate	report	innovate
Participant	x		x		
Activity Manager		x	x		
Cluster / Programme Manager			x		x
Learning Innovation Programme			x		x
Quality Assurance			x	x	x
ICTS				x	x
Training Director			x	x	x

This matrix is essential for understanding which indicators are useful for which roles in the organization. To figure out if an indicator adds value to the everyday process, consider the activity first. Check, which stakeholders are active in that way and therefore interested in the message conveyed by the indicator.

Evaluation is an aspect of all roles. In some form or another, they process information and reflect on the current or past training process. In some roles, this happens passively, with users having information prepared for them by other roles. In the case of stakeholders further down the table, evaluation implies an in-depth analysis of data and will sometimes go beyond the well-prepared data within dashboards.

Reporting is mainly performed within the Office of the Director of Training, directly by using the IBI or indirectly through the Quality Assurance team. Evaluated data is aggregated and prepared for presentations.

Innovating implies thorough research on how to change the process, creative thinking, ideas on where improvements can be made, and testing to figure out correct answers. A/B testing (cohort testing) in particular is a formidable tool for finding profound answers that advance training quality. Innovation also sometimes implies creating or introducing new tools for the entire organization (or small parts of it initially), which naturally need to be aligned with the ICTS. Several roles within the Centre act as innovators within their own domains and therefore require transversal data that goes way beyond user-friendly indicators. Rather, they need in-depth raw data to drive and thoroughly explore new ideas.

As explained in the above stakeholder descriptions, several dashboards are required. At first glance, an obvious choice would be four dashboards, for the micro, meso and macro levels, with the micro level separated into internal micro for activity managers and external micro for end users.

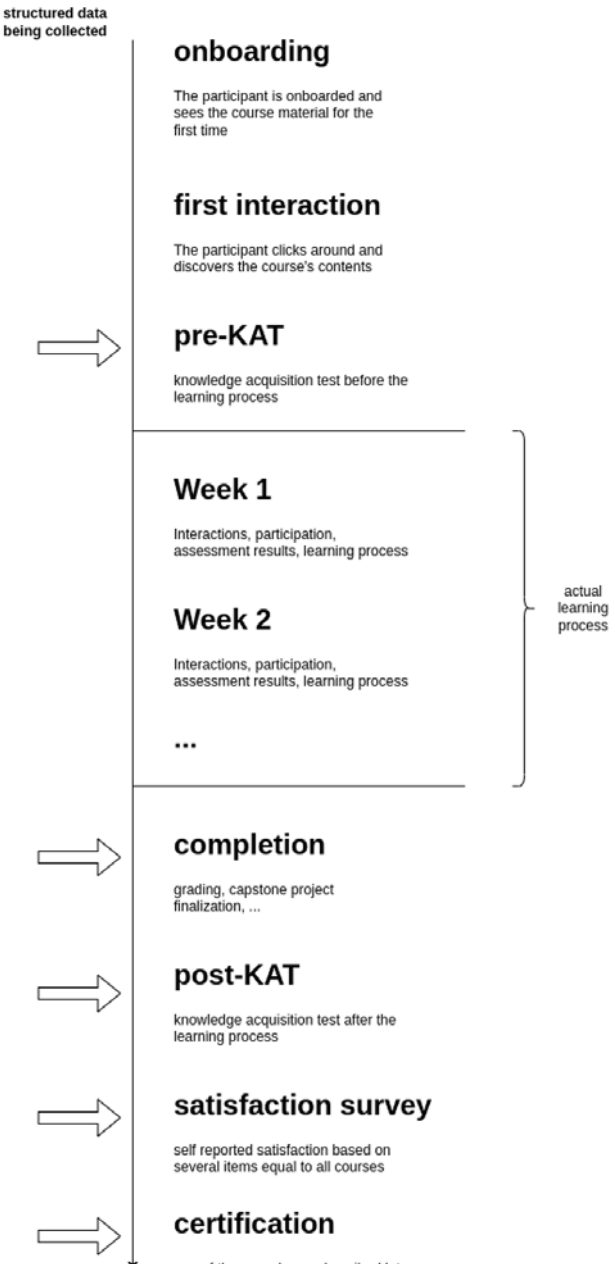
Not including any one of these dashboards would be a significant omission as they each need to show fundamentally different data, as described above, and an essential perspective would be missing. Greater differentiation, on the other hand, would be detrimental to the user journey and would create confusion as to which dashboard to use.

There is an important distinction that is often omitted in published studies concerning dashboards. There are several positions in the Centre that require a transversal view of learning processes, as explained above, namely the Office of the Director of Training including the Quality Assurance team, and the Learning Innovation Programme. The uses envisaged include A/B testing (cohort analysis), report generation, analysis of courses in need of support, measuring the adoption of new tools and so on. Often, this position is taken only by the research group conducting the study on dashboard construction. In our case, there are several people working full time on improving training quality, so it is worth designing the dashboards with this situation in mind. The solution to this challenge is twofold: close communication with the ICTS can mediate some of the need for information by producing information directly from the data, while some of the data might be extractable from existing Excel reports containing raw data. However, this approach inevitably leads to a focus on easy-to-produce metrics and failure to dig for data due to the hassle involved. More time will be devoted to this consideration in the instrument analysis.

Naturally, when reporting numbers for the biennial report, the questions asked and the data presented should be different from the data used when analyzing the situation for the longer-term gain of the educational process. The dashboards, especially on the macro level, should allow for a differentiation between biennial KPI reporting and deep fundamental analysis of educational processes that is not suited for regular reporting but is desperately needed for learning innovation. This is another key differentiating factor that needs to be taken into account where the macro level of the dashboards is concerned. Due to the good coverage of KPI reporting, this report will actively omit these KPIs.

Objective Analysis

From what we have gathered from the interviews and prior work, there are four major objectives to fill in the delta between the status quo and what can be considered as being on Step 3 of the learning analytics sophistication model. This section is not meant to be seen as a full list of all the objectives a stakeholder might have when using the relevant dashboard, but rather a list of objectives not accounted for on the currently offered IBI and eCampus dashboards that have emerged from the interviews.



As we write this report, the Centre's focus is continuously expanding to cover lower and more in-depth levels of measurement and analytics. Historically, data has been collected almost exclusively on the macro level. Now the world of analytics and data is opening up towards didactically inclined stakeholders and internal training staff. This is the perspective taken in this report and therefore the focus is predominantly on the meso level, as discussed with most interviewees. Before it is possible to influence learners' cognition, behavior and emotions, the varying formats of online courses need to be further standardized and streamlined (to which end, ongoing efforts are being made within the Learning Innovation Programme), since it is not feasible to design separate dashboards for dozens of different courses. In conclusion, this report will focus mostly on the meso level, while limiting recommendations on the micro level to well-known indicators usually applied in these scenarios.

Objective I - Measure Early, Measure Often

To quote Andreas Klemmer: "Currently the Centre is looking out of the back window in terms of learning analytics". While collecting data on knowledge acquisition before and after a course is an incredibly efficient way of collecting after-the-fact data about course success, it does not allow intervention while a course is taking place. As prior consultancy has shown, this is largely due to the effort made to collect data in support of existing KPIs, mostly outside of the learning process. Via the dashboard indicators, we suggest a first set of parameters that need to be adjusted in an ongoing process. By proactively defining learning KPIs in addition to existing knowledge-acquisition and application KPIs, the Centre puts itself in the driver's seat and in a position to actively intervene in educational processes.

The previous graphic illustrates the current data-collection points. Ideally, there should be several arrows in each row for collecting real-time learning data. Currently, data is only collected before and after the complete journey. This is largely due to the shift from comparatively short face-to-face courses to distance learning on a very large scale. Given the intended further expansion of distance learning, collecting relevant data early and often needs to be a major focus.

Objective II - Introduction of Meso level and Transversal Functionality

Due to the vast number of activities and participants, it is highly unrealistic to expect anybody overseeing more than a few activities to be aware of all learning activities and their respective data. Currently, therefore, unit chiefs and cluster managers are not able to take advantage of the data collected because it would require them to go into each micro dashboard and figure out the state of each course. At the same time, the macro overview of all courses is both unhelpful to them and misleading because certain correlations are expected to be different for each cluster of courses. New indicators need to be crafted to best represent the Centre's plethora of different didactic modes in an abstract summary on the meso or macro layer. Additionally the different objectives of different stakeholders need to be taken into account when considering which data to show on the respective dashboards. A special requirement to keep in mind is the transversal view necessary to accommodate the everyday work of the Quality Assurance and Learning Innovation teams.

Objective III - Differentiate measurement of effectiveness

Currently, the progress of a participant in a course is tracked exclusively through completion and the grade assigned for a completed activity. While this provides a solid basis for analysis and makes course outcomes comparable on a large scale, it does not necessarily allow for sensible conclusions to be drawn about the participant's engagement or fundamental understanding of the topic at hand. There needs to be a way of measuring participant's involvement that does not rely purely on exposure rates, but takes into consideration whether or not the participant can be considered to have understood whatever learning outcome the lesson was to convey. For example, simply detecting whether a participant has entered a forum post does not convey any information about his or her understanding. On the other hand, analyzing the content of the said forum could potentially measure how involved the participant is in the topic.

Objective IV - Answer questions; don't show data

While it is important to increase the data literacy of all stakeholders using the future dashboards, as pointed out in the 2020 report, the design of the dashboard-user experience can also help to ensure that the data is used in an effective way. When designing dashboards, it is essential to ask the right questions; hence the importance of objective analysis, which should be repeated regularly in the ongoing effort to improve learning design. Only when such awareness is created it is useful to think what story the dashboard needs to tell in order to answer these questions. When dashboards are constructed the other way around, it is inevitable that they will be based on the narrative for which data is available anyway, instead of answering the needs of those who have carefully considered what questions to ask in order to advance the quality of learning.

Consequently, a dashboard should not include any information that has not been asked for in the form of a question that reflects a stakeholder's everyday concerns. Rather, it should focus on the user journey that best fits the requirements defined in the "Plan" stage of the PDCA cycle, and should offer that information directly without the need of interpretation. Put another way: **a good dashboard is not an Excel report, but rather the condensed information that someone has already pulled out for you, which then enables you to draw conclusions on how to act on it**. Using graphic representations rather than text output further increases usability.

In turn, this implies a need to warn the user when the data displayed on a dashboard could wrongly be thought to answer a particular question, when in fact it answers a different question. For example, a user might pose the question of how high the engagement rate is within a specific cluster of courses. Given the different modes of training at the Centre, the dashboard should not display the requested data until the mode of training has also been specified, to avoid confusing engagement in a distance-learning course with engagement in a face-to-face course. Showing the average engagement levels of both distance and face-to-face courses would be not only unhelpful but also misleading.

Further objectives discussed in the interviews are better kept for a later stage. They are expanded upon in the final section of this report.

Data Analysis

At its core, the Centre is a hybrid-learning provider, therefore differences in course design need to be respected and highlighted in all instances. It is important to acknowledge that the Centre's activities are diverse and require different indicators to best capture data. In this section, the various aspects of the Centre's training activities will be examined. Three dimensions have been identified: Modes of Training, Activity Types and Certification Types. Finally, the available student metadata will be listed and described. The data relating to course distribution and participant statistics have been collected in cooperation with the Office of the Director of Training.

Modes of training

The definitive data prepared for regular reports, which is already an aspect of biennial reporting as conducted today, provides statistical insights into the distribution of the available modes of training.

Face-to-face

Historically, face-to-face training was the Centre's mode of choice. Pre-Covid-19, around 6,000 participants came on the Turin campus each year and spent a number of days or weeks there, until the course completion. Naturally, courses that take place on the campus do not generate much data. The same applies to off-campus training, which had an outreach of over 5,000 participants per annum before Covid-19. During the pandemic, the number declined to below 2,000 participants for both on- and off-campus training.

Quoting Gael Lams, Chief ICTS: "The eCampus is used only as a distribution platform for PDF files. Because the mere download of a file does not make it possible to measure learning progress, there first needs to be a change in how reading material is displayed on the eCampus." Forcing the participant to view the PDF in an online viewer where viewing time and engagement can be measured would on the other hand decrease usability (offline reading not recorded, reader probably not as advanced as native-speaker PDF readers, etc.).

Apart from the distribution of PDFs, there is not a lot of data that can be generated without significantly restructuring the way face-to-face activities are conducted. It remains to be seen how numbers develop after the pandemic but, as shown by the numbers for 2018 and 2019, the ratio of distance-learning scenarios was increasing even without the influence of Covid-19. According to the Office of the Director of Training, there are already plans to make an accompanying eCampus course mandatory for all courses .

Online Activities in general were attended by over 50,000 participants in the first year of Covid-19, as opposed to 14,000 before the pandemic. Roughly one third of these participants cannot be considered to be active participants. They should be filtered out whenever possible when analyzing the learning process and engagement or knowledge-acquisition rates.

Tutor-Supported Online Courses

These online courses usually range from 15 to 20 participants, having a "soft" minimum requirement of up to a few dozen, and consist of online activities that are tracked by completion. There are hundreds of available courses of this type (36 per cent of all distance-learning courses), and 28 per cent of all participants followed this type of courses this type in 2020. The E-Learning Design Lab 2021, led by Alessia Messuti, has been taken as an example for the purposes of this analysis. While a similar course might typically have been completed within one or two weeks, due to its asynchronous nature this course took two months. The metrics applied to tutor-supported online courses can be very similar to those applied to MOOCs (see next section), with one major difference: several of the tools used are harder to track, for example webinars and manually created assignments that are rated by the tutor. These require additional thought because the process of partaking in or creating assignments is by nature harder to measure automatically, though on the other hand they often generate some kind of grade.

Webinars were introduced to the eCampus in 2021 and are used in several courses. However, they are not contained in the data-collection cycle, and there is no integration of usage data between Webex, Zoom and eCampus. Fortunately, there are a couple of standard solutions for Zoom, or open source alternatives to Zoom, that allow for granular tracking of an individual participant at any time during the webinar. The technical aspect of this discussion will be expanded upon in the roadmap section. In conclusion, it is essential to look deeper into the activity involved, far beyond whether someone opened a zoom meeting of handed over an assignment or wrote a blog post. Rather, the results of such efforts should be examined and interpreted, which is luckily possible with the current state of technology.

Self-Guided Online Courses - including MOOCs

In only 9 per cent of the digital courses offered as self-guided online courses are 63 per cent of the participants learning digitally. Most of these courses are presumably designed in such a way that one might consider them to be MOOCs. Participants are guided through a curriculum and have to complete several assignments in each segment of the course (usually one a week). Once all assignments have been completed, a certificate of participation is generated. We will further examine the different kinds of learning activities in the next section.

These are the easiest courses to measure activity in because, by definition, every assignment completed by the participant has to be processed on the platform and can therefore be tracked. However, this does not mean that all the necessary data is available. Even in MOOCs, many of the critical steps a participant takes towards achieving the learning goal are still not measurable on the platform, e.g. taking notes on a piece of paper or reflecting on the contents of the course outside of measurable screen time. While this may seem obvious, it is important that it be taken into consideration when drawing conclusions from data. Some people might spend more time reflecting on a video they have watched, while others re-watch it a couple of times. This difference in measured activity is certainly not causal of learning success. Therefore, especially with MOOCs, the dashboards should not attempt to show granular qualitative usage data at any point, but

should rather smooth out these statistical errors by taking into account longer spans of time or many participants at once in an average or distribution function.

Blended Learning

Little emphasis will be placed on this mode of training for this first iteration of the dashboards, given the low number of participants concerned. These courses can be viewed more or less as a mix between tutor-supported online courses and face-to-face courses. There are, however, a few important distinctions to be made, which means that, for the purposes of the dashboard, blended learning has to be treated as a separate class. Because knowledge transfer will usually take place online in these scenarios, and knowledge application in the shared sessions, usage patterns will differ from pure online courses. If the same engagement indicators are applied, there will be less activity measured. One remedy might be to compare activity on the platform with that of other users, rather than using absolute thresholds. This will be expanded upon in the instrument analysis section.

Other activities

Non-training services are by definition left out of this analysis.

Activity types used

Downloadable non-trackable. Where PDFs and links are concerned, measurability is poor but they are much used and very popular. A significant point is that the downloading of optional material might imply high mental involvement in a topic. To increase the impact of these activities, it may be worth considering making answering a question about the contents of a PDF or link mandatory, to help participants reflect on what they have read, assuming that this is not already taken care of in the following activities.

Interactive non-trackable. There are hundreds of tools to engage the learner in what would otherwise be a static PDF. The tools used include Padlet, Student Genially, h5p and many others. Some of them can be integrated via the Tin Can application-programming interface; others can be embedded via an iframe. Fortunately, at least interaction and mouse movement are trackable in both cases. A download button is usually still offered for convenience. This obviously removes the ability to track user interaction, but seems to be a necessary evil given current user expectations.

Videos. Especially in tutor-guided courses, videos are used to convey a message or a piece of knowledge to participants. It is easy to measure how often, for how long and with how many interruptions a participant watches a video. In some courses, there are video lists or longer videos.

Forums. Moodle's own forum solution is often used to share news with participants and allow for discussion of course content. There are several options for tracking data here, ranging from simple activity measurements like counting a participant's replies to applying NLP principles to analyse the depth of a particular contribution. The usefulness of these measurements depends for

the most part on the general usage of the forum. If there is little traffic, the validity of conclusions drawn from the data decreases. In order to avoid misinterpretations, such analysis should be shown only as part of a compound indicator that also takes into account other metrics.

Quizzes / KATs. Historically quizzes have been the easiest and most reliable way to track knowledge acquisition and have been used at the ITCILO. While there are some downsides to using quizzes, particularly when capturing learner understanding of highly complex topics, a well-designed quiz is an effective tool for measuring learner progress in the form of a summative assessment. Integration of quiz response data is easy and already excellently performed by the ICTS.

Webinars. Introduced in 2021, some courses are using webinars as part of the tutor-guided online learning experience. There is great potential for analyzing both participants' engagement in a webinar and the words actually spoken. In the short term, a tracking solution should be embedded in the video-conferencing software. The resulting data is valuable for measuring engagement, especially if the webinar is held in regular periods. In terms of the recording itself, it is even possible to create content-related indicators: epistemic network analysis can be used to detect the topic a given person is talking about. Analysis of participants' vocabulary can be used to compare one participant with another and assign an involvement rating that increases with the use of certain automatically extracted words (by trawling the training material). While this may be a long-term goal, extracting the ratio of seconds during which a user is actively using the conferencing software, as opposed to being "tabbed out" or idling with video shut off, can already provide considerable insight as regards their engagement.

Written Assignments. Much like forum posts, written assignments can be analyzed in terms of length, depth of content and timely submission, relative to other participants taking part in the course. Measurement of learning success will be limited in the case of PDF assignments. While technically possible, analyzing the content would require a lot of effort, which might not been justified when other opportunities for data collection are more promising.

Moodle Profile Completion. As the name implies, completion of a Moodle activity is granted when the Moodle profile contains sufficient content. This is a useful tool for getting participants into the habit of completing assignments in the very first week, when no topics have yet been covered. Since it is not usually used to support the learning process, the Moodle profile completion function will mostly be excluded from the report.

Certification types

Because the type of certification a participant receives is often determined after officially starting a course, it is hard to measure correlations related to certification type. There are three certification types at the ITCILO: certificates of participation, certificates of achievement and diplomas. Diplomas are generally regarded as the highest form of achievement of the three; participants have to achieve results in a number of eligible courses and complete a capstone project. For a certificate of achievement, a sufficiently large course must be completed with some form of

final assessment and a pass grade. Certificates of participation are issued to participants who complete most of any given course.

Lastly, in collaboration with other educational institutions, Masters degrees are offered. For the purpose of these dashboards, they will be considered as having the same implications as diplomas. In both cases, participants need to earn a sufficient number of eligible course points, even though duration and formalities may vary.

Given the different requirements of the certification types, it is safe to assume that participants start their courses with the set goal in mind of obtaining a certain type of certificate. This information therefore gives valuable insight into the motivations of participants, which can in turn be used to raise the quality of training. For example, participants with the goal of obtaining a certificate of participation can be expected to be less committed to completing their course with distinction. There is a need to be clear as to which of the groups any indicator refers to, since the meaning of the information extracted might differ significantly. An appropriate filter will therefore be introduced in the dashboard prototypes.

Participant dimensions

When analyzing participant performance directly or indirectly, several dimensions can and should be taken into account. These dimensions are especially relevant in cohort analysis (A/B Testing) since, by controlling for these variables, much more reliable and conclusive insights can be generated. In the following paragraphs, only tentative hypotheses are stated. For the dashboards, it is important to be able to extract this information in order to further confirm or negate these hypotheses, or generate new ideas.

Region. Studies have found that different cultures have different degrees of learning success using different methods. For example, participant-facing dashboards have been shown to be less impactful in so-called “tight” cultures. This information could be used to pre-emptively structure courses in different ways, based on the composition of participants’ countries of origin.

Currently, regions of origin are tracked by the ITCILO in a way slightly more specific than that of the ILO parent organization: Africa, Americas, Asia, Europe (non-EU), Europe (EU), Oceania, North Africa, Middle East and Western/Central Asia.

Participant in a multi-course track / **Repeat customers**. As described in the certification-type section, the intended goal of certification will most certainly have an impact on learning processes. Controlling this variable is therefore essential to understanding participants better.

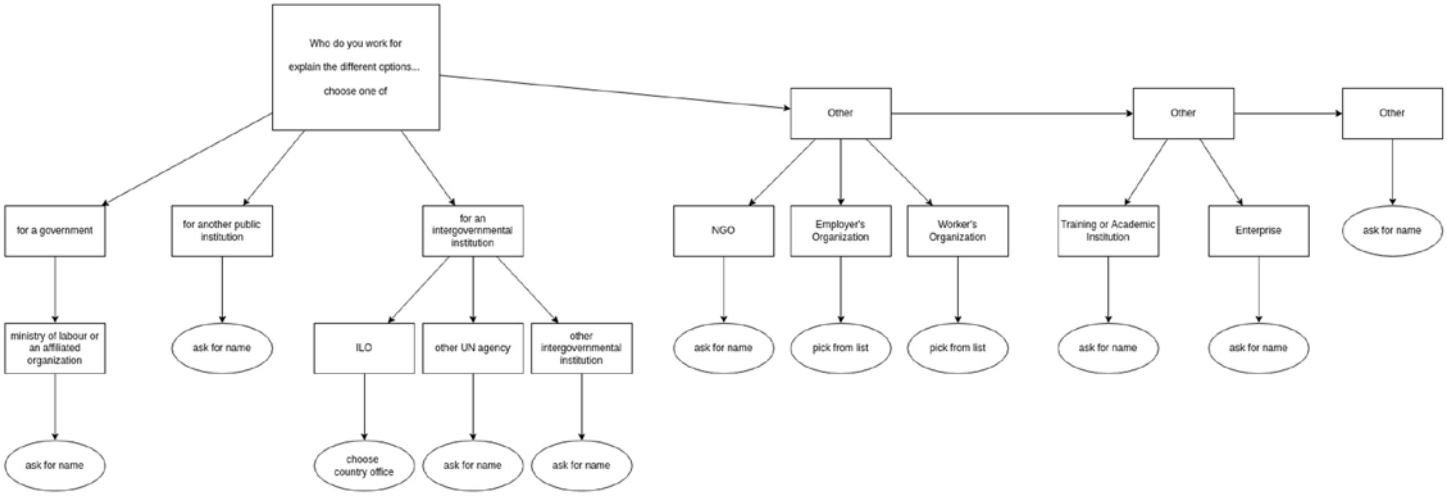
Self-paid. Similarly, the fact that a participant is paying for the course out of their own funds is going to influence the depth of their commitment where the learning outcome is concerned. A participant may either be paying out of their own pocket, or be sponsored by their organization / funded by an ILO scholarship.

Professional context. Participants in courses come from different contexts and backgrounds. Naturally, therefore, their interests and approach to the learning process will differ. It remains to be seen whether there might even be differences in learning parameters, for example the distribution of learning periods over time, which in turn could be used to adjust learning scenarios to better suit learners.

Participants are asked to enter details on registration. Clean-up is necessary because in email communications incorrect or imprecise choices are often made. Participant can choose from the following options: Employer's organization, Enterprise, Government (other than Ministry of Labour), ILO, Intergovernmental organization (other than UN), Ministry of Labour, N/A, Non-governmental organization, Other, Public institution (other than government), Training/academic institution, UN organization (other than ILO) and Workers' organization.

By advancing the notion of repeat customers and reusing cleaned-up data, this issue will become less relevant, albeit always present due to the nature of the process. To completely eradicate this systematic error, it is recommended that a question tree be included in the registration process. By asking two or three questions sequentially instead of leaving this complex choice to the user, it is possible to further increase data quality. For example: "Are you working within a government?"→"Are you working at a Ministry of Labour?"→"Choose either 'Other' or 'Government'". By applying this strategy, two systematic errors can be fixed. Firstly, when two categories overlap (e.g. training institution and government), control is regained over which has priority. Secondly, human error is much less likely thanks to the simpler choices presented.

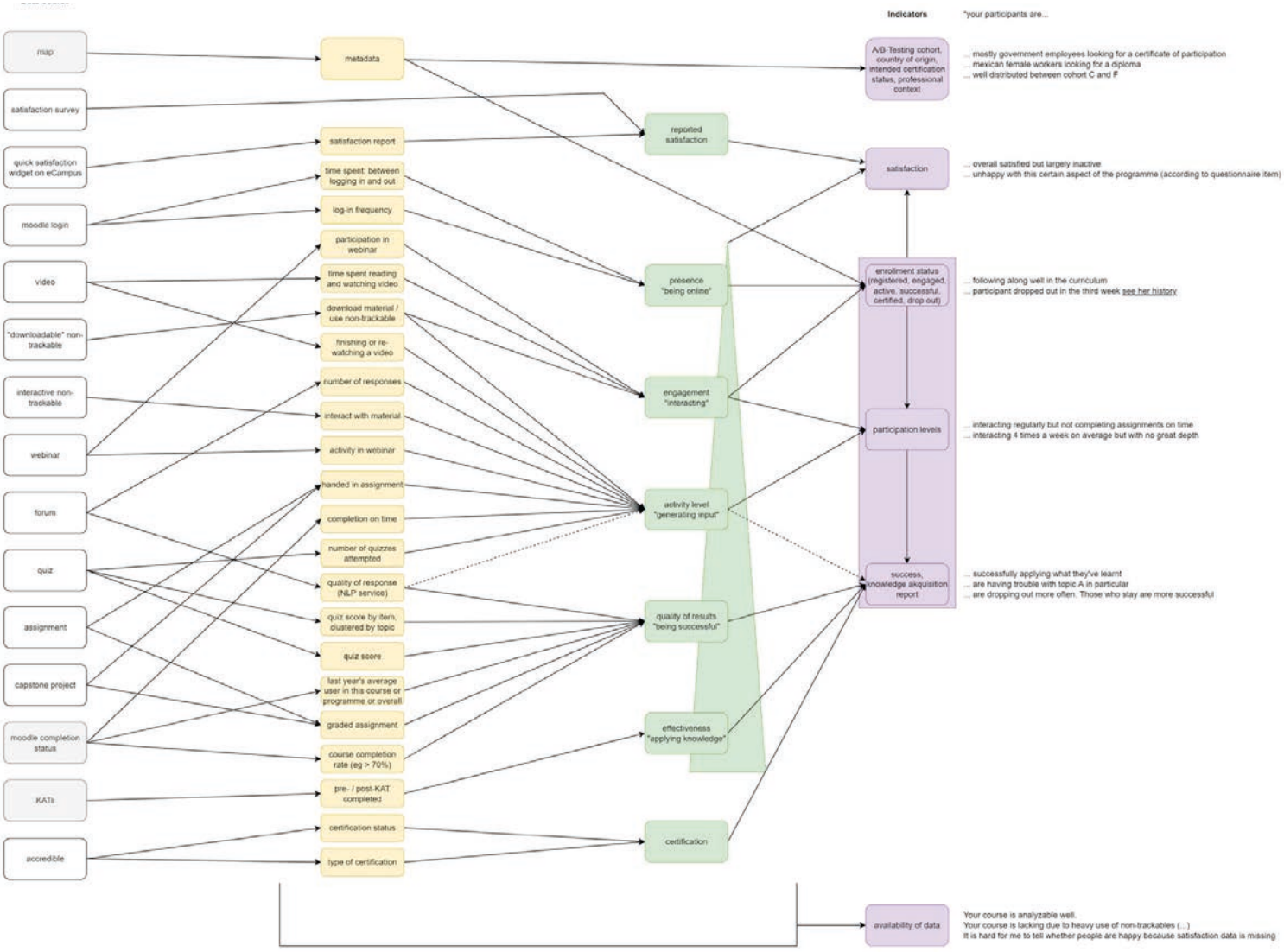
Below we propose a possible question tree. Its main goals include 1) giving no more than three clearly distinct choices at any one time, 2) providing dropdown lists wherever possible, and 3) putting the most likely choice first to minimize the expected number of questions. A participant will never have to answer more than three questions.



Instrument analysis

The obvious approach to selecting the correct indicators would be to start from a list of indicators and select the most fitting instruments (for example based on Schwendiman et al.). However, this shortcut would be detrimental for two reasons: firstly, the compound indicators required are unlikely to exist in the literature exactly as they will be needed for the Centre's dashboards. Secondly the ITCILO, with its Learning Innovation Unit, is in the beneficial situation of being able to test even the most cutting-edge technologies, if only for a few courses initially. We have therefore attempted to deduce the correct indicators and dashboard structure from first principles, as analyzed above, then determine what deriving the data needs to be shown, and finally apply UX design principles to make them as digestible as possible.

Visualization of data and what is really measured



The above illustration shows:

- available data sources (mostly activity types from Moodle, as described in prior sections);
- aggregated data points and events from these sources (yellow);
- the category of data, which can serve as a low-level indicator for some users (green);
- high-level indicators which can easily be communicated to the user (purple);
- example messages which might be given on a dashboard.

The category of data selected is classified in terms of depth of interaction. Starting only with metadata and self-reflected satisfaction levels, which are currently used as KPIs, the metrics become increasingly involved. These five levels are borrowed from the 2015 report, to which a link is provided in the Annex. We have refined them to make them more distinct in the data that is reported, and have also aggregated them further into the purple high-level indicators. Depending on the perspective of the user, the green or purple data might be shown. Quality Assurance and the Innovation Programme need access to the more detailed green layer, whereas an activity manager or unit manager not involved in learning design might want to access the purple indicators, which already draw some preliminary conclusions for his or her benefit. Naturally, the way in which these conclusions are drawn needs to be continuously monitored and improved upon. Without the work done by the Innovation Programme and Quality Management, none of these aggregations can reasonably be calculated automatically, since they are certain to be unreliable and imprecise at first while the Centre develops its expertise. Skilled supervisors are therefore needed to interpret, correct and iterate on them for all other employees of the ITCILO.

Compounding

The graphic above shows the compounding process in several steps (white -> yellow -> green -> purple). Compounding these metrics is a delicate procedure since it is the deciding factor in delivering relevant and accurate indications. Unfortunately, it is imperative not to forego this step, as the alternative would require an extraordinary degree of data literacy and time investment on the part of all dashboard users, which cannot realistically be expected. The result would be the user “drowning in data”, as Andreas Klemmer put it in 2020. Because of the complexity of the task, and as mentioned in the stakeholder analysis, certain institutions at the Centre should always be allowed a deeper look behind the compound indicators to enable them to improve the indicators and better understand the data. An “advanced” view will therefore be considered, giving graded access to more layers of complexity, as illustrated in the graphic above. Following the same thought process, it is also essential that the automatic compounding process give access to detailed in-depth data from the data sources themselves, rather than only data aggregated for the prior logical step. The above graphic visually implies discrete steps, which are introduced only for better readability. In summary, these layers should be seen as similar to an onion, whose skin can be peeled back layer-by-layer if the user so desires.

When compounding the data points to provide more meaningful compound indicators, some critical aspects need to be considered. Simply averaging the numerical values of specific indicators would result in important data being lost. For example, the fact that someone participates heavily in webinars does not imply above average activity if such is not the case for quizzes and other

materials. Quite the contrary: this scenario could imply a lack of motivation in self-regulated learning phases. Critical relationships of this kind could be accounted for manually using heuristics, or a deep-learning framework could be built to automatically analyze and adjust. Since this is not something readily available "off the shelf", due to the in-depth customization to the Centre's needs, we will refrain from recommending generalized and unspecified solutions here. Instead, a few examples of heuristic solutions are provided. In the roadmap section, we discuss how such deep-learning algorithms could be established in the future.

Relative dimensions

It can be argued that using absolute values when reporting KPIs often leads to higher variance in behavioral changes on the part of the user, especially when the data is only superficially interpreted. Relative reporting implies an immediate context in the user's mind, while absolute reporting broadens the focus. In the case of an activity manager, for example, actively controlling the context of interpretation by using relative values it is therefore more likely to produce the desired action. This can be achieved, for example, by showing last year's data or data from similar courses above any given indicator. The less a user is involved with the data, the more this approach supports decision making in a desired direction. The deliberate structuring of dashboards by including or omitting data related to desired behavioral change, and displaying relative comparisons as a means to provoke action in the user's mind, is an apt tool for expediting the standardization and formalization of online courses in an informal way, i.e. without needing to introduce new corporate rule sets. Lastly, another relative dimension of use, especially for the Learning Innovation Programme, is the comparison of cohorts, which can be applied to A/B test a new tool or didactic approach.

To determine similarity, it is obvious that topical relations should be highly prioritized. However, there are other attributes to be taken into consideration: a similar course can be found by examining the metadata and learning history of participating learners (both on the same course and more especially across related multi-course curricula such as diplomas or Masters degrees). Borrowing from modern recommender systems, the probability of similar personas as determined by "nearest neighbor"-type algorithms can be used to correlate similar courses, because these are more likely to be seen by the same kinds of participants.

In the following sections, we will explore in more depth how each indicator could be applied, and what heuristics could be applied to compound data in a sensible way. Exemplary messages (blue banners) are used to give an idea of what can be communicated to the user using the respective indicator. Each indicator description stated here would ordinarily be seen as a group of indicators. In order to provide a helpful overview, we will summarize some indicators, which will be displayed differently but present very similar structural data.

Examples of visual representations are given by means of screenshots. Naturally, there are many ways of showing the results and putting them in perspective with other information. These screenshots are merely ideas and starting points. They can and should be discussed continuously and amended to convey the desired message before they are implemented in a production dashboard.

Contextual indicator

This indicator takes into account metadata from the map. It provides an overview of different attributes, described in the participant dimensions section. We will not detail this indicator further, given the already well-established and sophisticated representation of this data in the IBI. Similar approaches can be applied here. The suggestion is that simple messages be derived from this data to help activity managers and their teams gain an overview of whom they are facilitating. When one has data concerning the participants, the training can be more personal, even before the first face-to-face interaction via video call. From the interviews conducted and from prior experience of similar projects, we believe it is detrimental to show graphs of this indicator in the first instance; only predefined messages with the respective ratio or metric should be shown. When interacting with the indicator by hovering or clicking, a useful graph can optionally be shown to allow a deep dive into the data, whenever necessary. However, our feeling is that the specifics of members' distribution are better covered in the advanced view of the IBI than in an appropriately restricted version of it.

Compounding is straightforward, more or less copying the data already collected from MAP and showing predefined messages when certain conditions are met.

Satisfaction Indicator

In essence, the satisfaction indicator gives quick access to the data collected from satisfaction surveys. As with the contextual information, automatically stripping the data of all items except those providing outlier data is a good way to condense information. A list of the five most deviating items - in positive or negative direction - can be shown and optionally expanded. Each will be assigned its current value and relative comparisons with the previous year and other similar courses.

In pursuit of Objective I (measure early, measure often), it would be prudent to retrieve rudimentary satisfaction data from participants as they are progressing through the course. This could be achieved by using a simple pop-up that could be shown once a week with a simple thumbs-up/down input field and an optional comment-text field. Alternatively, a slightly more complex five-step Likert scale using satisfaction emoji could be used. If there is a notable abnormality in the correlation with participants' activity levels, it is worth mentioning in this indicator:

your participants are happy, but not especially active your participants are happy and active, good job

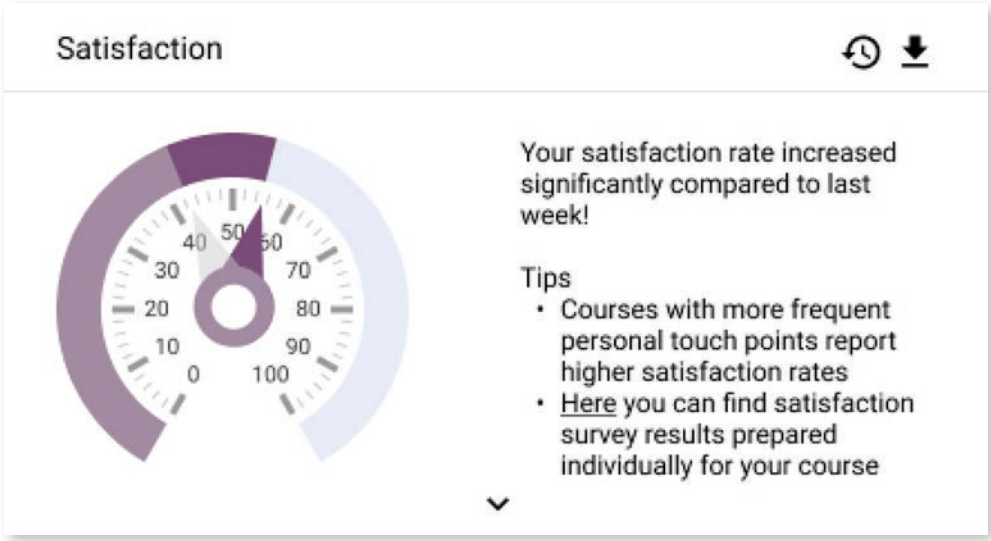
your participants are not happy, but very active compared to…

This indicator would need to communicate differently when it refers to the satisfaction survey, which is not completed until after the course:

your participants have been happier and more active than last year

your participants have been more successful but less happy than those from other similar courses

Compounding: if collected, the ongoing satisfaction data could be shown as a distribution over time. They should not be mixed with the survey results. The items from the survey results should be averaged and significant derivations from the same averages for similar courses should be calculated.



Enrolment status

Enrolment status is an indicator of the current state of activity on the course. Its main purpose is to give an overview of whether and how many participants are properly on-boarded and have started actively participating in the course. It takes into account the aggregated data from the presence-, engagement- and activity-level layers, as well as the metadata collected from the MAP. Participants are separated into five states of enrollment:

Registered: the participant has only registered and no further action has been taken, apart from logging in.

Engaged: the participant has started spending time on the course (for example, at least 10 minutes) and can therefore be considered to be successfully on-boarded:

your participants are almost all on boarded properly, except for: <name>

Shown after the first couple of days:

not all of your participants are engaged with the course.
Help <name>, <name> and <name> to get started.

Active: the participant has interacted with the course and generated learning data by starting a quiz, using an interactive piece of content, watching a video (almost) completely, etc. The user can be considered an active learner.

Successful: the participant is not falling behind schedule and has received sufficient grades (if appropriate). Consequently, there are two possible variants of "unsuccessful": falling behind or having insufficient grades. Both should be analyzed separately and taken into consideration when displaying statistics: your participants are on track to be more successful than last year's cohort

Certified: the participant has received some form of certification for this course or a multi-track course that encompasses this course. Both should be tracked and shown separately.

Dropped out: the participant has failed to take part in essential activities for too long and has been marked as a "drop-out":

you have managed to maintain a low level of drop-outs compared to last year

Much of this data is already shown on the current activity-manager dashboard on the eCampus. This view can be expanded to show more detailed status indications for each participant.

Compounding is straight forward; calculation of the respective data sources is described elsewhere.

Participation level

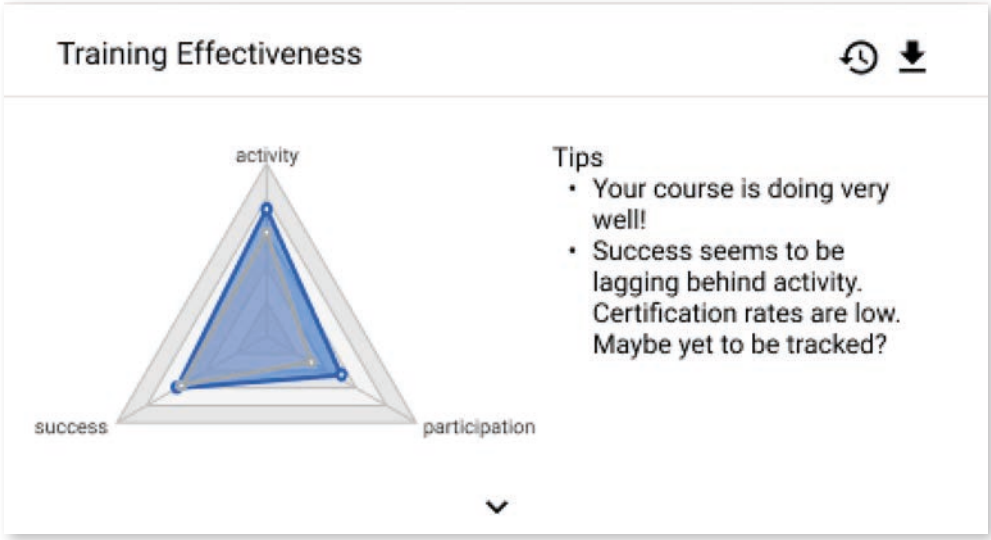
While the enrolment-status indicator already provides high-level information about participants, it is essential to go into detail and understand what exactly is happening at the micro level and have the opportunity to make a tangent analysis at the meso level. From the perspective of a cluster manager, for example, courses might be performing decently but a little worse compared to the previous year. There is no abnormality in the overall participant-enrolment indicator apart from slightly worse numbers than before. The solution in this case lies in the introduction of a new tool, used on most courses, which results in equal changes in performance across all other dimensions. When participation levels are viewed by activity type, the problem can be tracked down to the new tool, which apparently is not well received by participants. This indicator only takes into account interactive users and will filter out users who have signed up merely to retrieve more relevant data about those who are truly active, and displays it over time to allow inferences as to changes in participation in a certain week of the course. This indicator is concerned with generating an aggregated participation score and distribution across different activity types. More information about a participant's completion of activities can be extracted from the existing activity-completion table available to the activity manager's team.

Compounding: All participating factors shown in the graphic should be separately aggregated into their own metric on a per-learning-activity basis. Only those with enough data should be kept; others should be neglected for the purposes of the calculation. In the case of videos, for example, the time spent and the number of interruptions per session might be collected, while for activity in a webinar the time tabbed out of the window might be subtracted from the total minutes of participation. Each metric should therefore be represented as a positive absolute number which increases linearly with better performance. To aggregate the metrics, they need to be normalized to ensure that they are comparable. In our opinion, taking learning-minutes as the unit is the only viable option. Each data point which is not already measured in minutes should be somehow converted to minutes. Forum entries can be converted by calculating number of words * depth of content ratio * seconds per word. For as long as the content depth cannot be calculated, it can be omitted. Downloading a non-trackable unfortunately needs to be assigned a fixed number of minutes, which can only be modified by page or word count if it is a PDF file or similar statistic. The question of whether the user actually reads all of it remains unanswered. Finally, before they are aggregated, all the different metrics can and should be weighted, since watching a video does not have the same value in the learning process as active participation in a webinar. A score can be computed by comparing the added sum to the average across all or a cluster of courses.

Success indicator

The success indicator aggregates data from the activity, quality-of-results and effectiveness layers. Most importantly, it shows quiz scores and assignment grades, and also takes into account the completion rate. As a next step, analysis of the written forum entries can be included. As shown in several papers, the quality of forum threads can be assessed to a large extent automatically. Fundamentally, depth of content is used to evaluate quality, which correlates highly with learning success. Last but not least, the pre-KAT and post-KAT evaluations, which are mandatory for all courses, can be used as a reliable after-the-fact measure of success. When transitioning towards training a neural network to automatically measure learning success, the improvement from pre- to post-KAT can be used as a goodness-of-fit test for the success data collected on the ongoing course, assuming the KATs have been carefully validated.

Compounding is achieved by averaging the success ratio as a percentage value of all data points. If a specific activity is not assigned a grade, it can be easily calculated from its contents: the completion rate can be directly used as a success ratio, as can all directly given grades or quiz results. For ungraded assignments and other activities such as forum entries, NLP can be applied to detect the relative content depth compared to other users, where the best user would receive 100 per cent and the worst 0 per cent (which usually would be a short response in the case of a forum: “Thank you”. Until NLP systems are in place, the content length can be used for making an estimate. This again can be shown as a spider chart, with each activity type or submission separated, or as an aggregated total score, for which the different activities first need to be weighted.



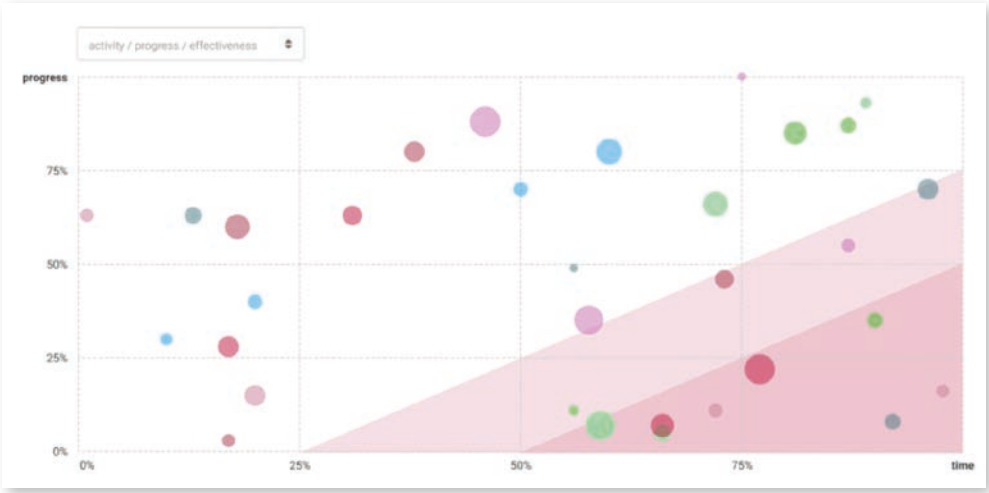
Course progress by activity, participation and success

Especially on the macro level, a bird's-eye view of all participants is necessary. For this purpose, a stacked bar chart is well suited to display both the total number of active participants over time within the average course and their respective depths of involvement.

In this chart, a given bar represents the nth week of a course. Given the varying lengths of courses, a common number should be chosen, with all other courses interpolated respectively. If there is no common ground to be found for course length, a smoothed line chart can be used. Towards the end of the course, one should see the share of successful participation increase (as measured by intermediary results), while the share of participants who are merely active decreases. The total number of participants will decrease over time, depending to a greater or lesser extent on different parameters. Of course, the mode of training is highly relevant to this data and users should be allowed to make distinctions by mode of training. Depending on the objective, which should be discussed regularly, it is also possible to use aggregated numbers of participants instead

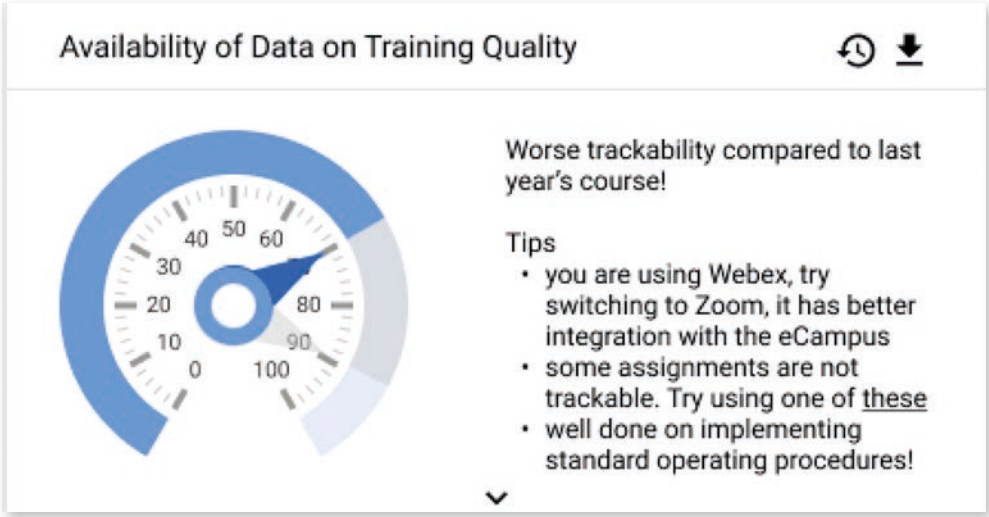
of weekly measurements, or add the certification dimension as a fourth color, at the same time extending the timeline to after the official end of the course.

The following example of an indicator as used in the prototype shown later demonstrates a high-level overview at the meso level. It gives an overview of which courses are badly placed in terms of several dimensions. See the prototype for more in-depth information.



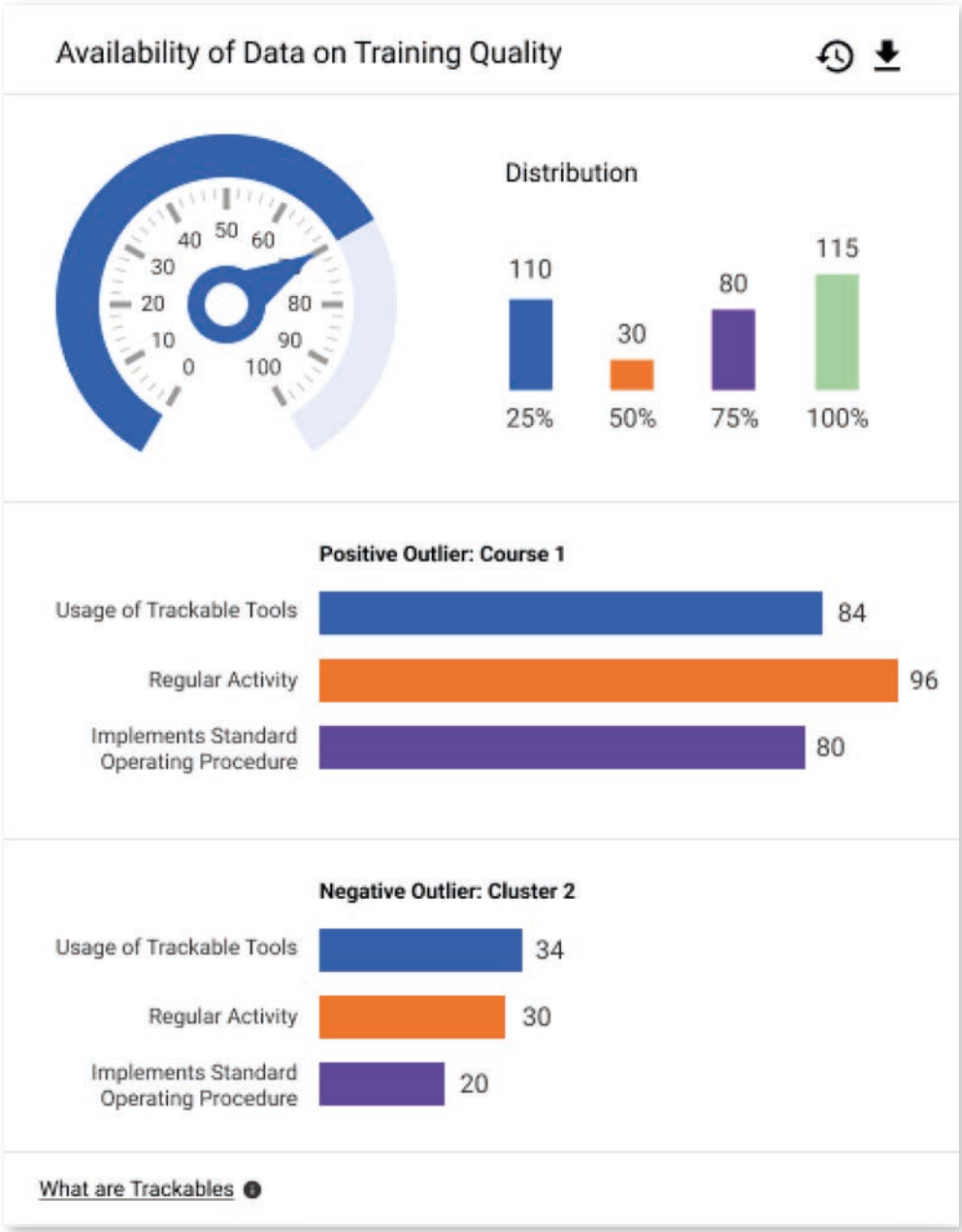
Availability of Data

This meta-indicator is an important step towards achieving better data quality and in essence provides an outside evaluation of how well a course can be measured in terms of the indicators used.



All the metrics and data points used in compounding should be taken into consideration here. The result is twofold: a list of data points that have not been collected due to technical limitations, and a score for the course or cluster. This score then neatly indicates how well a course is doing in terms of learning analytics and gives a first clue as to where internal training might need to be offered. The insights gained from this indicator can also be used in the dashboard to show the level of confidence in other indicators. For example, when there are no KAT-tests but a KAT is essential in evaluating success, these indicators are statistically less significant and should be suppressed before they can be misinterpreted by a less skilled end user.

Compounding: The data-availability KPI is well approximated by the ratio of activities on a course which provide data vs those which do not. All activities can be extracted via the Moodle API, which also serves the type of the activity.



Starting point for a prototype

The following link contains several ideas for indicators which could be shown on existing dashboards.

https://www.figma.com/proto/LYLumcsM17Q1g8iF9m2vye/Dashboards-ITCILO?node-id=225%3A1190&scaling=min-zoom&page-id=0%3A1&starting-point-node-id=225%3A1190

We avoid direct placement into the existing dashboards since the established common language should not be broken with. Precise positions should be discussed in detail with the person in charge of the ongoing design of the dashboards. Rather, we simply overlay them in the rough area where we think the indicators would fit in well with the existing user flow.

You eat with your eyes first

To achieve a high level of adoption and participation in the continuous development of the newly introduced tooling, it is important to use state-of-the-art design methods that appeal to the user. The so-called "Aesthetic-Usability Effect" ensures that poor functionality is gladly forgiven by end users provided there is a visually appealing interface. This effect is inversely proportional to expert knowledge of a field. In other words, the less people know about the subject, the more important visuals are. By providing appealing interfaces, which is hardly a serious problem given the marvels of present-day technology, the Centre can bridge the gap between motivated pioneers in the field and other staff in whom data literacy has yet to be instilled. Pioneering a project is significantly easier if there is an appealing narrative to accompany it (Tesla).

Avoid habituation

There is one particular phenomenon we want to highlight. Indicators in LADs are designed to guide the user's attention towards a specific actionable that management deems desirable. However, the effect of habituation strongly counters this objective. After seeing a seemingly identical chart dozens of times in the same place without anything new to catch the eye, users will soon grow tired and scroll over it - or worse, develop an aversion to the widget. To counter this, there are two measures that can be taken:

1) Only show what is relevant. If it is not relevant anymore, hide the information so that it gains more attention when it resurfaces. This applies mostly for "seasonal" statistics, such as the certification rate in the micro view. An activity manager will not be interested in this statistic until the very end of the course (the situation will be different at the meso level).

2) Hide indicators that convey little information. Even if a graph is applicable to the user's situation, it might not contain insightful information. For example, showing that activity rates are performing within a standard deviation of the mean will prompt the user to question the existence of this indicator. An indicator would better be hidden in this case, even if this means that only one or no indicators are left on the page. Expressed in the language of front-end designers: giving the remaining information space to breathe results in better buy-in with users,

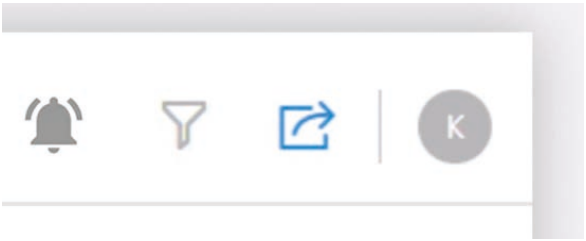
thanks to a more pleasant experience and better perceived information content.

Explanation of Controls

We suggest several controls for many of the indicators offered. Not every control should apply to every indicator, as you can see in the demo mockups.

The history icon controls the time dimension. It allows the user to relate every indicator to its prior versions. Depending on the indicator, this could be an overlaid shadowed version of itself or a simple bar chart which plots historic values. We have found that a common follow-up question to getting a piece of information is the question of how this relates to the history of the data.

The second control helps advanced users access the data they are interested in. It gives access to source data tables which can then be processed further with Excel or any other processing software. This is especially useful for gaining insights for advanced reporting or for developing new ideas for innovation meetings, as discussed further below. Continuous development of the indicators assumes creative thinking on the part of all staff, therefore drill-down data should be easily accessible if required.



At the top right, there is a new notification icon which indicates relevant changes in data. As explained above, there should be triggers which are defined manually at first and can be automated in a later release. These triggers give rise to a notification to be sent out via email and shown in the notification center. These might include course or participant outliers regarding the activity or success dimensions. They serve as proactive pointers towards indicators that could also be discovered manually but, because of conspicuous activity, are worth nudging the user towards.

The way forward

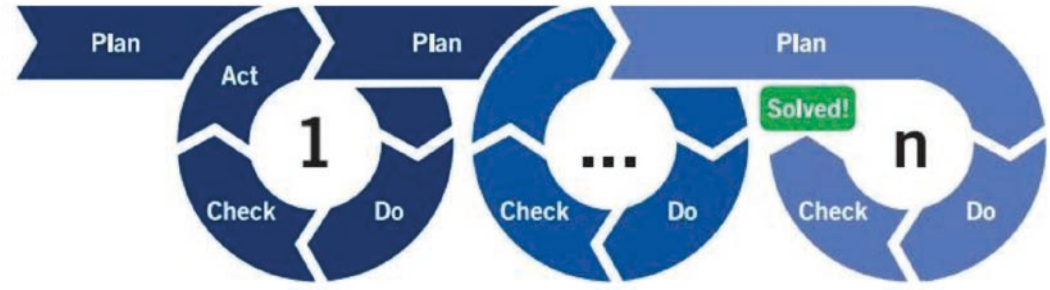
PDCA in Learning Analytics Dashboards

As mentioned in the introduction, this report shows first and foremost how to continuously adapt the dashboards to the changing questions posed by users and, more importantly, to consider new questions that are relevant to the current challenges faced by the different stakeholders. The dashboards will only ever be as good as the questions asked while developing and using them. Hence, it is essential to establish an officially endorsed process or "ritual" that encourages joint discussion of the topic across the different programmes. As recommended in the previous report, the PDCA cycle is well suited to providing a structure for these innovation meetings.

The procedure described in this chapter can be used for tackling various problems and opportunities relating to the issue of learning quality within the Centre. Given the nature of our report, we will focus on learning dashboard design, but the process and meetings can and should also be used for related issues:

- Evaluate and improve LA indicators and therefore LA dashboards.
- Test new didactic approaches.
- Introduce new tools to accommodate and improve existing didactic approaches.
- Experiment with cutting-edge tech on a small group, while making it possible to share the new knowledge across the whole team.

The PDCA cycle is well suited to accommodate all kinds of stakeholders in learning analytics: experts, users on various levels, outside observers and uninvolved staff. Since the goal of this process is innovation, it may be beneficial to invite outsiders to enrich our own perspective. In the current pandemic situation, buy-in will be increased. Quick adoption and communication of this process could therefore be beneficial in saving overheads, as opposed to managing the process of change later, when the focus may be back on face-to-face activities and it will be more difficult.



In the following section, we will start with the "Do" (second) stage and consider the "Plan" stage at the end of the circular structure.

Do - everyday work

The "Do" stage encompasses day-to-day courses, tutoring and coaching. Daily business takes place in between innovation meetings. This stage happens informally and takes up the time between joint meetings. To gain enough insights to warrant another iteration of the cycle, day-to-day work should be allowed for at least four uninterrupted weeks.

Check - in preparation of meeting

In the "Check" stage, available data and visualizations are used to compare expectations with outcomes and this information is used to answer questions that have been asked in prior cycles. This step should be completed before the meeting by everyone attending. Existing dashboards should be used and evaluated, and indicators created as a result of prior meetings should be read and interpreted thoroughly. We suggest allowing a further week after two weeks of uninterrupted day-to-day work to aggregate findings and prepare for the meeting. Ideally, these findings would be collected using a form, to spare participants the additional task of formatting information, and then be distributed to all participants before the meeting to save valuable time.

Questions on the "Check" form to be collected before the "Act" stage might include:

- Did your team conduct A/B tests this month? How do expectations compare to the actual results?
- Which indicators were added or changed in the previous period? Do you feel that these were helpful changes? Is there still room for improvement?
- What changes came up in your workflow that lead to new or changed requirements as to which indicators you need or do not need anymore?
- Are there any other general topics that you want to drive forward in the upcoming meeting?

Act - in meeting

Starting from the improved baseline worked out in the "Check" stage, the process is now critically examined and proposals are made on how to improve the reporting. Every stakeholder will have found shortcomings in their data visualizations that prevent them from using them effectively, or they will have new requirements and questions they want to have answered in the next cycle.

Several shareholder groups should be present at these meetings and should take on different roles.

The objective of the **Office of the Director of Training** is to increase the overall quality of training. It is therefore recommended that a member of the Office should chair the meeting, should lead the discussion and guide the respective expert participants, while allowing them to find the best solutions for their respective fields. Given the complex nature of learning analytics, it is not advisable to seek common ground between the programmes, but rather to let each unit decide what is best for its area of expertise, at the same time seeking to define a common understanding

of innovation processes as described here. We also recommend that the Office of the Director of Training be responsible for deciding which participants to invite.

The **Programme managers** are in the innovator role during these meetings. They will have used the dashboards or aggregated other data with their teams and are therefore well informed as to the nuances of the problems involved and the solutions attempted during the previous PDCA cycle. The objective of the programme managers is to share their views and spark healthy debate as to whether an idea might be suited for others, and as to the advantages and disadvantages of particular developments.

Guest Visitors could also be invited to a meeting on account of their insights into a specific topic. Their role would be that of experts reflecting on a case study or on data that had been collected. Presentations should be used sparsely to keep meetings short, and should be subject to prior consideration.

A typical sprint meeting might be structured as follows:

1) Two-minute **presentation** of findings and changes from each programme. This would also be an opportunity to celebrate findings or changes that had worked especially well
2) Each participant **contributes** a question or discussion topic, which they deem relevant for the future development of training quality at the ITCILO. There should be novalidation of topics at this point. Every member should be allowed to scrutinize whether the right questions are being asked and answered by current dashboards.
3) Dot voting to **prioritize** topics. Each participant is allowed three votes and visually places (digital) dots on the respective topics.
4) Open **discussion**. The topics are discussed in order of priority. As many topics as possible are covered within the given time constraint (most likely 30-45 minutes). Other topics may be discussed informally on another occasion or at the next meeting.
5) Conclusion and **requirements**. The moderator of the meeting should by now have a "wish list" and therefore requirements that will enable the relevant ideas. No feasibility considerations need to be taken into account at this point, as this will take place at the "Plan" stage.

Taking everything into account, the meeting should take no longer than 60-90 minutes. It is more valuable to conduct these meetings regularly, like an inner-circle ritual, than to ensure that all items are covered. In the end, the PDCA process is about continuous improvement in relation to lengthy planning cycles and this meeting should reflect that spirit.

Plan - after meeting

To start a new cycle, the proposals from the "Act" stage are taken into consideration and concrete steps are taken to provide better information on the dashboards next time. The results should be summarized and, to maximize impact, published for the benefit of outside staff who have not participated in the meeting. The technical steps are taken to install the necessary indicators on the dashboards.

This stage should take another week. It could be argued that the technical measures might take longer to implement due to their unforeseen complexity. In this case, in order to maintain the improvement rituals at regular intervals, the changes or new indicators should be implemented during the cycle and used in the next iteration. This idea is derived from SCRUM methodology, which is nowadays applied in most projects involving technical implementation.

The **ICTS** has an important role to play at this stage. The feasibility of the requirements collected from the "Act" meeting should be analysed. A concrete action plan and development sprint should be planned in collaboration with decision-makers and the Training Director's Office. It is important to allow the technical experts to judge where some low-hanging fruit can be more easily picked, rather than other seemingly simple changes. At the same time, the "Act" meeting should be based on first principles, not swayed by technical considerations. Technical considerations should not determine the direction of innovation, but must lead the way when it comes to deciding on the concrete action plan based on the results of the innovation ritual.

The specific outcomes of this stage might include:

- A newsletter with a brief summary of the meeting's topics and discussions, which can possibly be shared with the whole organization;
- Feedback on, e.g. in the form of an email from the ICTS with suggestions on which indicators to implement quickly and which to avoid for the time being. This should be discussed with the people who conceived the idea and the Training Director's Office, the better to judge relative importance;
- An indication of the resources and time needed to implement the necessary steps.

When planning next steps, measurement needs to be treated as a central component of any proposed innovation or change. Quantitatively and qualitatively comparing outcomes in relation to the current situation is essential in coming to informed conclusions about where to go next. Measurement should be a factor in all four phases of the PDCA cycle: when plans are made, the metric to be used or evaluation to be conducted should be decided at the same time as the change itself. In the "Do" phase, measurement needs to be monitored, most probably by the ICTS. At the "Act" meeting, all participants need to hold one another accountable for formulating changes and ideas for which there is a measurable outcome.

Not all measurements need to be quantitative: one could just as well conduct a qualitative interview with selected participants after a change has been implemented to measure its impact. Regardless of the type of measurement, an appropriate indicator will most likely be created in the long run, therefore the analyzability indicator also needs to be updated to include a heuristic for how well any course is trackable in terms of the new indicator.

Learning quality KPIs

When leading the meeting, it should not be a priority for the Office of the Director of Training to home in on governance KPIs, but rather to detach itself from existing KPI structures so that new and improved metrics can be found. To truly report on quality of learning, it is important to attempt

to cover the innumerable subtle differences in learning design necessary to accommodate the various topics represented at the ITCILO. These nuances can only be discovered and reported on by the activity managers, tutors and coaches working directly with

participants. While not the primary focus, there is also an opportunity here to improve reporting to the Board and the public by using new metrics to drive home important aspects yet to be included in the KPI structure. Some KPIs are already recommended in the macro dashboard described earlier.

Selected details of implementation

Data Warehouse vs Data Lake

Contrary to popular belief, these two options are not mutually exclusive; they simply serve different purposes, which may well be needed at one and the same time. A data lake is a large, unsorted collection of raw data, often decentralized across several systems. A data warehouse is characterized by its structured definition of used data and its well-defined processes for data movement. It can be relied on to be correct if constructed properly.

Currently in the case of LA data within the ITCILO, the same IT system is used, which also serves the financial and governance needs of business intelligence. As explained previously, Kettle is used to feed data into micro-strategy at regular intervals. Within Kettle, a fixed data format is assumed to be able to communicate well with its subsystems and report generators. To a certain extent, this system can be applied to meet the challenges described in this report. Especially when reporting new high-level indicators to the IBI, it is prudent to draw on existing solutions. While the system does not technically deliver “live” data, a daily fetching interval from the eCampus could very well be construed as providing sufficient regency to the user.

In the long term, establishing an overarching data-lake system has several significant advantages. Primarily, a lot of overhead is removed from the process of establishing new streams of data. In data-science contexts, a data lake greatly speeds up many everyday processes in contrast to a warehouse. This especially applies to creating and training neural networks based on data that is constantly changing. The more often data sources, processing, interpretation or rendering to end users change, the more likely a data lake is to bring measurable benefit in terms of time-to-market for any newly introduced functionality.

The future data lake can still be integrated with the current data-warehouse solution and exchange data seamlessly, and Kettle can continue to serve data to technical systems where structured and pre-fetched data is needed. In the long term, to avoid significant overhead, very new and highly unstructured event-based data sources should not be forced into the data-warehouse strategy. For consistency, MOOCs should be better integrated into the current data flow (listed in MAP) and regarded as “first-class citizens” due to their increasing importance.

Course similarity

An important first step is to establish a way of measuring similar courses. Without this, the relativization of indicator metrics based on other courses would not be possible. As almost everything else suggested here, this can be approached in several steps increasing in complexity and precision:

- Use courses that are assigned to the same cluster in MAP.
- Use courses that have similar attributes as measured by a Levenshtein distance or similar metric (mode of training, participant dimensions, etc., depending on the data analysis section).
- Apply fully-fledged nearest-neighbor algorithm to participants in order to deduce likeness of courses.

Roadmap

A couple of key questions present themselves when one is confronted with actual implementation of the described indicators, in order of the data flow: 1) collecting missing data, 2) extraction via Kettle, 3) processing and possibly taking the time dimension into consideration, 4) tracking the success of new indicators, 5) creating new stakeholder interfaces, and 6) integrating new interfaces into existing systems and workflows.

Collecting missing data: When it comes to collecting sufficient data for the aforementioned indicators, some data has already been collected, while other data is yet to be created. The following list constitutes an overview of what methods would need to be established:

- Collect events from Moodle.[2]
- Collect interaction events directly from users' clicks and movements.
- Extract interaction data from the video platform provider. Currently, WebEx and Zoom are used. Both are proprietary solutions which do not grant access to the lower levels of the application. Data collection is therefore difficult. Alternatives include Jitsi or Big Blue Button, both of which can be self- or cloud-hosted. In both cases, there are popular attention-tracking plug-ins that can be used to provide the necessary data.
- Interpret content depth in forum posts and written assignments. The *tf-idf* algorithm can be used to extract all words from the Moodle course, ordered by relevance. Then, based on the words relevant to the course, a user's coverage of these words can be used as an indication of their degree of reflection on the topic. In addition, the time other users spend on a post, taking into account the length of the submission, is an indication of content depth.[3]

Extraction: Given the current situation, pre-processing needs to be performed on several of the data sources, since the vast amount of unstructured data cannot be processed without major changes to the current system. This could be achieved using a micro-service architecture. An "aggregation" micro-service might provide an API which fetches data internally, for example from Moodle's event database, processes it and serves it as structured and well-defined data to Kettle.

[2] https://docs.moodle.org/dev/Events_API

[3] http://ceur-ws.org/Vol-2068/esida1.pdf

In addition to the above-mentioned new solutions, the Moodle learning analytics indicators can be put to use whenever there is a good fit for an indicator. These are mostly low-level, direct measurements of events, but provide a shortcut for some calculations which otherwise would have to be performed manually based on event data.

Processing: Currently, internal mapping needs to be performed to make the data ready for use in micro-strategy. It is possible that this step could be worked around for those dashboards that do not run on micro-strategy. Some data might even be processed within Moodle by adding a custom plug-in instead. For the data, which needs to be processed by Kettle, it is necessary to strip the time dimension, given the otherwise significant overhead of comparing two snapshots within Kettle. Data can be reduced in such a way that relative values to a prior state are already calculated and served to Kettle as fixed values to be further processed. Neither of these open considerations can be fully answered without an in-depth analysis of current systems but practicable solutions do exists, as discussed with the ICTS.

Tracking: Certain metadata is needed for the aforementioned rituals and meetings concerning learning innovation. Notably, now there is no way to assign a user and course pairing to an A/B-testing group. By allowing for A/B labels, trial runs of experimental features are possible for evaluating and directly comparing certain ideas with the control group, who will not receive the update. This data is essential for making informed decisions in the continuous improvement cycle.

Creation of new interfaces: The technical realization of the respective interfaces is highly dependent on where they are going to be embedded. On the macro level, at first glance micro-strategy is probably able to cover most use cases. The micro dashboard can be integrated directly into the eCampus Moodle. Therefore, Moodle plug-ins or theme adjustments need to be created. As we are aware, there is an external software partner responsible for reworking the custom Moodle theme, who can surely assist in this task. There are several extraordinarily well-thought-through graphing libraries for web frontends that can be used (apex charts, d3, ...). Some parts of the meso level can be integrated into Moodle as well. Others will probably be integrated into micro-strategy. If, contrary to expectations, neither is possible, another small micro-service can be set up to serve a simple frontend, which can then be linked from either of both existing platforms.

Integration: To increase buy-in as previously discussed, it is beneficial to embed indicators directly into existing interfaces. In effect, existing views can be “upgraded” to be a dashboard. As mentioned in the previous paragraph, this can mostly be achieved through Moodle and micro-strategy. There are several technical possibilities for integrating new charts, which do not need to be discussed in detail here (iframes, plain links, web components, platform specific plug-ins, etc.). During and after the implementation of the first indicators, we should closely monitor their perceived discoverability by uninitiated users.

These steps are best conducted multiple times as part of an agile process, much like the already adopted PDCA cycle, as explained at length in the 2020 paper. As a first step, viable short-term indicators could be determined, taking into account what the Learning Innovation Programme deems especially beneficial. In time, increasingly difficult changes can be attempted. By following

this approach, a process of continuous change can be introduced, measured and corrected. Flaws in logic while planning certain features, or low adoption rates due to bad usability, can be fixed relatively quickly. Close collaboration between Quality Assurance, the Learning Innovation Programme and ICTS will be necessary.

Standardization of Online Learning Offers

Creating a course from a basic idea of the learning design to be achieved is tedious and will result in wildly varying outcomes, depending on the person producing the course. To ensure high quality across all offers, and especially to allow for the tracking and analyzing of learner performance, some aspects should be standardized. This should be the ongoing result of the continuous PDCA process, which sets and refines minimum quality benchmarks for all courses.

Before conducting an in-depth analysis of which parts of training should be formalized and standardized, it is important to note that this section is not meant to convey the notion of data as an end in itself. It stands to reason that, due to the growing desire to collect data as part of the overall trend of learning analytics, everything needs to be cut from the same cloth if it is to be comparable. The KAT tests may also have conveyed this notion, asking only high-level questions that barely scratch the surface. For the future, however, we recommend a different mindset, in which standardization always follows first principles and merely provides a necessary shared skeleton for each course.

For example, in the shared PDCA ritual the programme managers might discover that all programmes conduct a number of assessments prior to the final assessment, which are used to indicate progress. Only some of them include multiple-choice questions. The correct procedure in this case might be to make weekly assessments mandatory and leave each activity manager the freedom to design and grade the assessments in a way that best suits the content of their particular course. In this way, standardization follows first principles and does not dictate any specific design choices. Subsequently, as mentioned previously, the KPI reporting structure should also respect the innumerable subtleties in learning design and allow for free-flowing innovation to occur, without adding rigid constraints to the process.

Tools

While it would be detrimental to force activity managers to use any particular tool apart from the basic eCampus platform, we would suggest that tools be recommended based on the desired outcome. The tools currently recommended are already examined in terms of trackability, reliability and quality, and have been chosen as being superior to their competitors. Details of these tools could be published in an internal knowledge base to save all activity managers time during course design and to enforce the minimum quality criteria set for tools used in ITCILO courses.

Here is an initial draft of a possible tool list:

Outcome	Best choice	Unreliable integration	(Currently) untrackable
Quiz Assessment	Moodle onboard quiz	Kahoot, Articulate	
Video	Moodle onboard		YouTube, Vimeo
Webinars	Big Blue Button Moodle-integration	Zoom	Webex
…	…	…	…

Standard operating procedure

At the end, the effort invested in this procedure will only have an impact on training quality if the insights gained reach all staff quickly. The individual results of each of the PDCA iterations amount to a collection of good practices and lessons for the whole organization that can be collected in a knowledge base. With a working title like "Standard Operating Procedure for Course Activities", this collection of documents could serve as a common basis for all activities within the organization. A standardization metric could be devised to determine what percentage of courses follow the procedure.

The knowledge base should also convey examples of good practice, not only for a complete course set-up in different scenarios but also for specific niche edge cases. Using these templates will both encourage adoption by saving time and refute the idea that following standard procedure generates too much work. Examples of good course design can regularly be highlighted and shared with colleagues across the organization. This will help to extend the reach of the innovation ritual from the inner circle to all activity managers. Good ideas, including ideas from other areas or external collaborators, can thus be propagated more quickly to all staff.

The current state of standard operating procedure is influenced by two major factors: the structure of the eCampus and the necessity of implementing KAT tests. The creation of a standard operating procedure applicable to all staff provides a great opportunity to revisit some of these topics and communicate the reasoning behind decisions that have been made to the learning designers who use the knowledge base.

In the future, more in-depth analysis of different didactic approaches will be necessary to provide staff with a solid example of best practice for their specific scenario. Initially, examples of good practice for some of the more common didactic approaches could be developed jointly by a member of the innovation team and a staff member from a particular team or cluster, where the programme manager might play the role of mentor. The result of this one-off collaboration could serve as a scalable basis for similar re-designs within the same programme and, if possible, in other programmes as well.

Implications for Professional Development and Recruitment

Due to the ever-increasing demand for a high degree of data literacy and good digital skills on the part of all Centre staff, we suggest that a concerted effort be made to improve the average skills level in these fields. In the Learning Innovation Lab, the ITCILO has a powerful and unique instrument at its disposal; namely an internal academy for building the necessary skills with and for its own staff. Due to the Learning Innovation Centre's involvement in the innovation process, the programme already has all the information needed for providing training opportunities to ITCILO staff, and this is already happening on a regular basis (with several courses in progress on e-learning, online learning design and learning analytics). We recommend making these opportunities mandatory in certain situations and focusing on quickly spreading the knowledge gained. In our view, this represents the single biggest bottleneck to achieving excellence in the areas discussed in this report. As long as innovation and carefully considered standard operating procedures reach only those who understand them, organizational transformation cannot be achieved.

Naturally, these considerations should also be taken into account in the recruitment process. When hiring activity managers, data literacy and familiarity with digital tooling in training scenarios should be very high up on the list of selection criteria. The analyzability indicator can give insight into the progress being made in this area and should be used to drill down and find out which clusters and courses are lagging behind. Without putting any staff on the spot, this information can be used to recommend helpful training materials as an aspect of regular professional staff development within the Centre.

Recommended by Tom Wambeke, the "gapminder" principle can be a powerful tool for sparking curiosity and motivation among staff. The Gap minder website[4] opens with the statement: "You're probably wrong about…" and focuses on the UN's SDGs. Questions about what you are certain to be able to assess are made and give three viable options. Very often, the correct answer is not what you think. This principle could also be applied to common e-learning myths and misjudgments which activity managers could be trapped in. For example, there might be a question about the distribution of learning activity over time. The results can spark curiosity concerning coherences in data and phenomena that would not otherwise be explored.

Future Possible Objectives

While the cornerstones for a continuous improvement process have already been laid in previous sections, dedicated attention needs to be paid to examining complex correlations in data or entirely new areas of learning innovation that go even further. Here are some examples, described very briefly:

- When a substantial portion of courses are standardized, reiterate the micro levels of dashboards to make use of more in-depth indicators targeting certain course designs. In this report, we are restricted to very abstract indicators, given the present lack of standardization. With a more

[4] https://www.gapminder.org/

in-depth analysis at the micro level, much more thought could be given to such ideas as self-regulated learning to allow for metacognitive, cognitive, behavioral and emotional stimulation, as described in the earlier 2020 report.

- Epistemic analysis of user contributions on courses could lead to an earlier and more automated understanding of the real level of understanding of a given topic on the part of participants. Some assignments not critical to the total course grade can probably be graded automatically using topical analysis methods
- Deep learning applied to compounding of indicator data can lead to controlled automatic improvement of reflected data. There is a complex dialogue to be had about the automation of meta processes such as this.
- Measure the impact of the dashboards on staff and participants. Up to this point, dashboard design has had to rely heavily on known good practices and gut intuition. A better way to progress would be to conduct a thorough investigation of the isolated effect of each of the aspects, for example by means of cohort or A/B testing.[5]
- An orchestrated effort is required to reinforce the use of certain didactic scenarios, e.g. constructivist learning. This can already be achieved to a certain degree by adding examples of best practice to the knowledge base. Thanks to the existence of the Learning Innovation Programme, there is an opportunity here to prototype examples of best practice for new scenarios that can then be integrated into courses, after they have been experimented with thoroughly and documented in the knowledge base.
- To support some of these objectives, tools for deep analysis will be necessary. For example, much more insight can be gained into the use of webinars, video consumption and free text exercises than at present.
- Deeper integration of arbitrary TinCan modules can be helpful in supporting more e-learning content (e.g. Articulate courses). By deeply synchronizing data from and to this content, activity managers are given more freedom in course design and choice of tools.

Next steps towards Level 4

By considering the problems described in this report and the actions to be taken, a solid foundation for the move towards Level 4 exists. In the following paragraphs, we dare to take a brief look into what lies beyond universally available and accepted dashboards.

The current design of the eCampus does not include any means of proactive automated communication with participants. Under the current plan, participants need to be activated manually by the tutor, coach or activity manager. This could be largely automated as soon as good baselines have been established for the activity and engagement rates to be expected on similar courses. The indicators described in this report could then serve as the basis not only for manual intervention but also for proactive communication via email or push notification with the participants concerned. In this way, activity managers, tutors and coaches could focus on learning success and effectiveness beyond mere activation.

[5] https://link.springer.com/content/pdf/10.1007/s11423-019-09693-0.pdf

On yet another level, a smart learning algorithm could be used to identify which participants are at risk from the very beginning, based on similar profiles that have been at risk in the past. For example, a study of MOOCs showed that participants from so-called "tight" cultures are more prone to drop out of self-guided MOOC-style courses. Many such correlations can be found and taken into account in the various processes. Since participants are sometimes evaluated not on merit but on correlations, this leads to important concerns about the ethics of such systems. This issue needs to be discussed and resolved before such systems are put in place.

By the same token, the dashboard can proactively notify staff as to which participants are at risk, without the need for regular checking of detailed dashboards. Rather, a well-designed home screen could be prepared for them before their work day begins, greeting them with a list of courses, participants or activities which require their attention. Automatic actions could be suggested to further assist the repetitive processes and allow more time for human interaction. This type of predictive analytics is a clear sign that Level 4 has been reached on the sophistication model and, in turn, implies yet another significant process of change. While the transformation to Level 3 is underway at the moment, the advent of predictive analytics will mean yet another big jump in the effectiveness and quality of the training delivered by the Centre.

With the at-risk detection features we have described and the more personalized learning achieved by closer attention to the micro-level dashboards, the Centre is well on the way to being an exemplary organization in terms of its use of learning data.

ANNEX

List of shared vocabulary

Activity manager: a person designing a course and leading course execution.

Cluster / Programme manager: a person leading several different courses and / or units.

Dashboard: any section of a tool that presents data concerning the current situation.

Indicators: visualizations or displays of condensed pieces of information that should directly answer questions relevant to the process.

PDCA meeting / Innovation ritual: a monthly sprint meeting attended by programme managers and leadership to reflect on the current state of education at the Centre.

Mode of training: face-to-face, tutor-guided or self-guided.

Trackable / Non-trackable activities: online activities within the eCampus which are either easily trackable or which do not provide detailed usage and interaction statistics.

Certification type: three common certification types. There are very different notions of success when considering.

Participant dimensions: region, multi-course track, self-paid, professional context. These dimensions should always be taken into account when reasoning is based on captured data.

INFO

FOR FURTHER INFORMATION PLEASE CONTACT

International Training Centre of the ILO
Viale Maestri del Lavoro, 10
10127 Turin - Italy

communications@itcilo.org
www.itcilo.org