12-12-2023

Optimizing Database Systems in Online Educational Platform sCoolMakers: Balancing Flexibility, Ownership and Data Structure

Jonathan Hombroek
University of Amsterdam, j.hombroek@uva.nl

Emőke Takács
ERI Hungary – European Research Institute, t.emoke@eri.net.in

Toon Abcouwer
Universiteit van Amsterdam, a.w.abcouwer@uva.nl

OPTIMIZING DATABASE SYSTEMS IN ONLINE EDUCATIONAL PLATFORM SCOOLMAKERS: BALANCING FLEXIBILITY, OWNERSHIP AND DATA STRUCTURE

Jonathan Hombroek
Faculty of Science
University of Amsterdam
j.hombroek@uva.nl

Emőke Takács
ERI Hungary – European Research Institute
t.emoke@eri.net.in

Toon Abcouwer
ERI Hungary – European Research Institute
toon@abcouwer.nl

Abstract:

Efficient data management and lifecycle strategies are critical for online educational platforms to maintain content quality, avoid redundancy, and ensure long-term viability.

This research studies the challenges of an online educational platform managing its database system while maintaining data ownership and a clear structure. We identified the need for strategies to prevent redundant content, the use of non-relational databases for managing educational content's lifecycle. Data lifecycle management, long-term data practices and determining responsibilities have emerged as critical components of effective data management.

The findings help to develop a comprehensive data lifecycle management strategy for online educational platforms, improve content quality, reduce redundancy, and realize long-term, sustainable data management effectiveness. They highlight the importance of incorporating non-relational databases into the platform to track changes, maintain data integrity, and inform content updates. Developing a save-state mechanism and establishing ownership and responsibility frameworks ensure data consistency and content duplication.

Keywords: Data Management, Data Lifecycle, Online Educational Platforms, Temporal Databases, Redundancy Prevention, Data Integrity, Ownership and Responsibility

I. INTRODUCTION

Since the beginning of this century, there has been a growing interest in sustainability and the circular economy. Educational institutes are increasingly incorporating these concepts into their curricula to educate students on their importance and relevance for the future [Brown, 2014]. To facilitate this process, various online platforms have been developed. An example is the sCoolMakers platform, which allows schools and other institutions to create, share, and download lessons on topics such as sustainability and circularity. These terms have gained an increase in popularity over the past decade, due to growing global awareness of environmental challenges, and the recognition of the need for long-term resource management.

The platform aims to address this need by allowing institutes to create and upload lessons that are tailored to their specific needs and preferences. This also means that similar or identical lessons may be uploaded by different institutes, causing the database to become cluttered and difficult to navigate through [Brown, 2014; Korhonen et al., 2018]. While customization and flexibility are valuable to institutes, it can also lead to the duplication of content on the platform and make it difficult for users to find the specific lessons they need. Brown [2014] states that digital education is a powerful tool for promoting sustainability but highlights the importance of designing educational resources

in a way that is engaging and accessible to learners. Institutes inadvertently creating almost identical lessons can lead to wasted resources [Brown, 2014].

This raises the critical question of how to prevent or circumvent this problem, to ensure that educational platforms remain useful resources for the users. This research aims to explore possible solutions by analysing literature and conducting a case study of the sCoolMakers platform. Our goal is to propose practical and effective strategies for managing lesson content on the platform, while maintaining the flexibility and customization that is valued by its users. We promote sustainability and circular economy in education by ensuring that educational resources are accessible and efficient. We also address the problem of data lifecycle management (DLM), which refers to the systematic management of data throughout its entire lifecycle, starting from its creation until its disposal. The data divided into various phases based on specific criteria progresses through these stages fulfilling many requirements or performs different tasks. An efficient DLM approach offers a well-organized and structured approach to managing a company's data, leading to improved data security and availability, among other key objectives [“Data Lifecycle Management | IBM”, n.d.; Marques, 2013]. DLM plays a critical role in the sustainable and secure operation of online platforms. As the database grows and potentially becomes flooded with similar lessons, it is increasingly challenging to manage and organize the content effectively [Marques, 2013; Zhenjia, 2019]. An efficient DLM strategy can provide a systematic approach to data management, ensuring that it remains well-organized and easily accessible for users. Moreover, data accuracy, consistency, and quality are paramount in the context of an educational platform like sCoolMakers. An effective DLM process helps maintain these qualities throughout the data lifecycle, ensuring that the lessons remain accurate, relevant, and effective.

In addition, sCoolMakers must comply with data protection regulations to mitigate the risks associated with data breaches or loss. As online platforms deal with sensitive educational data, ensuring data security and privacy is essential [Zhenjia, 2019]. By implementing a robust DLM strategy, sCoolMakers can streamline data management, enhance quality, and ensure security, which is crucial for long-term success. Therefore, DLM is of paramount importance for sCoolMakers, enabling the platform to effectively manage and secure its data while ensuring accuracy and relevance throughout the data lifecycle.

This paper will look into already existing literature on topics we thought useful, after which stakeholder interviews are discussed. Based on the outcomes of both researches, a framework is formulated and tested, and its limitations discussed.

II. THEORETICAL GROUNDING

We reviewed theories in three aspects: [1] Temporal databases and save-state mechanisms, and [2] data lifecycle management, and then identified knowledge gaps [3].

TEMPORAL DATABASES

A temporal database stores data with an explicit time component, allowing the management of changing data [Jensen and Snodgrass, 2009], dealing with historical or time-sensitive data. Temporal database management system [DMBS] requires temporal query and data manipulation language and temporal constraints [Steiner and Norrie, 1997]. The data of a temporal database is time-stamped. There are two prominent temporal aspects. The time references may capture either the database's past or the current states, leading to a transaction-time database; the states of the reality modelled by the data results in a valid-time database; or both aspects creating a bitemporal database [Jensen and Snodgrass, 2009].

Transaction-time databases record the timestamp of the transactions that modifies the data, then associates with the data changes resulted. Valid time denotes the period during a fact is true with respect to the real world. Transaction time is the period during a fact is stored in the database. These two time periods do not have to be the same for a single fact to be true [Steiner, 1998; Steiner and Norrie, 1997].

The two distinct concepts of time - valid time and transaction time - allow to differentiate various types of temporal databases [Steiner, 1998; Steiner and Norrie, 1997]. Historical databases store data in terms of valid time, whereas rollback databases store transaction time data. Bitemporal databases store both valid and transaction time data. Non-temporal DBMS stores only one state of the real world, typically the most recent one. Such databases are commonly referred to as snapshot databases, see figure 1 [Steiner and Norrie, 1997].

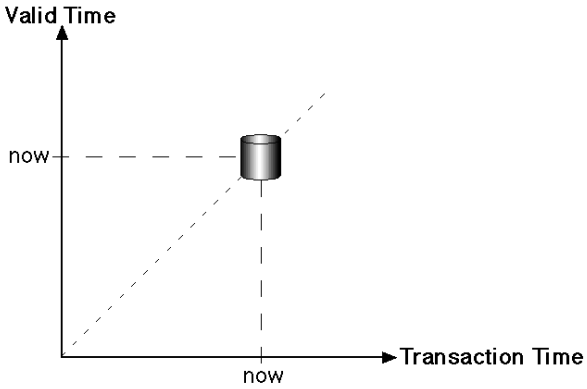

Figure 1. Depiction of a snapshot database

Bitemporal DBMS preserve data history both valid time and transaction time. The history of saved data [transaction time] is restricted to past and present database states because it is controlled directly by the system, which has no knowledge of future states. A table in a bitemporal relational DBMS can be a snapshot table [which stores only current data], a valid-time table [which stores data when is valid in relation to reality], a transaction-time table [which stores data when recorded], or a bitemporal table [stores both valid and transaction time] [Steiner, 1998; Steiner and Norrie, 1997]. When creating a table, a more complex version of SQL enables users to define the required type of table. Existing tables can also be modified. It accommodates temporal inquiries, temporal modification statements, and temporal constraints [Steiner, 1998; Steiner and Norrie, 1997].

Figure 2 [Steiner and Norrie, 1997] depicts the states saved in a bitemporal database. Of course, a temporal DBMS does not maintain each database state independently. It saves valid time and/or transaction time for each tuple separately.

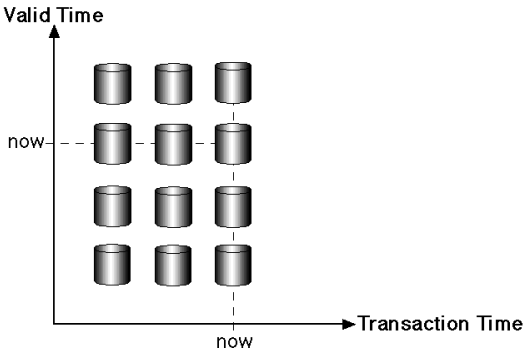

Figure 2. Depiction of states stored in a bitemporal database

Transaction-time temporal databases may also store information about the transaction itself, such as the user executed it or the reason for change [Jensen and Snodgrass, 2009]. This is useful for many reasons. They allow the retrieval of past versions of data, which is helpful for auditing, debugging, or analysing changes to data over time [Steiner and Norrie, 1997; Jensen and Snodgrass, 2009]. They ensure data integrity by providing a complete history of changes to data and allowing the identification of changes made in error [Steiner and Norrie, 1997; Lorie, 1977]. They can track the performance of transactions and identify potential bottlenecks or inefficiencies in a database system [Jensen and Snodgrass, 2009].

In the context of the sCoolMakers platform, temporal databases may be useful to provide a means to store and manage data that changes over time, like the stored educational classes. They could track changes made to the lessons by different schools by implementing a temporal database, ensuring that it contains a complete history of all versions of each lesson. This allows users to access and compare different lesson versions without creating multiple copies of the same content, which would clog the database and make it difficult to use. [Jensen and Snodgrass, 2009].

However, implementing a transaction-time database will not solve all the issues. To address the problem of similar contents, developing a save-state mechanism can be a solution [Lorie, 1977]. It ensures that the database remains in a consistent state by detecting and correcting inconsistencies in real-time [Lorie, 1977]. It prevents the flooding of identical content by maintaining a transaction-time database that captures past and current states [Lorie, 1977]. Whenever a new lesson is added or modified, the save-state mechanism can be triggered to detect and correct inconsistencies in the database [Jensen and Snodgrass, 2009].

DATA LIFECYCLE MANAGEMENT

Data - information, facts, statistics, or figures that can be collected, stored, and analysed to gain insights and make informed decisions - is an essential component of our modern society. Data collection is the systematic process of gathering observations or measurements. Data is used in a variety of fields, including business, healthcare, finance, and education, to name a few. It can assist organizations in making more informed decisions, understanding trends, identifying patterns, and solving problems.

As data is becoming increasingly important in many industries, so does managing them. In higher education ecosystems, it is essential that the data within and its life cycle is properly managed [Sant-Geronikolou, 2018]. As libraries are abandoning their old norms and traditional structures, new ideas and breakthroughs are required to create new practices [Anglada, 2014]: building accurate knowledge about user behaviour for improving existing services or create new ones, reshaping data collection procedures in libraries to provide rigorous, transparent, and reliable data for efficient and trustworthy service evaluation; and investigating novel approaches to capitalizing on these procedures within a broader educational context [Sant-Geronikolou, 2018].

Content sustainability proper data lifecycle management [DLM], the process of managing data throughout its lifecycle, including its creation, storage, maintenance, use, sharing, and eventual disposal - is vital in new libraries. It ensures high quality data, managed and protected properly, and complies with legal and ethical requirements. Effective DLM necessitates the use of appropriate technologies, policies, and procedures, and the participation of stakeholders throughout the data lifecycle. DLM should either be a shared responsibility between both users and hosts, or could potentially be outsourced, depending on the context of the entity.

Data is a valuable asset that organizations rely on to make informed decisions, so ensuring its sustainability is critical. Sustainable data management practices help ensure that data is accurate, reliable, and trustworthy, reusable and effectively shared. Long-term DM practices reduce the risks of data breaches, data loss and data misuse.

Libraries' sustainability is assessed using the relationship established between the values added and their costs. Anglada [2014] analyses this relationship and proposed a ‘sustainability formula’, seen in figure 3, to figure out whether the data that is stored sustainably or not.

$$\text{Sustainability} = \frac{\text{Value}}{\text{Cost}} = \frac{\text{Use} - \text{Dysfunctions} + (2 * \text{Perception})}{\text{Cost}}$$

Figure 3. Anglada’s Sustainability formula

Two of the formula's elements, costs and use, are self-explanatory, but including 'perception' and the weight assigned to it must be justified. Anglada [2014] states that cognitive biases or prejudices are not always grounded. Dysfunctions are ‘unfulfilled expectations’ [Anglada, 2014]; a gap between expectations and reality.

The formula originally supported the argument of physical libraries replaced by their digital counterparts [Anglada, 2014; Sant- Geronikolou, 2018] but can be applied to their contents, too. Applying the first half of the formula to data, we gain new insights on data sustainability, and thus essentiality to remain in the database. Considering the relationship between the value of the data (e.g. what unique information does the data store) and the cost of the data (e.g. size of the data in regards to storage waste) is vital.

The valuable data asset of online educational platforms like sCoolMakers, create, deliver, and improve educational content. Because of this, we believe these platforms can be seen as their own libraries, full of content that should adhere to the same formula. Implementing effective DLM practices ensure the platform's success and address challenges of maintaining data flexibility, ownership, and structural integrity.

sCoolMakers manage data at various stages of its lifecycle. It creates educational content, stores data, maintains, shares and updates them with other parties, and disposes them when outdated or irrelevant. Each stage must be managed properly to ensure data availability, quality, and compliance with relevant regulations and guidelines.

KNOWLEDGE GAPS

There are no explicit knowledge gaps related to strategies for preventing redundant content on the sCoolMakers platform based on the text on temporal databases. The text primarily focuses on the concept and characteristics of temporal databases, including valid time and transaction time definitions and temporal database types. However, one potential knowledge gap related to the use of temporal databases could be to investigate how temporal databases can be used to manage the lifecycle of educational content on the platform. How to track changes in educational content over time and ensure that outdated or redundant content is removed from the platform to avoid duplication, or how temporal databases can be used to analyse educational content usage on the platform and inform decisions about content prioritization for updates or revisions.

Incorporating a temporal database into the platform can be useful for tracking changes of educational lessons over time. It can ensure data integrity; assist detect errors and improve database performance. A save-state mechanism, however, can be developed and integrated into the database to avoid the flooding of similar content. This can detect and correct inconsistencies in real-time, ensuring that the database maintains consistency while capturing changes made by various schools. However, implementing these mechanisms require responsibility to work for sCoolMakers’ and the end users’ needs.

Conducting exploratory interviews with experts in [temporal] database design is key to solving these knowledge gaps.

III. METHODS

HEVNER DESIGN SCIENCE RESEARCH MODEL

Hevner model [Hevner et al., 2004; Hevner, 2007; Hevner and Chatterjee, 2010] is a framework for designing and evaluating information systems research. It emphasizes the importance of the design science paradigm, which involves iterative cycles of designing, implementing, and evaluating information systems. It is applied in a variety of contexts, including e-learning, healthcare, and supply chain management [Hevner and Chatterjee, 2010; Maiztegui, 2023]. Its main objective is to create knowledge that experts in the relevant field can use to create solutions for practical problems. They concentrate on the decision-making process for feasible and useful construction of potential futures. Hevner's design science research cycle [Hevner and Chatterjee, 2010] goes through three different cycles, giving unique insights for creating and designing an artifact as shown in figure 4 [Hevner, 2007].

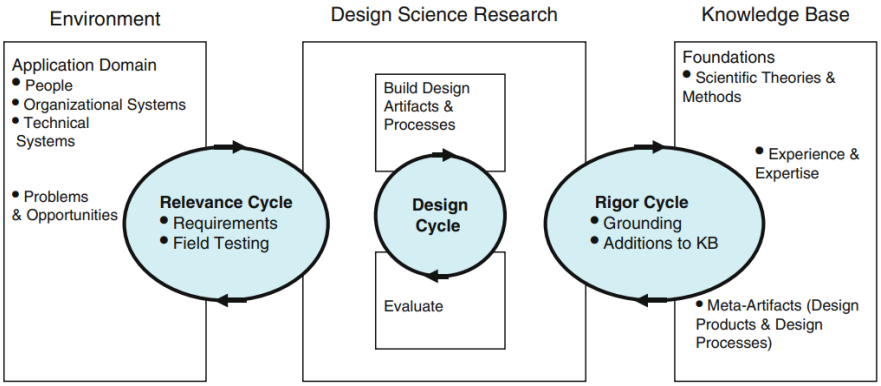

Figure 4. Design science research cycles

Using this IS research framework, we incorporate a focus on three essential research cycles. The *Relevance Cycle* serves as a link between the design science activities and the context of the research project. The *Rigor Cycle* acts as a link between the design science activities and the body of knowledge, experience, and expertise that underpins the research project. The main activity of creating and assessing design artifacts and research processes is iterated upon during the framework's central *Design Cycle*.

The Hevner model is used to guide the development of instructional design models and to evaluate the effectiveness of e-learning interventions [Maiztegui, 2023]. The design science paradigm is particularly relevant in the context of e-learning, where the design of effective online educational resources requires iterative cycles of design, implementation, and evaluation [Hevner and Chatterjee, 2010; Maiztegui, 2023].

For practical applications, the model is used to guide the development of strategies for managing online contents, such as sCoolMakers. The iterative cycles of designing and evaluating its database with a focus on data lifecycle management can provide a framework for developing and testing strategies for preventing identical or redundant content, still allowing the customization of lessons to meet the needs of users. This way, we can potentially improve the quality and accessibility of sustainability education.

This method was used in the cycles below, where firstly stakeholder interviews were conducted (environment), and together with the theoretical grounding (knowledge base), a framework was designed.

IV. FIRST DESIGN CYCLE

HISTORY

The origins and problems of the sCoolMakers platform influenced our design choices. The first interview with a key figure in the development of its database emphasized the platform's mission. "*Our primary goal is to provide schools with a user-friendly platform that facilitates knowledge sharing on sustainability and circularity*." The dedication to empower sustainable education has been the driving force of the evolution of sCoolMakers.

One of the key strengths of sCoolMakers lies in its flexibility, allowing institutes to customize lessons according to their unique needs and preferences. "*We wanted to offer institutes the flexibility to tailor the lessons according to their specific needs and preferences*." This approach empowers schools to deliver content that is relevant to their students and addresses their specific curriculum requirements. This customization causes content duplication. Multiple versions of similar or identical lessons uploaded by different institutes, clutter the database, hindering users' ability to find the specific content they seek.

sCoolMakers uses Bubble.io, a no-code development software that allows building diverse applications without the use of traditional coding. The default structure and functionality of Bubble's databases did not perfectly correspond with the requirements of sCoolMakers. "*We needed a more robust system to handle the customization and avoid content duplication*.", indicating the need for a more powerful mechanism to handle modification and eliminate duplicate content. Prewritten codes and styles make users relying on Bubble software and the fully customized program runs into issues. Despite the obstacles, sCoolMakers remains committed to providing a unified user experience. The developer stated, "We are actively exploring potential solutions to improve the database and eliminate content duplication." sCoolMakers hopes to improve the platform's user experience by addressing this issue, saving time and providing simple access to relevant content.

Bubble.io appears to be a suitable software for building a small application that may only display information or have a minimal UI, especially for those who lack skills in programming. Bubble's simplicity and instructions are helpful, but when the goal is to construct a larger, more complicated, customizable application, it is advisable to reevaluate its suitability and explore more bespoke alternatives.

PROLIFERATION OF IDENTICAL CONTENT

The second interviewee, a database architect, suggested to tackle the problem by creating a ‘Make your own menu’ restaurant type database, a complex multi-layered application.

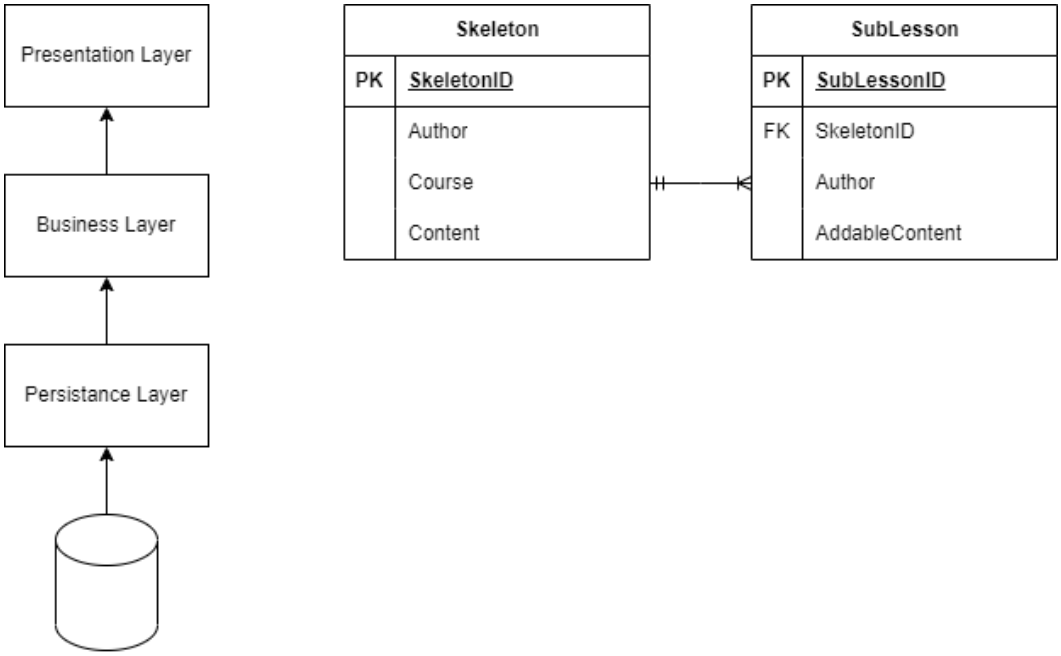

Figure 5. Construction of the ‘Make your own menu’-restaurant type database

Its hefty disadvantage is the lack of dynamic freedom. “*It forces a structure that may work well in certain cases but adding 'addable content' may not solve the problem. We don't know how this content will look, or the created menus solve things in a dynamic way. It is a 'solution', but not a practical one*.”

Due to this major risk of losing flexibility, ideating other solutions, DBMS could make use of reflective programming, a technique that allows a code to manipulate itself during runtime, flexibility and less code duplication. A non-relational DBMS is a more complex solution and infinitely expandable.”

By storing it in a format such as XML, JSON, TOML, or, if necessary, a document-oriented database, the problem can be approached in a more flexible manner. Data can be structured on a case-by-case basis, with a general overarching structure established by using fixed tags within an XML specification. This structured data can then be processed in the business layer to create a user-friendly product. “*Adopting this approach allows for the easy assembly of lessons through an internal conversion process that can rely on POCO, POJO, POPO, PORO, etc*.” This approach offers the advantage of enabling the implementation of elegant solutions using Design Patterns and Reflection [Reflective Programming], which can subsequently be passed to a presentation layer.

Because the data is stored and structured in a more flexible matter with adding tags, it is possible to add and store tags together. This, in turn, can create one ‘older’ file per lesson with a bonus of creating a lesson and a reference or path to its parent file. Extra lessons can be created, which is done by writing additional parsers within the business layer.

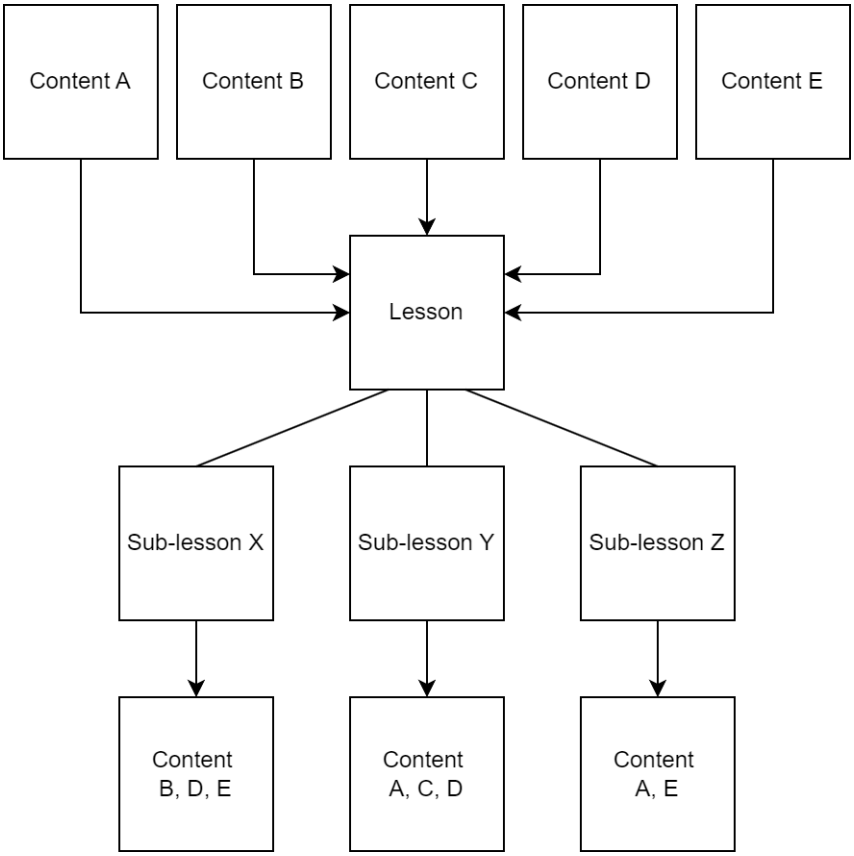

Figure 6. Summarizing diagram of the suggestion

For each sCoolMakers topic there is a main lesson, acting as a skeleton: a bareboned shell with a basic content. The remainder of the main lesson consists of different pieces of content, created by either the creator of the main lesson or the corresponding sub-lessons. These blocks need labels and tags. If other parties want to use the lesson or its pieces, they can customize and build upon the original, whilst the main lesson stays the same. These sub-lessons will exist of only specific pieces of content of the main lesson and can add new pieces. Instead of changing the original content block, it would result in the creation of a new one. This way, the main lesson and the sub-lessons are independent.

The problem of this solution is the dilemma how to balance the scales of content. The platform could allow practically infinite amount of options for users, where everyone can customize their own lessons into the finest of details. However, this would result in an abundance of data points, which would be near impossible to maintain [with current technologies]. To resolve this problem, the platform needs to limit the amount of data points the user can create and access. This would of course result in less flexibility for the user but an important balance to figure. Does sCoolMakers want more flexibility at the cost of data management, or would they rather lessen the dynamic experience to minimize the management of data?

DATA MANAGEMENT

To determine which way the scales of balance tip, it is critical to gather additional understanding on data management of online education. The next interview focused on the question how to effectively manage the retention and deletion of outdated content while ensuring that relevant content is maintained.

Given how important flexibility is to sCoolMakers, limiting the number of data points users can use seems contradictory. "*Finding the right balance is key. By defining specific criteria for identifying outdated content, such as relevance, accuracy, and*

curriculum changes, [sCoolMakers] can establish guidelines for whether content is still relevant. Once content no longer meets these criteria, it should be flagged for review or deletion."

Creating the criteria and acting on them are difficult tasks, especially when dealing with massive amounts of data. A suggested automation through artificial intelligence algorithms or machine learning models can assist identifying information that requires inspection based on the criteria. Algorithms can assess the relevance and quality of material by analysing characteristics such as content usage, user comments, and metadata. This would aid in identifying and prioritizing older materials for review or removal.

Another approach was suggested, one which ensures that sCoolMakers' valuable and up-to-date content is preserved while outdated content is appropriately managed. It's essential to engage the platform's user community in this process. "*Encouraging users to flag outdated or inaccurate content, providing a reporting mechanism, or implementing a user-driven content review system can be effective*." By implementing a community-driven moderation system, the criteria of whether data is still relevant or not are put up by the user's. Additionally, sCoolMakers could establish a team of moderators or subject matter experts who regularly review and curate the content. This approach would help maintain the platform's integrity and ensure that relevant and reliable content is preserved. This method could be properly deployed if the amount of data is not overwhelming.

sCoolMakers must find a balance between allowing maximal flexibility to users and ensuring effective data management. "*One approach to addressing this challenge would be to limit the number of data points users can create and access. By establishing realistic restrictions, the platform may save a manageable amount of data while still allowing users to significantly tailor their lessons*." This strategy recognizes the significance of data management while still allowing users some flexibility. A moderation crew could handle all the content, while the platform controls the amount of data.

It is ultimately up to sCoolMakers' developers and administrators to determine the optimum balance of flexibility and data control. They must measure the benefits of personalization against the challenges of data management and examine the trade-offs. Finding a solution that fits the demands of both users and the platform itself is critical. By striking this balance, sCoolMakers will be able to provide a dynamic and personalized experience while maintaining good data management within the present technology restrictions.

Another major aspect raised during the interview was the matter of ensuring compliance with data privacy standards and preserving user data throughout the DLM process. "*Data privacy and security are paramount in today's digital landscape*." It was suggested that sCoolMakers employ strong data protection measures, such as encryption of sensitive information, role-based access controls, and frequent audits to ensure compliance with relevant data protection standards. sCoolMakers should establish clear and transparent policies surrounding data retention and deletion, and seek for informed consent from users, giving them control over their data.

NEW CONCEPTUAL FRAMEWORK

sCoolMakers provides a user-friendly platform for knowledge sharing on sustainability and circularity in schools. Currently it uses a less robust system. Switching to a more powerful one is required to handle customization, eliminate content duplication, and provide greater dynamic freedom. Improving the database will also enhance the user experience, save time and simplify access to relevant content.

A main lesson for each topic should contain essential facts and necessary information with labelled and tagged content blocks. By associating the main lesson with the content, users are free to create and add their own lessons to it. Users could tailor lessons according to their specific needs and preferences, empowering them to deliver relevant content aligned with their curriculum.

Finding a balance between flexibility and data management, sCoolMakers should establish specific criteria, such as relevance, accuracy, and curriculum changes, to identify

outdated content. Content that no longer meets the criteria should be flagged for review or deletion with user-driven content reviews. Users should be encouraged to flag outdated or inaccurate content, report issues, or participate in a user-driven content review system. The criteria for content relevance can be determined by the active community of users with their input playing a significant role.

However, merely managing the existing content will not be enough to manage the data effectively in the long run. sCoolMakers should apply some realistic restrictions on the number of data points users can create and access.

V. SECOND DESIGN CYCLE

For the second design cycle, the new conceptual framework was verified with another expert, who has had years of working experience with content management and data engineering.

CONFIRMATORY INTERVIEW

The confirmatory interview was positive with the proposed system. It was mentioned that there is a need for a more robust system to handle customization, eliminate content duplication, and provide greater dynamic freedom. The idea of a non-relational DBMS combined with reflective programming and a flexible storage format was found to be a viable option. *"It allows for a more adaptable and expandable approach while minimizing data duplication.*" However, further consideration and evaluation may be necessary to assess the feasibility and implementation details of this solution.

The main critique was on the data management section of the framework. Merely allowing users to manage the data will not be an effective strategy in the long run. "*In general, you lack a clear data governance strategy. You mention allowing the users to manage your data for you, but I do not think this is a scalable solution*." Furthermore, there is currently an absence of discussion on data analytics and insights. How will it be decided what data is truly not relevant anymore? "*Data quality management involves processes such as data cleansing, validation, and monitoring to ensure the accuracy, completeness, and consistency of data.*"

We were given plenty of suggestions on how to circumvent some of the difficulties, ensure effective data management, and recommended to develop a comprehensive data governance framework that outlines data quality standards, data ownership, roles, and responsibilities. Additionally, implementing data stewardship programs can help foster data accountability and enforce data governance policies. There should be clear limits to the amount of data. To ensure the accuracy and completeness of incoming data, they recommended to implement data validation rules and checks during data ingestion. Regularly conducting data quality audits and performing data cleansing activities can help identify and rectify any data inconsistencies. In addition to this, proper data integration architecture should be implemented. "*Steps such as the cleaning, filtering, transforming and allocating of the right data are all important to data integration. Especially if you are working with content made by users themselves, how will you make sure all users submit the same quality of data?*"

There should be a focus on data privacy. "*As you deal with user-generated content and potentially sensitive data, it is essential to have a strong focus on data privacy and security.*" There are many difficulties with data privacy, such as how user data will be protected, secured, and managed to comply with relevant data protection regulations. To protect sensitive user data, it is critical to develop and enforce strong data privacy policies and procedures that comply with applicable regulations. Managing this would ensure that data handling and processing adhere to legal requirements and best practices. "*Implementing strict access controls, encryption techniques, and data anonymization methods should improve the security of sensitive user data. These safeguards can help to prevent unauthorized access and safeguard individuals' privacy*." Also, conducting regular security audits and penetration testing will allow for a proactive identification of

vulnerabilities in the data management infrastructure. Potential security risks can thus be detected and addressed quickly, ensuring a high level of data security.

FINAL CONCEPTUAL FRAMEWORK

Based on these inputs, we expect a solid advice for sCoolMakers and potential other platforms who are facing the same issues.

For sCoolMakers' DBMS, as suggested, each topic should be accompanied by a main lesson that serves as a foundational structure with limited content. The existing content in the main lesson is accurate and relevant. The remainder of the main lesson is made up of various content blocks that can be created by either the creator of the main lesson or the creators of the corresponding sub-lessons. These content blocks must be labelled and tagged appropriately. Other parties can customize and expand on the original lesson while leaving the main lesson unchanged. These sub-lessons include specific content blocks from the main lesson as well as new content. Instead of changing the original content block, changing it would result in the creation of a new one. This method ensures that the main lesson and sub-lessons remain distinct from one another. Because the sub-lessons are built with specific content blocks, even if the main lesson changes, the content blocks within the sub-lessons remain unaffected.

A proper data governance system, where both the integration and quality management of data is warranted, establishing a suitable data management lifecycle are necessary to map out a series of well-defined steps, like appropriate guidelines for quality standards, data ownership and responsibilities, an effective data integration architecture, scalability and privacy policies with procedures.

VI. DISCUSSION

The conceptual framework developed in this study provides useful insights and recommendations for sCoolMakers and other educational platforms looking to improve their content management and data governance. Platforms can improve customization, ensure data quality, and prioritize data privacy by implementing the suggested approaches, ultimately providing a better user experience and supporting the mission of knowledge sharing on sustainability and circularity in schools.

The design science research conducted in this research has more applications than that of ScoolMakers and can be applied to a wider audience. For example, information system educators and researchers could use this research design as a way of showcasing the importance of stakeholder engagement, and to identify gaps within artifacts.

It can close the education gap and increase access to high-quality educational resources. The platform's ability to customize and localize content allows it to meet a wide range of educational needs, including those of poverty-stricken communities or regions with limited educational resources. This can help to reduce educational disparities and promote equal opportunities for all students.

The proposed platform improves the learning experience and empowers individuals to become lifelong learners by encouraging educators and students to actively contribute and share their expertise. It can foster a growth mindset and a lifelong pursuit of knowledge, skills, and personal development outside of formal education.

Educational platforms foster digital literacy and 21st-century skills among educators and students. Users gain digital competency, information literacy, critical thinking, and problem-solving skills by creating and customizing content. These abilities are critical for thriving in the digital age and preparing individuals for the demands of today's workforce.

We emphasise collaboration and knowledge sharing aids in the development of a collaborative culture within the education community and foster a sense of collective intelligence and shared ownership of knowledge by encouraging educators and students to contribute and build upon existing educational content. Collaborative culture encourages

cooperation, empathy, and a sense of community, ultimately improving the learning experience.

Customization and expansion of educational content within the proposed framework encourages educators and students to be innovative and creative. The platform encourages creative teaching approaches and allows learners to express their ideas and insights in unique ways by allowing them to adapt lessons and content blocks. This creates an environment that encourages creativity, problem-solving, and the exploration of alternative learning paths.

One key part both we and the creators behind sCoolMakers find essential, is that the use of educational platforms crosses borders, allowing for global knowledge exchange and cross-cultural learning. Educators and students from all over the world can con- tribute, access, and benefit from a diverse range of educational content by leveraging the framework's capabilities. This fosters cultural understanding, intercultural communication, and appreciation for diverse viewpoints, resulting in a more interconnected and globalized society.

Our research bridges the gap between theoretical concepts and practical application in educational technology and using the Hevner model, the study provides a solid theoretical foundation by grounding the research in existing literature and theories on content management, data governance, and educational design. It provides practical recommendations and guidelines for implementing the proposed framework by incorporating real-world challenges and stakeholder perspectives.

The scientific impact of this study is that it advances educational technology research, bridges theory and practice, informs best practices in platform development, and opens avenues for future research. This study promotes evidence-based decision-making, fosters interdisciplinary collaborations, and drives innovation in the field of educational technology by contributing to the scientific knowledge base.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisors, Toon Abcouwer and Emöke Takács, for their unwavering support and guidance throughout this research journey. Their expertise and insights have been invaluable in shaping the direction of this thesis.

I am also deeply thankful to Emöke for her exceptional assistance in streamlining and condensing the content of this thesis. Her dedication and keen editorial eye have played a pivotal role in presenting this work more effectively.

REFERENCES

Anglada, L. (2014) “Are libraries sustainable in a world of free, networked, digital information?”, Profesional De La Informacion. 23 (6), 603-611

Brown, S. J. (2014) “Conceptualizing digital literacies and digital ethics for sustainability education”, International Journal of Sustainability in Higher Education, 15 (3), 280–290

Hevner, A. R. (2007) “A Three Cycle View of Design Science Research”, Scandinavian Journal of Information Systems, 19(2), 87–92

Hevner, A. R., and Chatterjee, S. (2010) “Design Research in Information Systems”, Springer Nature, https://doi.org/10.1007/978-1-4419-5653-8

Hevner, A. R. et al. (2004) “Design Science in Information Systems Research”, Management Information Systems Quarterly, 28(1), 75

Jensen, C. S. and Snodgrass, R. T. (2009) “Temporal Database”, Springer nature, pp. 2957–2960, https://doi.org/10.1007/978-0-387-39940-9_395

Lorie, R. A. (1977) “Physical integrity in a large segmented database”, ACM Transactions on Database Systems, 2(1), 91–104

Maiztegui, F. G. (2023) “Research Design”, Gabler Theses, pp. 7–15, https://doi.org/10.1007/978-3-658-40731-5_2

Sant-Geronikolou, S. (2018) “Understanding in-library use data lifecycle within Greek and Spanish higher education ecosystems”, Library Hi Tech News

Steiner, A. (1998) “A generalisation approach to temporal data models and their implementations”, https://doi.org/10.3929/ethz-a-001923958

Steiner, A. and Norrie, M. C. (1997) “A temporal extension to a generic object data model”, CTIT technical reports series, 265

ABOUT THE AUTHOR

Jonathan Hombroek is a Master’s student at the University of Amsterdam. He has a Bachelor’s degree in Information Sciences and is now pursuing a Master’s degree in Data Science, both at the University of Amsterdam. His interests lie in the field of Information Systems and Machine Learning.