Designing and Developing a Data Science Programme in Bhutan

Phub Namgay
Department of Informatics and Media
Uppsala University
Uppsala, Sweden
phub.namgay@im.uu.se

Pema Wangdi
Department of Mathematical Sciences
Sherubtse College, Royal University of Bhutan
Kanglung, Bhutan
pemawangdi.sherubtse@rub.edu.bt

Sangay Thinley
Department of Mathematical Sciences
Sherubtse College, Royal University of Bhutan
Kanglung, Bhutan
sangaythinley.sherubtse@rub.edu.bt

***Abstract*—Academic programme development primarily focuses on a product rather than a process. Thus, the dynamics among stakeholders and processes while designing and developing a programme receive little attention and are rarely studied. This study discusses the process of developing a four-year undergraduate data science programme at Sherubtse College, Royal University of Bhutan, through the analytic lens of routine dynamics. Routines are recurrent, recognisable action patterns enacted by different interdependent actors to accomplish a task. A need assessment survey was conducted to garner insights and perspectives of stakeholders on data science. Qualitative coding was employed to analyse and synthesise the survey responses to understand the needs and challenges of data science. Drawing on the survey responses, constructive recommendations from the stakeholders, and insights on data science, the programme committee successfully developed an academic programme that caters to the data science needs and demands in Bhutan. The programme was launched in the autumn academic session of 2021 with 29 students, guided by the aim of equipping students with sound theoretical knowledge and practical skills to perform complex data analytics and operationalise data technologies. This study contributes to the knowledge base of data science, namely the use of routine dynamics as an analytical instrument to study the underlying workflow processes and information exchanges in an undertaking. The methodical description of a programme development process by anchoring on an analytic framework offers valuable information for curriculum developers mandated to produce graduates with know-how in dealing with data of varying volume and variety.**

Keywords*—*Data science, Programme development, Routine dynamics, Bhutan

I. Introduction

Data science is an interdisciplinary field defined succinctly as the "science of data" or "study of data" [1]. From an ontological perspective, data are symbols that represent the properties of objects and events [2]. It is the *rawest* information from which knowledge and insights are derived [3]. Datafication phenomenon, defined as taking many aspects of life and converting them into data to be realised as a new form of value [4], has given rise to the rapid adoption of information systems and data technologies to conduct everyday work and life. It further amplifies the touting of data as 'a new oil' [5] or 'new Intel inside' [6]. A number of emerging technologies such as machine learning, artificial intelligence, augmented reality, the Internet of Things, and digital twins enable the current society referred to as 'Society 5.0' [7] and the new industrial paradigm referred to as Industry 4.0 [8]. An enormous volume of data generated at a great velocity in different structural formats is a chief element that enables the new societal or industrial paradigm. Thus, the need to equip people with the know-how to work with small data or big data has necessitated academia across the globe to offer a data science programme [8, 9].

With the high obsolescence rate of contemporary technologies and short shelf life of knowledge and skills, it is imperative for people who deal with or are going to deal with data and related technologies to be furnished with adequate knowledge and skills that are current with the demands of Society 5.0 and Industry 4.0. It is only possible through adequate practical skills supported by strong theoretical foundations in data science. The Royal University of Bhutan (hereafter, the 'University') thought it was an opportune time to offer a data science programme, given the growing demand for graduates with expertise and competency in dealing with data of varying magnitude. Nevertheless, numerous stage-gating processes must be fulfilled before an academic programme can be designed, developed, and offered to prospective students.

The research data stems from Bhutan, a tiny country between China and India, as in Fig 1. Bhutan is relatively a latecomer in the global digital ecosystem. Nonetheless, considerable progress has been made to actively contribute to the data revolution in the last few decades. Bhutan has witnessed rapid datafication in every aspect of society. The e-government policy has 'Digital by default' and 'Single source of truth' as its policy statement [10] to deliver digital public services driven by accurate and consistent data. Moreover, the Bhutanese government strives for a society driven by smart technologies, such as online government to citizen services, to supplement conventional analogue services across all sectors.

Fig 1. Geographical location of Bhutan

Similarly, private enterprises in Bhutan are also increasingly aligning their workflow with the Industry 4.0 paradigm to optimise business. Thus, a torrent of heterogeneous data is generated every day to drive technologies and services. A shift in the workflow alludes to the requirement for people with adequate know-how to harness raw business data for social and economic value. The risk of the data generated by the public sector or private firms ending up as 'dark data' [11] is heightened if tools, methods, and mentalities for exploiting data value are scant. In light of the rapid data-driven developments in the public and private sector, the time is ripe to offer a programme at the University that would address, solve, or cater to the needs and challenges of data science in Bhutan.

At the University, the programme development process has core patterns of action that are virtually recurrent whenever there is a proposal to offer a new programme. The process is underpinned by the Wheel of Academic Law—a definitive set of policies and regulations that govern the University's academic matters [12]. The patterns of actions enacted by various actors in scrutinising the proposal for a new academic programme to an eventual launch are recurrent and recognisable, involving multiple participants with interdependent actions [13]. In the current research, we are interested in the dynamics of the processes involved in developing an academic programme and how the performance of different actors through the enactment of patterns of action [14] has implications for translating a proposal into an eventual offering. Hence, routine dynamics (hereafter, the 'routines') [13, 15] is an appropriate theoretical framework to expound on designing and developing a curriculum in academia. Routines are recurrent, recognisable action patterns enacted by different interdependent sociotechnical actors across space and time to accomplish a task [13]. To direct, structure, and present our work, the overarching research question that underpins the study is:

RQ: How can we use an analytic framework to capture the process of designing and developing an educational programme in academia?

The research objectives are to analyse an academic programme development process through an analytical lens; document good practices that emerged in designing and developing a programme; and share contextual insights and experiences in designing and developing a programme in academia.

This paper is a story of our endeavour to design, develop, and offer a data science programme through a qualitative study of the programme development process anchored on the analytic lens of routines. The word 'we' or 'our' refers to the programme development committee (hereafter, the 'programme committee') of the Department of Mathematical Sciences at Sherubtse College. Every facet of the workflow in designing and developing the programme was analysed and synthesised using routines as an analytical lens. Subsequently, we graphically mapped the programme development process using a flow chart diagram incorporating the concepts and constructs of routines, as in Fig 2. The schematic diagram captures the ostensive and performative aspects of different actors involved in a programme development process. Moreover, the granular illustration of aspects of recurrent action patterns that emerge from different actors [15, 16] evokes the link between the macro-level dynamics of a programme development process and the micro-level action of various actors at the University.

After working on the project for almost one-and-half years, the programme committee successfully developed the Definitive Programme Document (hereafter, the 'programme document') of the data science programme, which took into account the objective issues and concerns of stakeholders. The four-year undergraduate data science programme was launched in the autumn academic session of 2021 with 29 students—21 boys and 8 girls. The first batch will graduate in spring 2025. So far, the programme is functioning smoothly, and the performance of the first batch of students in their first semester was good—26 out of 29 progressed to the next semester. The programme's overall efficacy remains to be studied. The exposition of an academic programme development process by anchoring on routines contributes to our understanding of the routine processes from a different analytical framework. Moreover, the study demonstrates the applicability of routines as an instrument for studying the processes in a workflow. This study also contributes to the literature on curriculum development in data science, namely understanding aspects of curriculum development through an analytical framework. The contextual insights and experiences presented in this study will be helpful for academic institutions with a plan or those mandated to offer data science or related programmes.

The paper is structured as follows. First, prior literature is presented, followed by a description of the study design and methods. Then, the analyses and findings of the research are described. After that, a discussion of the results, contributions to knowledge, and limitations of the study are presented. Finally, the paper ends with the conclusion of the study.

II. Literature Review

Data science is a field that requires skills from diverse fields such as computer science, mathematics, statistics, information sciences, and arts [6, 17], thus demonstrating the interdisciplinarity of the field [1, 18]. It is a relatively new academic field [9]. The interdisciplinarity of data science is captured in the following discipline-based formula by Cao [1] to elucidate the definition further, where '|' in the formula means 'conditional on':

Data Science = Statistics + Informatics + Computing + Communication + Sociology + Management | Data + Environment + Thinking

The formula resonates with Tang and Sae-Lim [9] that data science should represent a combination of subject areas from multiple disciplines. Data product, that is, deliverable from data, is the ultimate output of data science [1, 6]. It comes in various forms, such as value from data analytics to drive machine learning models or actionable insights for business optimisation.

Datafication has implications in the current data ecosystem as organisations and people increasingly attempt to make sense of their data [19]. Hence, it has become an essential discourse in data science. It is worth noting that datafication is more than converting analogue objects into digital data [4, 20]. For instance, Zoom datafied formal meetings and Twitter datafied stray thoughts into 280 characters. A skill imperative to deal with data pouring from every direction in a datafied society is the ability to analyse and synthesise vast amounts of data [9]. Big data analytics, for instance, helps gain insights into citizens' online footprint to introduce new digital services or optimise existing services. Datafying work and life also result in the fast growth of data from every dimension—volume, velocity, variety, veracity, value, and the like—thus needing people with adequate proficiency to gather, process, and analyse big, unstructured data [9, 21]. This has necessitated some academic institutes to focus on developing analytical skills in their data science programme [9].

At a fundamental level, data science aims to extract actionable insights from everyday organisational data for business optimisation, direct monetisation, or research and development. Rich theoretical knowledge and technical skills in data science to leverage data of varying magnitude is a competitive advantage. In an era of a torrent of data, one pertinent question we should ask is not 'what data to store, but what we can do with the data' [5]. Marz and Warren [3] highlight an unanticipated value that inheres within every large dataset. In a data context, value can be the material or monetary worth of data or data deliverables like analytic insights. Nevertheless, how an individual or organisation perceives value depends on their strategic goals for exploiting or operationalising data [22]. Data does not equate to value and needs systematic analytics to translate raw data into value [23]. It resonates with Monino [5] that data in itself is not power; it is using that gives power to data. There is no one-size-fits-all method for optimally harnessing value from data [24]. Organisations, especially profit-driven ones, need to process data on time for decisions and actions [25]; otherwise, data value will gradually decline, often referred to as '*Decay*'. 'Decay' is now seen as a critical dimension of big data [26].

The data dimension —volume, velocity, variety, and the like—significantly increases as society gravitates toward Society 5.0 and Industry 4.0 paradigms. People who can perform complex data analytics [27] and establish a data analytics culture [8] are vital to tapping knowledge and insights from data. The exploitation of socioeconomic value that inheres in data requires leveraging fundamental theories (e.g. the CAP theorem for distributed data store [3]) and practical technologies (e.g. Hadoop, MapReduce, and Lambda Architecture [3]) of data science. Public organisations and private firms are leveraging the data revolution by racing to extract optimum value from their enterprise data to drive decisions and actions [28], as informed decisions driven by data contribute significantly and positively to organisational performance [27]. The ability to capitalise on data science to gain a competitive advantage is an edge in a competitive climate [18]. Hence, data analytics skill is sought after across organisations overwhelmed with data to avoid an 'information paradox' situation—drowning in information but starving for knowledge [29].

The components of data science such as big data and complex computational analytics are disruptive innovations that reconfigure society. As data science takes centre stage in this era of data deluge [8], there is a shift in perception among people who are increasingly aware of the significance of data as a critical strategic asset [5], not a liability anymore, owing to its inherent value [22]. It resonates with Chen et al. [30] that data could be comparable to material assets and human capital. There is also a trend of collecting data without a predefined purpose [22], thus, resulting in people striving to capture the entire population or systems data, where *n=all*, to use statistical terminology, to harness optimum value by capturing the full resolution of the data points of the problem than analysing a sample of the population [31]. For instance, Lambda Architecture facilitates capturing the full resolution of data to answer a question by looking at the entire dataset from the past in the data systems using queries such as *query=function(all data)* [3]. It echoes the views of Anderson [32] that the data deluge is making scientific methods obsolete.

The use case of data science is widespread because people increasingly leverage and capitalise on data for insights and actions. Spanaki et al. [33] highlight the need for data capabilities and skills to build innovation from data in Agriculture 4.0. There is also an increasing demand for data scientists to translate business data into insights (e.g. Li et al. [34]). Chiang et al. [35] regard data scientists as a new breed and fastest-growing career in the twenty-first century and underline the attributes of data scientists, such as technical skills (programming, statistics, mathematics, and model building) and curiosity to make unexpected discoveries from big data. The rapid permeation of the data in science and engineering disciplines and concomitant demands to use data effectively have mandated data science programmes to address the data challenges of the real world through various mechanisms, such as the Lambda Architecture that connect big data with fast data [3] and big data process model to transform raw data into value [36]. We live in an era where better data and analytical tools would win the day [32]. It is evidenced by firms such as Google and Facebook that are born-digital and digital astute to leverage data science technologies to capitalise on data to deliver value propositions.

The extant literature also discusses the lack of knowledge and skills extensively to perform data processing activities, such as data gathering, data processing, and data analytics in various fields: chemical engineering [35], smart factories [34], and social science [31]. Hence, education of students and initiative from the government to enhance the curriculum [8] are crucial to face the challenges posed by the rapid datafication of society. Similarly, data analytics skills gaps in people across the sectors are also spotted [18], and responsible stakeholders push governments to fund and grow university data analytics programmes [8]. For instance, to overcome the shortage of people with data skills [37], higher education institutes in the United States offer blended programmes such as Data Science, Data Analytics, and Business Intelligence [9].

III. STUDY DESIGN AND METHODS

The following sections briefly describe the empirical setting and method adopted in the current research.

A. Empirical Setting

The plan to offer a data science programme, the first of its kind in Bhutan, at Sherubtse College has been in discussion for a long time. However, the real fruition of the idea began during the three-day programme development retreat to work on the data science proposal document. A twenty-three-page programme proposal document, sent onward to the planning and resources committee of the University, was the outcome of the retreat. The document includes a comprehensive description of the need to introduce data science in tertiary education in Bhutan, particularly at Sherubtse College. In addition, the document delineates Sherubtse College, amongst other sister colleges of the University, as the right college to offer the proposed programme. Sherubtse College is a publicly financed liberal arts and sciences college. This came as our strength to deliver breadth subjects imperative for effective delivery of data science; institutions that host a data science programme must be diverse in offering programmes [9]. The proposal also includes a provisional curriculum structure, subject matters, pedagogical approaches, and student admission criteria.

A proposal for a new programme typically stems from a host college at the University. Nevertheless, as deemed necessary and relevant to the University's overall objectives, the Academic Planning and Resources Committee (hereafter, the 'planning and resources committee') may also propose and instruct a relevant college to initiate a proposal [12]. Before enrolling students in a new programme, it has to receive planning approval as well as academic approval from the University. In consultation with the relevant department or school, the college management performs need analysis and aligns with the strategic plan of the university and college. Once the college is convinced of its relevance and significance, the plan is incorporated into the University's strategic plan with a timeline and intake number. Accordingly, the respective department is tasked to form a programme committee to work on the proposal, including a write-up justifying the need for the programme, the demand for the programme, the University's overall strategy, and pedagogical resources. To explore the demand and suitability of such a programme, the programme committee surveys and meets stakeholders for discussions, as discussed in the 'Analyses and Findings' section.

B. *Research Method*

A qualitative method was adopted for an exposition of the phenomenon of interest. We solicited primary data via a needs assessment survey of stakeholders' insights and perspectives on data science. The participants were randomly sampled from various organisations across Bhutan, as illustrated by the parenthetical value in the first column of TABLE 1. Our work was underpinned by the analytical lens of routines [13] to study macro and micro dynamics of processes and recurrent action patterns enacted by various actors [38] involved in programme development. Routines study the situated action and recurrent patterns of action [15, 39] that emerge from multiple interdependent actors while accomplishing a task. *Action* is steps in the process of accomplishing a task [15]. *Pattern* is the repetition of an action over time [40]. Routines are robust tools to capture the ostensive—abstract idea of routines (routines in principle); performative—the actual performance of routines by actors (routines in practice); and generativity—capable of producing or reproducing something by actors [13]. An actor in routines can be an intermediary or mediator. Intermediary connects different human or non-human entities through recurrent action patterns to transport meaning without transforming it. A mediator, in contrast, has the potential to transform, translate, and modify meaning to create new things through their generative properties [41]. Routines have been used to study artificial intelligence lab [41], agile software development project [16], information systems design [42], and creativity crisis in a video game development studio [43]. Nevertheless, thus far, to the best of our knowledge, routines have not been explicitly used to study workflow processes in academic programme development.

IV. ANALYSES AND FINDINGS

The following sections describe the analyses and syntheses of the need assessment survey and expositions on the enactment of routines by different actors involved in designing and developing a data science programme. English is a working language in Bhutan, so the survey does not require translation. Open coding and axial coding [44] were used to qualitatively code, analyse, and synthesise the respondents' answers. We expended efforts to ensure that the open codes and axial categories that emerged from the coding analysis remained faithful to the data [38]. The concepts, constructs, and ontologies of routines were used to capture the aspects of activities of actors and patterns of action that emerge from them while being involved in designing and developing an academic programme.

A. *Analysis of the Need Assessment Survey*

The survey received responses from 30 participants across various organisations in Bhutan, as illustrated by parenthetical values in the first column in TABLE 1. Their perspectives on different aspects of data science appropriate to the proposed programme and knowledge of the current demand of the job market were crucial in developing a robust programme. Thus, the involvement of different stakeholders who constitute facets of the triple helix—university, government, and industry [45] was imperative to gathering real-world insights and perspectives on data science. When asked about the need for data science graduates in Bhutan, the overwhelming majority reported 'Yes' (more than two-thirds), and the rest said 'Maybe'.

An open-ended question was asked to garner perspectives on data science from the respondents. The analysis of their answers using axial coding [44] is presented in TABLE 1. Only some representative open codes are included due to space concerns. An empty cell (with '–') in the table indicates that the coding of answers from the respondent category did not yield any valuable code. The coding facilitated a deeper insight into the responses based on three axial code categories: 'Data Science Career Opportunities', 'Data Science Knowledge and Skills', and 'General Data Science Perspectives'.

TABLE 1. OPEN AND AXIAL CODE OF THE NEED ASSESSMENT SURVEY RESPONSES

Respondent Categories		**Axial Code Categories**		
		Data Science Career Opportunities	**Data Science Knowledge and Skills**	**General Data Science Perspectives**
Government Organisation (20)	Open Code	Demand for data science; Data literate human resources; Citizen data scientist	Evidence-based decisions; Data analytics know-how; Dark data situation; Frequent curriculum revision	Data-driven policy and decision; Data valorisation; Data as an agent for national development
Corporation (6)		Require specialist in data science	Balance theory and practice; Data analytics skills	New oil; Potential of raw data; Data possibilities
Private Organisation (1)		Positive prospects for data science programme	Require people with data knowledge and skills	Societal benefits of data science
Autonomous Organisation (1)		–	Prioritise skills development	Sceptical about data science
Research and Corporate (1)		Opportunity for data science	–	–
State Owned Enterprise (1)		Require data scientist	Exploit data for insights	Organisational data asset
Total Open Codes		24	33	25

A persistent criticism of the tertiary academic programme in Bhutan is the imbalance in theory and practical. It is often described as theory-laden. The serious disconnect between what students study in their undergraduate and the demand in the job market is also raised by various stakeholders. A respondent from an autonomous organisation noted the following:

"Please focus on providing skills."

Data skill such as 'analytics' is the keystone of data science [1, 9] and a tool to realise the business value of data [18]. It is corroborated by qualitative codes from inductive axial coding in TABLE 1 that the respondents emphasised knowledge and skills. Tang and Sae-Lim [9] highlight the inadequacy of data science programmes in delivering analytical skills in their study on data science programmes in the information schools (iSchools) of the United States. Mikalef et al. [18] also underscored the value of and the need for soft skills to augment graduates' data analytics skill set.

Some responses underline the need for the programme developers to be cognisant of the current state of the job market and consider stakeholders' perspectives. For instance, a respondent from a government organisation highlights an abundance of data but siloed and inaccessible to extract actionable insights, a case that is characteristic of a 'dark data' situation [11]:

"Data is in abundance in Bhutan. However, the issue is these data are scattered everywhere, within different agencies, ministries, organisations, etc. Getting access to these data is difficult, and most of the time they are lying idle ... Graduates with the skills learned in data science are in much requirement in the job market today. Yes, there are job prospects for the graduates of data science in Bhutan."

B. Routinisation of Programme Development Process

The link between the macro-level dynamics of the workflow processes and the micro-level action patterns of actors involved in developing an academic programme at the University is analysed from a routines' perspective. Fig 2 un-black box the enactment of routines by various actors while being involved in developing a data science programme as a flow chart [46]. We used 'situated action' [47] and 'relationality' [41] of actors as a unit of analysis, thus foregrounding action [15], which is vital in routines. The figure likewise details the flow of data and information as a boundary object within and across different entities, such as committees, panels, and stakeholders.

A programme development process is a concerted effort of diverse actors, such as committees, stakeholders, college president, and college library, as illustrated in Fig 2. The ostensive and performativity of actors are manifested by their situated actions while executing a task. An actor acts as an intermediary or mediator [41] across different phases of the programme development. In Fig 2, the dashed box and solid rounded box represent the intermediaries and mediators, respectively. A dashed arrow represents the exchange of data or information with limited power to transform meaning. In contrast, a solid arrow represents the ability to transform

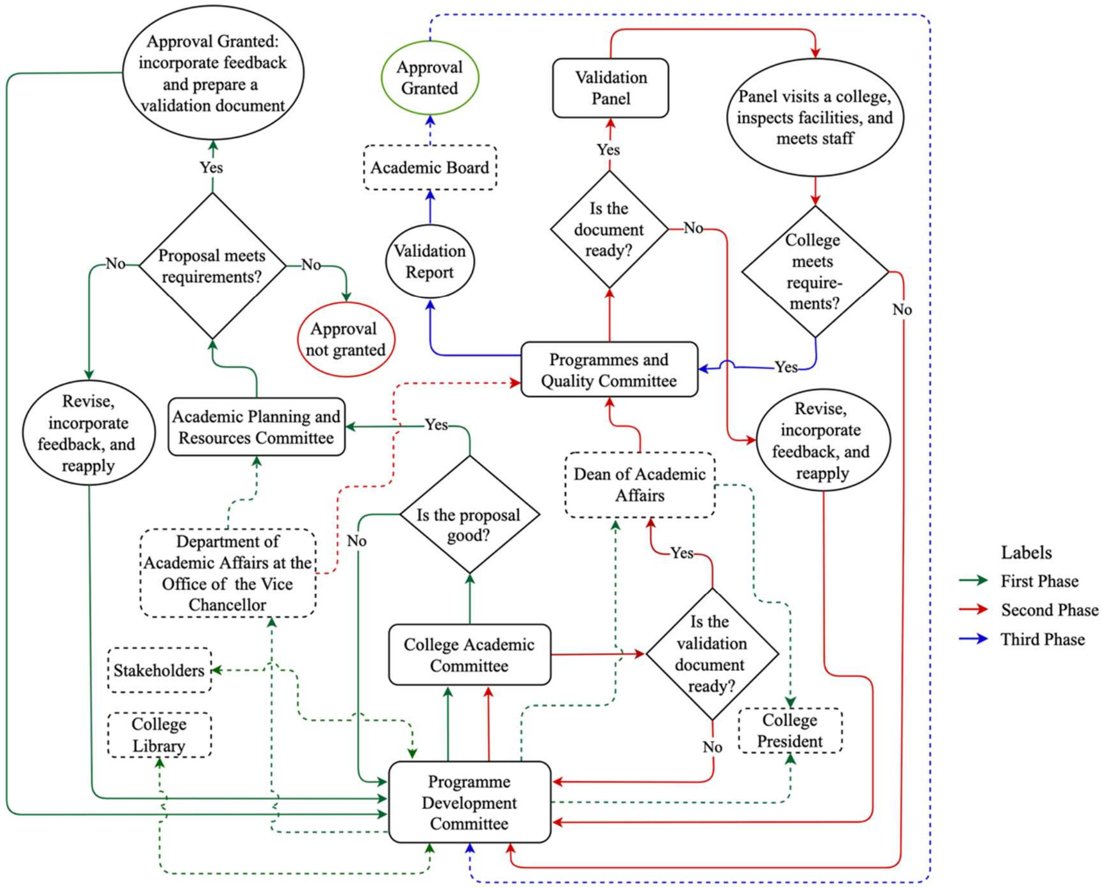

Fig 2. Recurrent, recognisable action patterns in the academic programme development workflow at the Royal University of Bhutan

meaning and the potential to generate new things. Mediators have an implication on subsequent transformation in the data and transfer of data between the actors. The generative potential of actors and what results from their generativity while enacting routines are central to our analysis.

In the interest of space concern, it is impossible to provide an exhaustive account of each actor's situated action, recurrent action patterns, and generative effects. Only the first, second, and third phases, indicated by green, red, and blue colour, respectively, in Fig 2, of the programme development process usually observed at the University were selected for an exposition. The ostensive aspects of routines are abstract, general ideas that guide performances; in simple terms, they are a 'schematic form of routines' or 'routine in principle' [13]. In Fig 2, we capture the ostensive aspects of routines, such as how policies and frameworks guide an actor's action and action patterns while being engaged in developing a programme at the University. The performativity constitutes 'routines in practice' and embodies specific actions of routines [13] in the programme development process workflow, which we describe as a vignette below.

Phase 1: Programme Proposal Document for the University's Academic Planning and Resources Committee

The first phase of programme development begins by translating an abstract idea of offering a data science programme into a concrete proposal document onward to the planning and resources committee. It involves preparing a document that includes a brief write-up on the subject matters, curriculum structure, pedagogical approaches, and a list of topics. Additionally, the document also mentions aspects of resources in the college to offer and sustain the programme. The programme committee submits the proposal to the College Academic Committee (hereafter, the 'academic committee').

The performativity of the academic committee of the college is displayed through their enactment of routines when departments submit a proposal to offer a programme: check the completeness and justifications, after which suggestions and feedback return to the programme committee for improvement or grant approval for further submission via the Dean of Academic Affairs (hereafter, the 'academic dean') of the college, as illustrated in Fig 2. The academic dean submits the final version of the proposal to the planning and resources committee of the University. The programme committee and the college president and/or academic dean defend the proposal when the planning and resources committee call for a programme proposal meeting. It is an intellectually tense and contentious moment. A similar experience was discussed in the study of the 'Always Playable' Project through the lens of routines at Ubisoft's Montreal studio [43].

Based on the documents and justifications received and the outcome of the meeting with the programme committee, the planning and resources committee will: (1) approve the incorporation of the proposed programmes in the University's forward academic plan as a programme that helps to fulfil the University's obligation to provide relevant and good quality programmes; (2) approve the proposed student numbers; (3) approve the further development of the proposal in preparation for submission to the Programmes and Quality Committee (hereafter, the 'programmes and quality committee') of the University; and (4) agree to the incorporation of the resource requests in the University's budget proposals [12, E1 Planning Approval for a New Programme]. The routines enacted by the planning and resources committee act as a critical passage point [14] in the programme development process at the University, and the outcome of the planning and resources committee meeting is non-binary.

The two 'Nos' of the decision flow in Fig 2 of the first phase are worth describing. 'No' on the left means that the proposal is not discarded but is not yet ready to accord approval. The likelihood of the proposal for a new programme being discarded is high, as illustrated by 'No' on the right of the Academic Planning and Resources Committee's decision in Fig 2, if the planning and resource committee deem that there is no strong basis for a need for a new programme. Other reasons for the proposal being discarded include the proposed programme does not meet the University's overarching strategic goals or that resources are inadequate to sustain the proposed programme. The non-binary decision also illustrates variations in the enactment of routines [13]. The proposal meeting determines the fate of a proposed programme. Nonetheless, the planning and resource committee accepted our proposal for the data science programme in the first meeting itself and thus alluding to going ahead with the next phase.

Phase 2: Programme Validation Document for the University's Programmes and Quality Committee

After planning approval is granted by the planning and resources committee, the logical next step for the programme committee is to develop a validation document. Red arrows in Fig 2 represent the enactment of routines by actors in the second phase. The validation document is more rigorous with meticulous detail. It includes crucial information such as the programme's aims and objectives, rationale and justification, human resources, academic resources, subject matters, and reading lists. The document is then submitted again to the academic committee of the college for consideration and subsequently submitted to the University's programmes and quality committee by the college's academic dean.

It is evident from the workflow in Fig 2 that programme validation is an iterative process if the validation document does not make it in the first go. The programmes and quality committee decides whether there is an acceptable basis for validating the proposed programme based on the documentation and preparedness of the proposing college. The programmes and quality committee will consider and determine issues such as logistical and resource shortcomings that require addressing; staff's readiness and cognisance of issues posed by a new programme and mitigation plans; and sufficiency of the information provided, failing which the committee will seek for additional information or direct for re-submission. The performativity of the committee is demonstrated by actions such as producing validation reports and written constructive recommendations for improving the validation document. Thus, from routines ontology, the programmes and quality committee act as mediators, as indicated by the solid rounded box in Fig 2.

The college usually proposes panel members for programme validation [12]. The programme committee will then revise the document taking into consideration the recommendations of the programmes and quality committee, which is then sent to members of the panel to study in preparation for the validation. The panel visits the college for validation. During the visit, the ostensive and performativity of the panel are illustrated through activities such as meeting privately to collate the main issues related to the programme; engage in detailed discussion with the staff; meet the college management to confirm the adequacy and allocation of resources for the programme; meet students to establish their learning and teaching experience at the college; and assess the adequacy of facilities that are proposed to support the effective operation of the programme. The panel members present a written report on the final day of the validation to the college. Based on the findings from their visit and interactions, the possible result of the assessment by the panel is: (1) approve without conditions with a review planned to take place within the normal duration for the periodic review of a programme; (2) approve without conditions with a review planned to take place in a period less than the normal duration for the periodic review of a programme; (3) approve upon meeting specified conditions; or (4) not approve, and the college will be invited to reapply, taking into account the panel's comments. In the next logical phase, the validation report is then sent to the programmes and quality committee to make a considered recommendation to the University's Academic Board.

Phase 3: Programme Validation Report for the University's Academic Board

In the third phase of the programme development process workflow, the Academic Board will consider the recommendations of the programmes and quality committee and the validation report to approve the launch of a new programme. The outcome of the validation or review is a revised programme document called the Definitive Programme Document, which is submitted electronically to the Director of Academic Affairs at the Office of the Vice Chancellor of the University. In Fig 2, from the initial proposal to offer a programme to the final approval by the Academic Board of the University, the recurrent action patterns enacted by different actors involved in the programme development process are living and generative systems [48]. The generativity of actors is demonstrated across different phases of the programme development process, with material deliverables in each stage such as proposal document, validation report, and programme document. From the expositions above and further corroborated by Fig 2, the progression of events in space and time through the enactment of routines by different actors culminate in the programme document for the data science programme approved by the Academic Board of the University. After the Academic Board approves, the programme is launched, and the programme committee is responsible for implementing it.

V. Discussion

The following sections briefly discuss the data science at Sherubtse College, contributions to knowledge, and limitations of the study

A. *Programme Development Process Through the Analytic Lens of Routine Dynamics*

The recurrent action patterns enacted by a web of actors in Fig 2 are generally followed for developing an academic programme at the University. However, there might be subtle variations in the programme development workflow due to endogenous or exogenous factors [13]. As actors, especially members of the committees, acquire experience developing a programme, the process becomes part of their procedural memory [48], easing the design and development of an academic programme. The programme development routines at the University occur in a stable environment as new programmes cannot be offered without undergoing proposal scrutiny and a rigorous validation process. For this reason, the outcome of the planning and resources committee meeting is non-binary.

Actors with a stake in developing an academic programme at the University share their insights and experiences through ostensive and performativity. For instance, the college president is not actively involved in the programme development process, but they are an essential source of crucial management support (e.g. participating in the programme proposal meeting) and thus play a critical role as an intermediary, as indicated by the dashed box in Fig 2. At the University, the Wheel of Academic Law [12] serves as a meta-routines (that is, routines for changing routines [15]) to guide the programme development process from initiation to completion. Moreover, it codifies the recurrent action patterns whenever there is a proposal for a new programme. Thus, viewing the programme development process workflow through the analytic lens of routines allows one to comprehend the complex nexus of activities, actors, actions, and routines in accomplishing a task.

B. *Data Science Programme at Sherubtse and Its Quality*

The Bachelor of Science in Data Science programme is Sherubtse College's endeavour to develop the data proficiency of prospective students, which has implications for human resource development [8]. The in-house designed and developed programme is robust and current with the needs and demands of the twenty-first-century job market, namely the needs and challenges of data science in Bhutan. It was launched in the autumn academic session of 2021 with 29 students with a vision of building data science literacy and developing the data analytics knowledge and skills of students who want to pursue a career in data science. Different entities with a stake in the proposed data science programme played a crucial role in actualising the proposal through their constructive feedback and recommendations.

The programme takes into account the current fourth paradigm of science that is exploratory in nature and is characterised as 'data intensive; statistical exploration, and data mining' [31]. In addition, the programme considers the tools and methods required to read constituents of data DNA—knowledge, insights, and potential [1]. The programme aims to build the student's data science repertoire with market demands. For instance, Python is a widely used programming language in the data science domain and the language recommended by the stakeholders. So, it is adopted as a fitting language for practical aspects of the programme. Similarly, as desired by the respondents to the need assessment survey, R language is also included in the syllabus to complement Python. Most of the modules in the curriculum are designed in such a way that it requires significant effort for hands-on experience working with data. An industrial attachment/internship of two months followed by a semester-long capstone project in the final year would equip students with adequate skills and confidence in dealing with data in their workplace. The respondents also highlight the management and stewardship of data. So, the curriculum includes data management and governance guidelines that are gaining traction in academia and industry, such as 'FAIR Data

Principles' to enlighten students on how to make data findable and reusable for machines and humans [49] and ultimately optimise the use and reuse of small data or big data.

An assessment of the quality of a product depends on the context and contextual metrics to assess quality. The data science programme of Sherubtse College is underpinned by the 'Computing Competencies for Undergraduate Data Science Curricula' framework of ACM [17] to be at par with similar programmes offered globally. Moreover, the interdisciplinarity of data science was taken into account by including a rich breadth subject to balance computer science, mathematics, statistics, and domains/business knowledge to develop the know-how to harness the social and economic value of data [22] in the evolving data world [1]. It resonates with Tang and Sae-Lim [9] that a single subject domain is insufficient to provide the content and skills required for data science. In the same vein, Monino [5] underscores the essentiality of excellent analytical and related skills such as understanding and manipulating large datasets and interpreting and applying the results. The encodement of the need assessment survey responses reveals an apparent lack of people with data knowledge and skills, as in TABLE 1, to distil meaning from a vast volume of data in various formats generated at great velocity from various sociotechnical systems [50]. The respondent's concerns were considered by balancing the theoretical and practical aspects of modules in the curriculum of the programme. In view of the short shelf life of technologies and techniques of data science, and knowledge and skills for that matter, the programme document of data science is a living document. Hence, the University mandates reviewing the programme after one cycle—six years for the four-year programmes—to ensure that the right knowledge and skills are delivered to the students to deal with the data challenges of today and tomorrow adequately.

C. Contributions to Knowledge and Limitations of the Study

Thus far, routine dynamics have not been applied to study curriculum development in academia to our understanding. So, this study makes a valuable contribution to the body of knowledge on data science, namely in the development of an academic programme in tertiary education. We demonstrated that routines could be used as an instrument to analyse workflow processes by incorporating routines' concepts and constructs into a flow chart diagram, as illustrated in Fig 2. Similarly, the systematic description of the use of routines to study processes in a workflow would be a valuable source of information for those who want to use routines as an analytic lens to study a similar phenomenon. Our study also makes a methodological contribution to the knowledge base of routines by demonstrating the applicability of routines beyond their usual application.

The study has generalisability and replicability limitations. We traded off these issues to study an interesting context worth exploring via an analytical framework of routine dynamics. The current study only examines the dynamics of processes, actors, and actions involved in designing and developing an academic programme. For a comprehensive insight into the programme's efficiency, from content delivery to absorption of the graduates in the job market, a long-term study incorporating perspectives of stakeholders and students is required. This would help objectively assess the programme's efficacy relative to the demands of the job market and stakeholder expectations. Moreover, it would provide an avenue to enhance the programme further.

VI. Conclusion

The central focus of academic programme development is a product, and less attention is paid to subtle dynamics between the stakeholders and processes in the workflow of designing and developing a programme. Academic programme development is an ensemble of constitutive processes, actors, actions, and routines. Thus, we anchored our study on the analytical lens of routines to capture the orchestration of ostensive and performativity of actors across different programme design and development phases. Notwithstanding the challenges of developing an academic curriculum in-house, the programme committee designed and developed a four-year undergraduate data science programme at Sherubtse College in Bhutan to equip students with theoretical knowledge and practical skills to deal with real-world small data or big data. This study contributes to the knowledge base of data science, namely the use of routine dynamics as an analytical instrument to examine processes in a workflow. Likewise, the systematic description of the academic programme development process by anchoring on routines offers valuable information for curriculum developers mandated to produce graduates with a broad spectrum of knowledge and skills in data science.

REFERENCES

[1] L. Cao, "Data science: A comprehensive overview," *ACM Computing Surveys,* vol. 50, no. 3, pp. 1-42, 2017. [Online]. Available: https://doi.org/10.1145/3076253.

[2] R. L. Ackoff, "From data to wisdom," *Journal of Applied Systems Analysis,* vol. 16, no. 1, pp. 3-9, 1989.

[3] N. Marz and J. Warren, *Big Data: Principles and best practices of scalable realtime data systems*. New York: Manning, 2015.

[4] K. Cukier and V. Mayer-Schoenberger, "The rise of big data: How it's changing the way we think about the world," *Foreign Affairs,* vol. 92, no. 3, pp. 28-40, 2013. [Online]. Available: https://www.jstor.org/stable/23526834.

[5] J.-L. Monino, "Data value, big data analytics, and decision-making," *Journal of the Knowledge Economy,* vol. 12, no. 1, pp. 256-267, 2021. [Online]. Available: https://doi.org/10.1007/s13132-016-0396-2.

[6] M. Loukides, *What is data science?* " O'Reilly Media, Inc.", 2011.

[7] M. Fukuyama, "Society 5.0: Aiming for a new human-centered society," *Japan Spotlight,* vol. 27, pp. 47-50, 2018.

[8] S. J. Qin and L. H. Chiang, "Advances and opportunities in machine learning for process data analytics," *Computers & Chemical Engineering,* vol. 126, pp. 465-473, 2019. [Online]. Available: https://doi.org/10.1016/j.compchemeng.2019.04.003.

[9] R. Tang and W. Sae-Lim, "Data science programs in US higher education: An exploratory content analysis of program description, curriculum structure, and course focus," *Education for Information,* vol. 32, no. 3, pp. 269-290, 2016.

[10] Department of Information Technology & Telecom, "e-Governance Policy for the Royal Government of Bhutan," ed. Thimphu: Ministry of Information & Communications, 2019, p. 10.

[11] P. B. Heidorn, "Shedding light on the dark data in the long tail of science," *Library Trends,* vol. 57, no. 2, pp. 280-299, 2008, doi: 10.1353/lib.0.0036.

[12] Royal University of Bhutan. "The Wheel of Academic Law." https://www.rub.edu.bt/regulation/ (accessed 31 March, 2022).

[13] M. S. Feldman and B. T. Pentland, "Reconceptualizing organizational routines as a source of flexibility and change," *Administrative Science Quarterly,* vol. 48, no. 1, pp. 94-118, 2003. [Online]. Available: https://doi.org/10.2307/3556620.

[14] M. S. Feldman, B. T. Pentland, L. D'Adderio, and N. Lazaric, "Beyond routines as things: Introduction to the special issue on routine dynamics," vol. 27, no. 3, pp. 505-513, 2016. [Online]. Available: https://doi.org/10.1287/orsc.2016.1070.

[15] B. T. Pentland, M. S. Feldman, M. C. Becker, and P. Liu, "Dynamics of organizational routines: A generative model," *Journal of*

Management Studies, vol. 49, no. 8, pp. 1484-1508, 2012. [Online]. Available: https://doi.org/10.1111/j.1467-6486.2012.01064.x.
[16] T. Gustavsson, "Dynamics of inter-team coordination routines in large-scale agile software development," in *European Conference on Information Systems*, 2019. [Online]. Available: https://aisel.aisnet.org/ecis2019_rp/40.
[17] A. Danyluk and P. Leidig, "Computing competencies for undergraduate data science curricula: ACM Data Science Task Force," 2021. [Online]. Available: https://www.acm.org/binaries/content/assets/education/curricula-recommendations/dstf_ccdsc2021.pdf
[18] P. Mikalef, M. N. Giannakos, I. O. Pappas, and J. Krogstie, "The human side of big data: Understanding the skills of the data scientist in education and industry," in *2018 IEEE Global Engineering Education Conference*, 2018: IEEE, pp. 503-512, doi: 10.1109/EDUCON.2018.8363273.
[19] M. Lycett, "'Datafication': Making sense of (big) data in a complex world," ed: Taylor & Francis, 2013.
[20] K. Ayankoya, A. Calitz, and J. Greyling, "Intrinsic relations between data science, big data, business analytics and datafication," in *Southern African Institute for Computer Scientist and Information Technologists*, 2014, pp. 192-198. [Online]. Available: https://doi.org/10.1145/2664591.2664619.
[21] C. K. Emani, N. Cullot, and C. Nicolle, "Understandable Big Data: A survey," *Computer Science Review,* vol. 17, pp. 70-81, 2015. [Online]. Available: https://doi.org/10.1016/j.cosrev.2015.05.002.
[22] W. A. Günther, M. H. R. Mehrizi, M. Huysman, and F. Feldberg, "Debating big data: A literature review on realizing value from big data," *The Journal of Strategic Information Systems,* vol. 26, no. 3, pp. 191-209, 2017. [Online]. Available: https://doi.org/10.1016/j.jsis.2017.07.003.
[23] D. Wang, "Building value in a world of technological change: Data analytics and industry 4.0," *IEEE Engineering Management Review,* vol. 46, no. 1, pp. 32-33, 2018, doi: 10.1109/EMR.2018.2809915.
[24] S. Sathananthan, "Data valuation considering knowledge transformation, process models and data models," in *International Conference on Research Challenges in Information Science*, 2018: IEEE, pp. 1-5, doi: 10.1109/RCIS.2018.8406649.
[25] M. O. Gokalp, K. Kayabay, M. A. Akyol, P. E. Eren, and A. Koçyiğit, "Big data for industry 4.0: A conceptual framework," in *International Conference on Computational Science and Computational Intelligence*, 2016: IEEE, pp. 431-434, doi: 10.1109/CSCI.2016.0088.
[26] I. Lee, "Big data: Dimensions, evolution, impacts, and challenges," *Business Horizons,* vol. 60, no. 3, pp. 293-303, 2017, doi: 10.1016/j.bushor.2017.01.004.
[27] K. Vassakis, E. Petrakis, and I. Kopanakis, "Big data analytics: Applications, prospects and challenges," in *Mobile Big Data*: Springer, 2018, pp. 3-20.
[28] N. Ismail. "Extracting value from data: how to do it and the obstacles to overcome." Information Age. https://www.information-age.com/extracting-value-from-data-123480490/ (accessed 2022, 12 March).
[29] L. V. Orman, "Information paradox: Drowning in information, starving for knowledge," *IEEE Technology and Society,* 2015, doi: 10.1109/MTS.2015.2494359.
[30] M. Chen, S. Mao, and Y. Liu, "Big data: A survey," *Mobile networks and applications,* vol. 19, no. 2, pp. 171-209, 2014. [Online]. Available: https://doi.org/10.1007/s11036-013-0489-0.
[31] R. Kitchin, "Big Data, new epistemologies and paradigm shifts," *Big Data & Society,* vol. 1, no. 1, 2014. [Online]. Available: https://doi.org/10.1177/2053951714528481.
[32] C. Anderson, "The end of theory: The data deluge makes the scientific method obsolete," vol. 16, ed: Wired, 2008, pp. 16-07.
[33] K. Spanaki, E. Karafili, and S. Despoudi, "AI applications of data sharing in agriculture 4.0: A framework for role-based data access control," *International Journal of Information Management,* vol. 59, p. 102350, 2021. [Online]. Available: https://doi.org/10.1016/j.ijinfomgt.2021.102350.
[34] S. Li, G. C. Peng, and F. Xing, "Barriers of embedding big data solutions in smart factories: insights from SAP consultants," *Industrial Management & Data Systems,* 2019. [Online]. Available: https://doi.org/10.1108/IMDS-11-2018-0532.
[35] L. Chiang, B. Lu, and I. Castillo, "Big data analytics in chemical engineering," *Annual Review of Chemical and Biomolecular Engineering,* vol. 8, pp. 63-85, 2017. [Online]. Available: https://doi.org/10.1146/annurev-chembioeng-060816-101555.
[36] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *International Journal of Information Management,* vol. 35, no. 2, pp. 137-144, 2015. [Online]. Available: https://doi.org/10.1016/j.ijinfomgt.2014.10.007.
[37] H. Chen, R. H. Chiang, and V. C. Storey, "Business intelligence and analytics: From big data to big impact," *MIS quarterly,* pp. 1165-1188, 2012. [Online]. Available: https://doi.org/10.2307/41703503.
[38] E. Baralou and D. D. Dionysiou, "Routine dynamics in virtual teams: the role of technological artifacts," *Information Technology & People,* 2021. [Online]. Available: https://doi.org/10.1108/ITP-03-2020-0109.
[39] B. T. Pentland and M. S. Feldman, "Narrative networks: Patterns of technology and organization," *Organization Science,* vol. 18, no. 5, pp. 781-795, 2007. [Online]. Available: https://doi.org/10.1287/orsc.1070.0283.
[40] M. C. Becker, "The concept of routines: some clarifications," *Cambridge Journal of Economics,* vol. 29, no. 2, pp. 249-262, 2005. [Online]. Available: https://www.jstor.org/stable/23602185.
[41] K. Sele and S. Grand, "Unpacking the dynamics of ecologies of routines: Mediators and their generative effects in routine interactions," *Organization Science,* vol. 27, no. 3, pp. 722-738, 2016. [Online]. Available: https://doi.org/10.1287/orsc.2015.1031.
[42] D. Sammon, T. Nagle, and J. McAvoy, "The ISD process as a 'live routine': The mindless behaviours of a narrative network," in *European Conference on Information Management and Evaluation*, 2011: Academic Conferences and Publishing International Limited, pp. 428-435. [Online]. Available: https://doi.org/10.1016/j.infsof.2014.01.007.
[43] P. S. Cohendet and L. O. Simon, "Always playable: Recombining routines for creative efficiency at Ubisoft Montreal's video game studio," *Organization Science,* vol. 27, no. 3, pp. 614-632, 2016. [Online]. Available: https://doi.org/10.1287/orsc.2016.1062.
[44] J. Saldaña, *The Coding Manual for Qualitative Researchers*. SAGE, 2021.
[45] L. Leydesdorff and H. Etzkowitz, "The triple helix as a model for innovation studies," *Science and Public Policy,* vol. 25, no. 3, pp. 195-203, 1998. [Online]. Available: https://doi.org/10.1093/spp/25.3.195.
[46] B. T. Pentland and M. S. Feldman, "Designing routines: On the folly of designing artifacts, while hoping for patterns of action," *Information and Organization,* vol. 18, no. 4, pp. 235-250, 2008. [Online]. Available: https://doi.org/10.1016/j.infoandorg.2008.08.001.
[47] B. T. Pentland and M. S. Feldman, "Organizational routines as a unit of analysis," *Industrial and Corporate Change,* vol. 14, no. 5, pp. 793-815, 2005. [Online]. Available: https://doi.org/10.1093/icc/dth070.
[48] B. T. Pentland, T. Hærem, and D. Hillison, "The (n) ever-changing world: Stability and change in organizational routines," *Organization Science,* vol. 22, no. 6, pp. 1369-1383, 2011. [Online]. Available: https://doi.org/10.1287/orsc.1110.0624.
[49] M. D. Wilkinson *et al.*, "The FAIR Guiding Principles for scientific data management and stewardship," *Scientific Data,* vol. 3, no. 1, pp. 1-9, 2016, doi: 10.1038/sdata.2016.18.
[50] F. Frankel and R. Reid, "Big data: Distilling meaning from data," *Nature,* vol. 455, no. 7209, pp. 30-30, 2008. [Online]. Available: https://doi.org/10.1038/455030a.