Managing Quality Constraints in Technology-managed Learning Content Processes

Claus Pahl, Neel Mani
CNGL, School of Computing, Dublin City University
Dublin 9, Ireland

Abstract: Technology-enhanced learning content processes consist of individual activities related to the creation, composition, consumption and analysis of content facilitated through services. These service processes are often enacted across different boundaries such as organisations, countries or even languages. Specifically, looking at the quality of learning content and other artefacts and the governance of respective processed through services in this context is important to control quality requirements. We assume a partially automated workflow process for the content lifecycle. We suggest a rule-based constraints monitoring of learning content processes. A learning domain ontology shall capture the key data/content types, activities and constraints, which forms the basis of a rule-based policy monitoring solution that takes content provenance data into account.

Introduction

Learning content and related data is increasingly processed, i.e. created, composed, delivered and consumed in distributed settings by different services in a technology-enhanced learning process (Rockley, 2003). As a consequence, maintaining quality across a boundary-crossing process is a challenge. This is a process that involves the instructor in the creation and delivery stages and the learner for instance in the consumption stages of the process (Murray, 2003; Melia, 2009). The focus of this paper is learning content and data quality in learning content processes. We aim to enable learning service processes for learning content manipulation and consumption based on a formalised content model, which requires:

- A layered content model consisting of (at the bottom) core content, the states and stages of processing on top of that, and a provenance layer linking content and processing to their origins and dates.
- An activities and operations framework that defines the processing and manipulation activities on content in the context of provenance data.

The W3C provenance model (W3C, 2014) plays a pivotal role here for logging process activities, but also as a metadata framework for constraints and rules. Provenance is about the activities (here create, compose or consume) that agents (here authors, instructors or learners) carry out on entities (here content). Our aim is to define a content-centric quality assurance framework that allows quality requirements to be defined as constraints to be monitored and managed, i.e., we need a framework for the definition, monitoring and handling of quality concerns as provenance-based workflow constraints. Quality in a learning content process refers to the quality of the content, but also the quality of the learning experience as a process. We are specifically interested in quality concerns that can be assessed automatically (but this may include manually provided data).

Our solution provides a novel contribution in the form of an ontology-based policy constraints definition. We explore quality management and governance for learning content processes, i.e., to define, monitor and analyse. In a wider sense, this is a governance concern. Our solution specifically extends process adaptation and customisation techniques (Ardagna, 2007; Erradi, 2008), e.g., generic policy monitoring for service processes and workflows (Wang, 2010), by a learner-specific configurations.

.

Scenario - Service-based Learning Content Processing

Service-based learning content processing covering the whole content lifecycle is a distributed problem. Learning content is created, searched, manipulated (composed and maybe translated, localised, adapted and personalised) and consumed by learners. This challenges quality assurance across the lifecycle of learning content in distributed service processes covering for instance the creation (Holohan, 2006), composition (Pahl, 2004), consumption (Murray, 2003) and assessment (Melia, 2009; Lei, 2003) of content. Fig. 1 illustrates a sample learning content lifecycle process.

Fig.1. Application Context: Learning Content Processing.

The starting point for the implementation of content quality assurance is an integrated content service process enabled by a content service bus, into which the different processing, integration and management applications are plugged into, see Fig. 2. This scenario defines our wider objective. Our specific aim here is to configure the quality component of these processes by a constraints policy language (Wang, 2010). The provenance model PROV (W3C, 2014) forms the abstract constraints layer. PCPL, the Process Customisation Policy Language (Wang, 2010), is part of the technical process platform and controls the process quality through its policy engine. We will show that this technical infrastructure is suitable to support key aspects of learning and content quality governance.

Now, we briefly outline a suitable architecture for quality constraints definition and workflow execution. The main components are a policy editor, a process engine, a monitoring system and a policy engine. The PCPL rule engine from Fig. 2 can be decomposed into policy definition, monitoring and policy validation engine. The implementation platform is here the Activiti process engine (http://activiti.org/) with constraint weaving. Content is managed by the Alfresco tool (http://www.alfresco.com/). This architecture allows the policies to be defined locally and then a generic process adapted to learning-specific needs and enacted by a central process engine (Activiti).

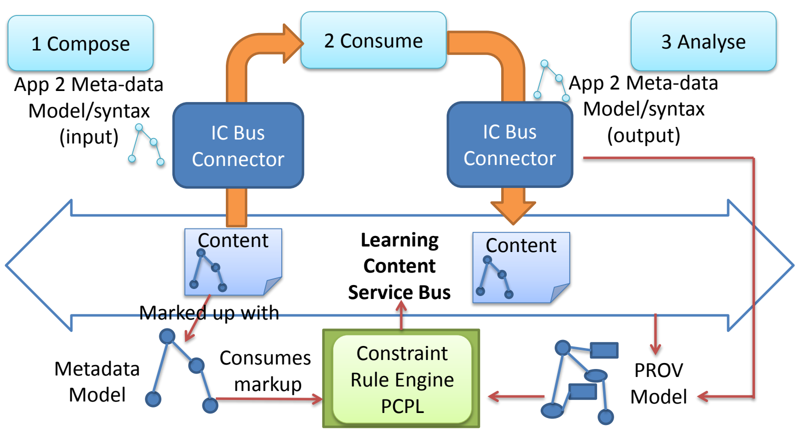

Fig. 2. Learning Content Process - Service Bus Architecture.

A sample learning process describes the creation and consumption of content through a sequence of services, Fig. 3, where the corresponding provenance model gives context. The provenance model accompanies a learning process model, consisting of the following steps:

1. Specify(ATH): A course content specification is created (node 15601)
2. Compose(LCR): Content from a content repository is assembled (node 15709)
3. Analysis(INS): Content suitability is analysed by the instructor (node 12401)
4. Import(LMS): the composed content schedule is imported into the LMS (node 16727)
5. Deliver(SRV): Guided by the LMS, individual services provide content for the learner (node 16723)
6. Learn(LRN): learner creates learning output (node 15771)

In the process description, we have added two categorisations: (i) activity: we distinguish standard activities in learning content process (specify etc.) and (ii) roles: we distinguish human and system roles in this example such as software services (SRV), individual learner (LRN), author (ATH), instructor (INS), learning management system (LMS) and digital learning content repositories (LCR). The node references in the process refer to the provenance model, Fig. 3. This graph may be built up using the provenance model as content passes through the process to create a process activity log, which forms the basis for quality analysis.

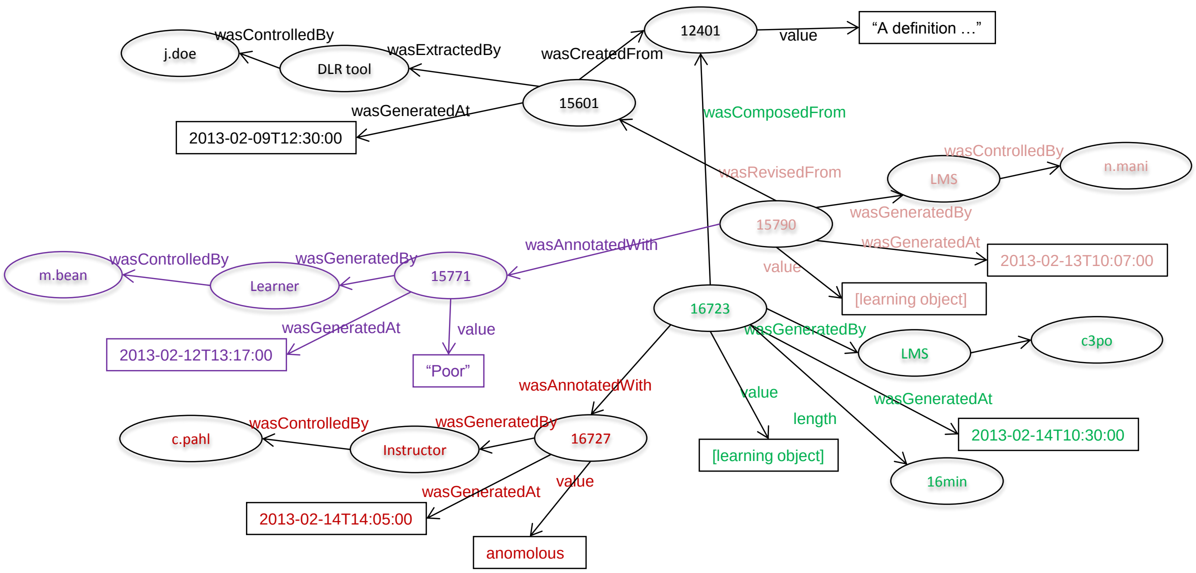

Fig. 3. Provenance Model for Learning Content Processing – in an RDF graph visualisation.

The next question concerns the quality aspects. We can identify the following concerns for process and content quality and process governance. The quality aspects shall be distinguished into four high-level constraint categories, which we try to motivate here through specific concerns:

- *Authorisation/access control*: Restricting access to content, following classical access control specifications (subject, access operation, object) because content is extracted from repository and might only be made available to a learner in specific situations (e.g. adaptation based on progress)
- *Accountability*: decide and control where are data about student content access and performance is kept to allow failure recovery, but also to address privacy and secrecy concerns
- *Workflow governance*:
  - Learning content status (e.g. specified, composed, consumed, contributed) and learning progress to be kept - ensuring that required stages of the learning content process are reached
  - Containment as a subprocess, e.g. learner tracking for assessment or adaptation to be included
- *Quality for content/process*:
  - Rating of content quality (e.g. poor, sufficient) by a learner or LMS recorded as part of provenance data
  - Learning experience as a rating of process quality (slow, satisfactory, etc.), equally recorded as provenance data

The individual constraints form policies that need to take into account the PROV structure: (i) single element type, e.g., PROV Timestamps (start/end or interval constraints) or Quality Rating (liveness constraints: should always be 'satisfactory' or better; safety constraints: should never be 'poor') and (ii) multi-element type, e.g., access control in terms of PROV (user, activity, content), status (content, 'generatedBy', activity) or governance (content, 'processedBy', user) – all presented here as simple RDF triples (http://www.w3.org/RDF/).

A Model for Content Quality Constraints

The different notations and formats involved are:

- Content: we assume RDF as the canonical meta-format, which facilitates controlled access to content as the targeted storage infrastructure and modelling of different content types. The content notations involved are text (assumed for simplicity) and PROV to capture objects with origins (actors) and operations (creation and manipulation).

- Process: For process modelling, BPMN, the activity workflow notation, is assumed as a graphical notation, which is complemented by jBPM and Java process engines for execution. Providing a runtime process execution environment is essential here.
- Constraints: PCPL, a process customisation policy language, provides a generic policy notion, extended to a process framework. Here, an integration with BPMN is implemented, which weaves the policy constraints into the process definition.

A layered domain model based on content, process and provenance data to support constraints shall be proposed. A provenance model can be maintained with the processing of content. In the provenance model (RDF linked data), the following is reflected (Fig. 3):

- change operators (activities): GeneratedBy, ComposedFrom, AnnotatedWith can be aligned to the standard content processing operations defined earlier,
- actors/participants (agent): m.bean, j.doe are named service providers that can be classified by our role categorisation scheme, e.g., the next expression j.doe:LRN->Consume links a service to a learner agent in charge of its execution,
- objects (entity): content being processed, e.g. consumed.

The provenance model can be presented as a process of change operations. This results in a 3-layered architecture (Fig. 4), where the upper layer (PROV) is defined by the provenance model (state-by-state changes), the middle layer (BPMN) is defined by the process model based on PROV content activities (changes in process) and the bottom (RDF-annotated content) refers to content processed by learning operations.

Our aim is to allow a generic content process to be adapted to situation-specific constraints, i.e. constraints that are specific to an organisation such as an education provider. Two principle solutions to deal with customisable policy constraints can be distinguished. A minimal invasive one weaves quality constraints into a process, where all constraints are monitored and managed by external services. On the other hand, an explicit extension of BPMN models constraints within the language itself and to map quality constraints into this BPMN extension is possible. Regarding the second option, BPMN constraints have been proposed as a BPMN extension (Wolter, 2007). While we adopt their constraints classification to some extent, our implementation will favour the less invasive solution suggested in (Wang, 2010) in order to achieve interoperability with a number of process and workflow engines.

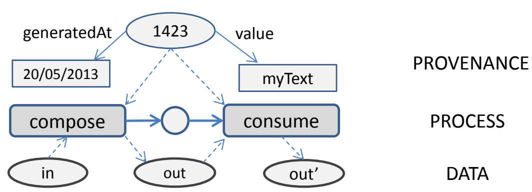

Fig. 4. Sample 3-Layered Service Process Model

Constraints and Rules

Quality constraints and their formulation as policy rules are at the core of our solution. We have already provided a classification of several quality and governance concerns as constraint types: authorisation, accountability, workflow governance and quality. The objectives of rule-based content process quality constraints for quality monitoring are twofold: firstly, optimisation, i.e., to improve quality of the learning content and experience during the process (by looking at ratings or performance measures) and, secondly, governance, i.e., to enforce access control and privacy rules.

Constraints are technically conditions on concerns. A rule associates an action related to a condition in the Event-Condition-Action (ECA) format, that checks on an event the correctness of a condition and triggers the execution of an action, if required by the condition. Thus, based on the four constraint types, we define four rules types to link conditions and actions – which shall be illustrated by some examples. The event is always the initiation of the analysis by the monitoring component.

- Authorisation/access control - example: to restrict access to content in repositories
- Accountability/audit/tracing - example: where are records or copies kept
- Workflow governance - example: *status = composed* -> *deliver(..)* or *status = delivered* ->

- Quality for content/process - *consume(..)* where the arrow -> separates condition and action example: content rating *studentRating = poor -> annotate(content)* or for process performance *time(compose) > t -> alert*

We use the PCPL policy language for process customisation to implement PROV-based constraint policies based on individual rules. Provenance constraints can be integrated in the process constrains as parts of conditions. PCPL policies consist of the following central notational elements:

- Objects: here learning content, processed by activities like compose or consume
- Activity states: capturing processing state and quality assurance state based on the learning activities
- Conditions covering the content context (owner, format etc.), the activity context or provenance/log data (authorisation, state etc.):
    - Performance/Time for processing, includes manual effort (called asynchronous) and the execution of software services (called synchronous)
    - Authorisation: who can process/access content including the location of objects (e.g. no externalisation/outsourcing allowed as a condition)
    - Existence of entity/object in a state: e.g. composed as a workflow stage
- Actions: process decisions, which cover the constraint violation handing strategies

Related Work

Current open research concerns for process quality and governance include customisation of governance and quality policies and the non-intrusive adaptation of processes with policies (Ardagna, 2007; Erradi, 2008; Riegen, 2010). Today, one-size-fits-all service monitoring techniques exist and provide support for software systems in classical sectors such as finance and telecommunications. However, their inherent structural inflexibility makes constraints difficult to manage, resulting in significant efforts and costs to adapt to the specific needs of domains such as learning technology. We discuss related work in the field of constraints and policy definition and adaptive BPEL processes. While we have also referred to BPMN, there is more work on WS-BPEL in our context Ardagna, 2007; Erradi, 2008; Riegen, 2010). While this looks at process monitoring, provenance becomes more important in open systems.

Some provenance-enabled workflow systems have been developed (Davidson, 2008). These workflow systems monitor workflow or process executions and record task names, execution durations or parameters as provenance information. Other work has focused on data (Hartig, 2009; Glavic, 2007), recording the owner and the creation and modification time for provenance. Various query mechanisms such as SQL, SPARQL, and proprietary APIs are supported for different provenance data storage solutions. However, for a document or content-centric service process system where the activities of processes are responsible for content manipulations and changes, a domain-specific should be defined in a content-centric way to capture provenance information at process level and, thus, to support provenance-based process adaptation. Our system is a hybrid approach, which supports both content-oriented and process-oriented provenance requirements, such as content and process activity access control. Moreover, the provenance query is integrated into a process customization policy model to enable provenance-based process adaptation.

Conclusions

Learning content is more and more part of a systematic creation and consumption process that is enabled and governed through a workflow or process engine. As this is a process often across platforms and locations, involving different actors including authors, instructors and learners, quality control is important.

We have proposed a notation for the description of quality and governance constraints for learning content processes. This is a domain-specific content constraints model, here applied to processable learning content. The content is of specific types – to enable processing we have assumed RDF/XML. Processing activities are content creation, delivery and consumption oriented. The categorisation of constraints is specific to the different types of quality and governance constraints for learning content processing. A layered, modular information model covering content, processes and constraints facilitates its implementation in a wider interoperable content integration system. Interoperability is a critical driver in the application context. PROV has played a critical role, for the monitoring and recording as well as supporting the adaptivity for domains (here learning workflow processes for creation and consumption). This application serves as a template for more specific constraints and policy definitions. Together with a user-based customisation architecture, workflow processes can be adapted to meet organisation-specific needs, e.g., for the different types of providers and courses in the educational sectors.

Acknowledgements. This material is based upon works supported by the Science Foundation Ireland under Grant No. 07/CE/I1142 as part of the Centre for Global Intelligent Content (www.cngl.ie) at DCU.

References

Ardagna, D., Comuzzi, M., Mussi, E., Pernici, B., & Plebani, P. (2007). Paws: a framework for executing adaptive web-service processes. IEEE Software 24(6):39-46

Davidson, S.B. & Freire, J. (2008). Provenance and Scientific Workflows: Challenges and Opportunities. ACM SIGMOD Intl Conference on Management of Data.

Erradi. A. (2008). Policy-Driven Framework for Manageable and Adaptive Service-Oriented Processes. PhD thesis. The University of New South Wales.

Glavic, B. & Dittrich, K.R. (2007). Data Provenance: A Categorization of Existing Approaches. 12th GI Conference on Database Systems in Business.

Hartig, O. (2009). Provenance Information in the Web of Data. Workshop on Linked Data on the Web.

Holohan, E., Melia, M., McMullen, D., & Pahl, C. (2006). The generation of e-learning exercise problems from subject ontologies. Sixth International Conference on Advanced Learning Technologies, pp. 967-969.

Lei, X., Pahl, C., & Donnellan, D. (2003). An evaluation technique for content interaction in web-based teaching and learning environments. IEEE International Conference on Advanced Learning Technologies.

Melia, M. & Pahl, C. (2009). Constraint-based validation of adaptive e-learning courseware. IEEE Transactions on Learning Technologies, 2(1): 37-49.

Murray, S., Ryan, J., & Pahl, C. (2003). A tool-mediated cognitive apprenticeship approach for a computer engineering course. IEEE International Conference on Advanced Learning Technologies.

Pahl, C. (2002). An Evaluation of Scaffolding for Virtual Interactive Tutorials. 7th Int. Conference on E-Learning in Business, Government and Higher Education. Montreal, Canada. AACE.

Riegen, M. von, Husemann, M., Fink, S., & Ritter, N. (2010). Rule-based coordination of distributed web service transactions. IEEE Trans on Serv Comp, 3(1):60-70

Rockley, A., Kostur, P., & Manning, S. (2003). Managing enterprise content: A unified content strategy. New Riders.

W3C: PROV-O (2014). The PROV Ontology. Retrieved, 10 Mar 2014 from http://www.w3.org/TR/prov-o/.

Wang, M.X., Bandara, K.Y. & Pahl, C. (2010). Process as a Service - Distributed Multi-tenant Policy-based Process Runtime Governance. IEEE International Conference on Services Computing SCC 2010.

Wolter, C. & Schaad, A. (2007). Modeling of task-based authorization constraints in BPMN. Business Process Management, pp. 64-79.