REGULAR PAPER

GoRIM: a model-driven method for enhancing regulatory intelligence

Okhaide Akhigbe[3] · Daniel Amyot[2] · Gregory Richards[1] · Lysanne Lessard[1]

Abstract

Regulators are under constant pressure to demonstrate if and how the regulations they administer, which impose many requirements on various systems and processes, achieve intended societal outcomes. Traditionally, regulators have relied on impact assessments, risk analysis, and cost–benefit analysis to assess compliance with regulations. These methods, however, are effort and time intensive and focus on the efficiency of regulatory processes rather than on the effectiveness of the regulatory initiatives meant to improve compliance to regulations and the latter's impact on intended societal outcomes. Goal-oriented modelling and data analytics approaches provide the basis for the development of more sophisticated methods and tools to better address the needs of regulators. This paper introduces the goal-oriented regulatory intelligence method (GoRIM), which enables effective management of regulations through modelling and data analytics. Through continuous monitoring, assessing, and reporting on efficiency and effectiveness aspects, GoRIM is meant to facilitate the analysis of feedback loops between regulations, regulatory initiatives, and societal outcomes. To demonstrate the applicability and perceived usefulness of GoRIM in addressing the first feedback loop between regulations and initiatives, we evaluated it through three case studies involving regulators from different contexts, with positive results. GoRIM extends the concept of regulatory intelligence beyond the analysis of compliance. It also provides practical guidelines and tools to regulators for making, in a timely way, evidence-based decisions related to the addition, modification, or repeal of regulations and related regulatory initiatives. In addition, GoRIM helps better identify software and information needs for enabling such decisions.

Keywords Data analytics · Evidence-based decision-making · Goal-oriented modelling · GRL · Regulations modelling · Regulatory intelligence

Communicated by Silvia Abrahao.

✉ Daniel Amyot
   damyot@uottawa.ca

   Okhaide Akhigbe
   oakhigbe@uottawa.ca

   Gregory Richards
   richards@telfer.uottawa.ca

   Lysanne Lessard
   lysanne.lessard@telfer.uOttawa.ca

1   Telfer School of Management, University of Ottawa, Ottawa, Canada

2   School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Canada

3   Business Consultant with the Government of Canada, University of Ottawa, Ottawa, Canada

1 Introduction

Regulations are an important mechanism for achieving desired societal outcomes related to the safety, security, and prosperity of citizens and communities [1, 2]. To achieve these objectives, regulations are accompanied by various regulatory initiatives such as programs and enforcement actions designed to influence social behaviours and ensure compliance [3]. While regulations are unavoidable in today's societies and can help ensure minimal system qualities, they can also create a burden for organisations who must respond to new or frequently changing requirements [4]. For example, the cost of complying with regulations in the United States of America has been estimated in 2014 at 10,000 USD per American worker [5]. Regulators are therefore constantly pressured to demonstrate if and how the regulations they administer achieve desired objectives at adequate costs [6]. This is an ongoing challenge, since regulators are also

responsible for evolving regulations as needed, requiring further assessment [1, 7].

Regulators typically manage regulations through a lifecycle approach consisting of three iterative phases: (1) the *make* phase in which relevant government policies are transformed into regulatory instruments; (2) the *operate* phase in which regulations and regulatory initiatives are applied; and, (3) the *review* phase in which regulatory instruments are evaluated in order to assess if they are meeting their objectives, or if they should be modified in order to do so. However, while data are generated through activities within the first two phases (e.g., compliance data), they are not currently being systematically leveraged to support the review phase [1]. This may be due to the lack of adequate software support and uncertainty regarding data requirements for that phase [8].

The evaluation of regulations has so far been mostly limited to methods focusing on administrative efficiency and cost [4], the regulator's political accountability [9, 10], or the appropriateness of the governance framework involved [11, 12]. Such methods, in particular those limited to evaluating costs, are inadequate for assessing societal impact and effectiveness [13]. Hence, they provide insufficient support for the review phase of the regulatory cycle, which also requires the exploration of relationships among regulations and regulatory initiatives, compliance levels, and intended societal outcomes. Moreover, while some impact analysis methods explore societal outcomes, they often require slow, complex, and expensive studies that cannot be undertaken frequently and hence cannot provide timely results to decision makers [1]. As such, it is common practice for regulators to settle for the performance of programs (a kind of regulatory initiative) that are used to support regulations, instead of the performance of the regulations themselves.

Logic models are graphical models that regulators often use to demonstrate the performance of programs supporting regulations [14]. A logic model is a depiction of causal pathways that presents the shared relationships among the resources, activities, outputs, outcomes, and impact of the program [15]. While logic models describe how programs operate, and while they can be used to coarsely evaluate the societal outcomes of the *program*, they do not necessarily reflect the specific requirements of the regulations they support. Moreover, they do not quantify relationships between model elements in terms of contributions to the success of the program and, indirectly, to the performance of the regulation. As such, regulators are not able to make informed and evidence-based decisions about regulations using logic models.

The term *regulatory intelligence* refers to the gathering, collating, and analysis of regulatory data from multiple sources to assess and ensure an organisation's compliance to applicable regulations [16, 17]. The fundamental concepts of regulatory intelligence can be applied to evaluating the compliance level of a target population by leveraging available regulatory data and business analytics tools [18]. The availability of increasingly large datasets and tools such as data analytics systems provide an opportunity to expand the concept of regulatory intelligence to better assess regulatory performance and management of the regulatory cycle.

The objective of this research is to propose a model-driven, tool-supported method for managing the regulatory cycle focused on the review stage discussed above. Such a method must enable the iterative assessment and evolution (i.e., addition, modification, or repeal) of regulations and regulatory initiatives in relation to the societal objectives they are meant to achieve. It should also provide timely and evidence-based decision-making capabilities to regulators, as well as transparency into their decision-making processes to facilitate communication with stakeholders such as government agencies and regulated parties.

To achieve this objective, we propose the *Goal-oriented Regulatory Intelligence Method* (GoRIM), a model-driven method that incorporates goal modelling (from the requirements engineering field) and analytics to support the timely and continuous assessment of regulations. GoRIM extends the current body of work on regulatory intelligence by modelling and supporting the analysis of relationships among the requirement clauses of a regulation, the regulatory initiatives that have been implemented to support this regulation, and the societal outcomes achieved by the regulation. The method provides novel support to regulators by enabling timely and evidence-based decision-making for regulatory intelligence, hence the ongoing evaluation and evolution of regulations and related regulatory initiatives in order to achieve intended societal outcomes. In this paper, we explore and test the development of GoRIM to address the first feedback loop: between the regulation itself and the impact of regulatory initiatives meant to ensure compliance.

GoRIM relies in part on the *Goal-oriented Requirement Language* (GRL), a modelling language part of the User Requirements Notation (URN) standard [19, 20]. URN is a language used to model and analyse requirements with goals and processes. Goal models specify the goals of stakeholders, their decomposition structure, and how they contribute to each other. We focus on goal-oriented modelling in this work because of the need to relate regulations to their objectives in order to assess their effectiveness [8]. URN has been successfully used to model various laws and regulations [21], to measure their compliance in domains such as healthcare and aviation security [22]. Moreover, GRL includes an *indicator* concept that is useful for measuring and analysing the performance of goals based on real data [23]. However, the use of GRL and URN has up to now been limited to modelling and assessing compliance with regulations [24]. GoRIM also integrates data analytics tools to leverage capabilities that are

important to professional users, such as visualisation, interactive exploration, and report creation [25].

This paper contributes: (1) a definition of GoRIM and its steps, (2) a description of modelling and analytics software needed to support GoRIM, instantiated with one particular combination of academic and industrial tools, (3) three case studies that illustrate and evaluate GoRIM in different regulatory contexts, while highlighting several software and information challenges and solutions related to GoRIM's steps, and (4) further evaluation based on a survey of GoRIM users, with positive outcomes. While these contributions will be of interest to policymakers and regulatory organisations in need of improved methods to monitor and assess their regulations, they will also be useful to software developers and researchers working in regulatory compliance or business intelligence.

This paper is organized as follows. Section 1 provides background on the current body of knowledge and existing challenges in regulatory intelligence. Section 3 describes our use of the Design Science Research (DSR) methodology in this research. Section 4 presents GoRIM and the process by which it can be used by regulators. The application of GoRIM is demonstrated in Sect. 5 through three case studies undertaken in different domains. The first case study describes GoRIM's use to analyse a regulation and related regulatory initiatives; the second one focuses on the use of GoRIM to assess the performance of a regulatory initiative; and the third one highlights how GoRIM can be applied in contexts with insufficient data. Section 5 presents the evaluation of GoRIM through qualitative and quantitative results, and the implications and limitations of the research are discussed in Sect. 6. Section 7 presents related work, and the paper concludes with a discussion on plans for future work in Sect. 8.

2 Extending regulatory intelligence

This research is motivated by the need to better support regulators in reviewing and evolving regulations in a manner that is both timely and based on evidence. Regulatory activities happen in a lifecycle composed of three phases that occur concurrently: (1) *make*, in which regulations are created; (2) *operate*, in which regulations are implemented and enforced; and (3) *review*, in which regulations are reviewed to evaluate their performance and revised as needed to ensure that they remain relevant [3]. However, a meta-analysis of information technology artifacts used to support regulatory compliance showed that while methods, guidelines, and systems are generally available to support compliance modelling and compliance checking tasks that are accomplished in the operate phase, very little support exists for compliance analysis and compliance enactment tasks accomplished in the review phase [8].

While the meta-analysis focused mainly on tasks conducted by regulated parties, these results reflect issues experienced by regulators in relation to understanding the data requirements needed to assess a regulation's performance [26] and collecting the large volume of data generated throughout the operate phase [27]. These issues translate into challenges in exploiting these data to measure the performance of a regulation and in a general lack of understanding of how to review regulations to demonstrate performance [1, 27, 28]. Methods have been proposed to measure regulatory performance (i.e., whether they are effective and efficient), including randomized experiments, quantitative observational studies aiming to identify causal relationships between a regulation and its impacts, and qualitative studies that explore multiple impacts of a regulation [1]. However, these methods are predicated on large-scale and longitudinal research designs that can be expensive, time-consuming (sometimes taking years), and difficult to implement. Hence, while these methods may provide reliable insights to policymakers, they do not support regulators' needs for timely decision-making based on frequent or continuous monitoring of available data. As a result, regulators currently lack support for timely iterating regulatory cycles in a manner that is based on evidence, which is required for accountability and transparency purposes [27].

Regulatory intelligence is a promising approach to addressing this gap. From a *regulated party*'s perspective, regulatory intelligence is widely used by companies in the pharmaceutical sector, a sector characterized by ever-increasing and changing regulatory requirements [17, 29]. In this approach, organisations systematically gather and collate regulatory data from multiple sources in order to analyse them for the purpose of remaining compliant [16, 17]. A core concept of regulatory intelligence is the creation of a feedback loop between the regulatory environment and a firm's regulatory strategy. The feedback loop is represented by the use of data collected and analysed from the regulatory environment as input into the regulatory strategy of a firm. It is possible to create a similar feedback loop from the *regulator's* perspective, by linking a regulation and its related initiatives to compliance data through dimensional models that enable regulators to monitor the compliance level of a regulated population using business intelligence tools [18]. However, the methodology proposed to implement this approach remains focused on compliance monitoring and does not provide the means to measure and update regulations [18].

We thus propose to expand regulatory intelligence to better support regulators' activities beyond compliance monitoring by establishing linkages to a regulation's intended societal outcomes, which further exploits dimensional models' ability to support performance-related measurement and decision-making. Such expanded regulatory intelligence should consider the regulatory initiatives used to promote and
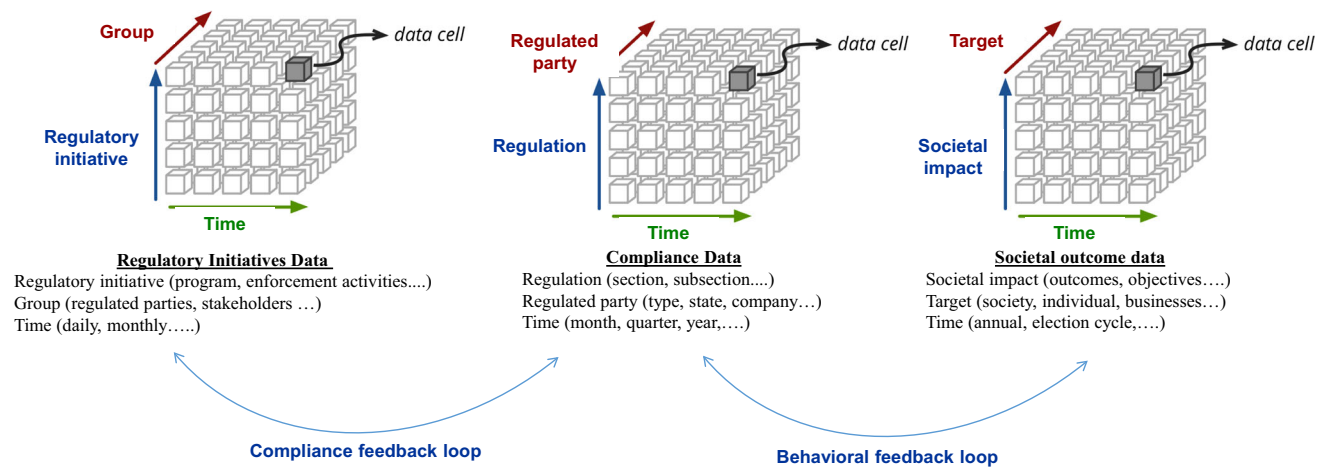
Fig. 1 Types and dimensions of data generated through the regulatory process

enforce regulations and to educate regulated parties, together with measures of whether societal outcomes have been achieved. Figure 1 depicts dimensions that form the basis of an extended regulatory intelligence framework encompassing regulation initiatives, compliance to regulations, and societal outcomes. Each dimension corresponds to a given phase within the regulatory cycle: regulatory initiatives data generated through the "make" phase, compliance data generated through the "operate" phase, and social outcome data required for the "review" phase. The exploration of relationships among these dimensions is key to supporting iterative decision-making within the regulatory cycle.

The framework thus specifies (1) a compliance feedback loop linking compliance and regulatory initiatives, and (2) a behavioural feedback loop linking compliance to intended societal outcomes. For these loops to support evidence-based decision-making, data from the regulatory process must be captured, structured in a coherent manner, and analysed using relevant software and statistical techniques. A first step in building these feedback loops is thus to map the rules and interrelationships among these rules contained in a regulation so that the feedback loops are specific to the different sections and sub-sections in each regulation. The analysis generated through such mapping and subsequent analysis can then enable regulators to make informed decisions about interventions that could improve a regulation and its supporting regulatory initiatives.

In order to move from this conceptual framework to tools able to better support regulators in reviewing and evolving regulations, we propose a comprehensive method (GoRIM) that can be supported by existing open-source and commercial software, with some adaptation and configuration.

3 Research methods

This research is situated within the Design Science Research (DSR) paradigm, which guides research focused on the introduction of innovative artifacts, whether in the form of models, methods, or information systems, that can both solve problems experienced by domain practitioners and generate generalizable design knowledge [30, 31]. DSR is increasingly used in engineering disciplines and computer science because it helps researchers move between problem and solution domains and between theory (or extant knowledge) and practice in an explicit and transparent manner [32]. There are varied ways to move among these dimensions; we followed one of the more comprehensive paths, thereby starting with a problem instance and abstracting to a general problem and its solution, followed by the evaluation of an instance of that solution [32, 33].

3.1 Design science research methodology

We used the Design Science Research Methodology (DSRM) [34] to guide the creation and evaluation of GoRIM. DSRM offers conceptual principles, practice rules and a process for carrying out and presenting research in line with DSR principles. The main steps within the DSRM process are: (1) Problem identification and motivation; (2) Definition of the objectives of a solution; (3) Design and development of an artifact; (4) Demonstration of the use of the artifact; and (5) Evaluation of the artifact. This process is not linear but iterative, with the results of the demonstration and evaluation steps typically triggering modifications to the artifact [34, 35]. We briefly outline the activities conducted in each step.

The *first* and *second* steps are highly interrelated and correspond to the development of a solid understanding of the problem domain. In the first step, researchers should define the problem and justify the value of addressing it. In the second step, the objectives of a solution should be inferred from the problem definition and the knowledge of what is possible and feasible [34]. In this research, both steps were supported by systematic literature reviews on the tools and

methods available to regulators and informal conversations with some of them [8, 24, 27]. The problem being addressed is hence defined as the lack of support for the review phase of the regulatory cycle, and the objective of a solution is defined as the provision of a method that can support timely and evidence-based decision-making by regulators. A similar approach was adopted in other DSR projects in the field of software and systems modelling and design [36].

An artifact, or generic solution, is created in the *third* step. Moving from objectives to solution should rely on relevant extant knowledge or theory [30, 34]. The framework underlying GoRIM was based on existing practices in regulatory intelligence [17, 29] and preliminary proposals for applying regulatory intelligence to support the needs of regulators [18], as described in Sect. 2. Moving from this framework to the creation of a method that can be used by regulators in their daily activities relied firstly on a systematic literature review of the modelling methods used to support regulatory compliance [8]. Results of this review showed that goal-oriented modelling methods offered more benefits for modelling regulatory compliance than non-goal oriented modelling methods because they can: (1) model both the intent and structure of laws and regulations, rather than solely their intent; (2) capture a broader range of constructs including regulations, stakeholders, objectives, outcomes, processes, and their relationships; and, (3) facilitate monitoring and assessing these relationships using indicators that represent varied types of measure. The design of GoRIM also relied on existing goal-oriented approaches and tools that have been developed for the domain of regulations, while extending them to fully reflect our proposed framework for an expanded regulatory intelligence (see Sect. 8—Related work). GoRIM thus extends the application of goal-oriented modelling for the domain of regulations by expanding the scope of data collected and by integrating off-the-shelf, easily accessible business intelligence solutions for the analysis of these data. Our approach to the creation of a solution reflects the fact that a full cycle of DSR is often accomplished through numerous research projects carried out over time [37].

The *fourth* and *fifth* steps represent two levels of evaluation: (1) a demonstration that the artifact can solve one or more instances of the problem, and (2) the evaluation of how well the artifact provides a solution to the problem—in other words, how well the solution achieves the objectives defined in the second step [34]. While a full DSR cycle would encompass both levels of evaluation, the first level is considered critical for novel artifacts since it shows whether the solution is a candidate for adoption in practice before it is implemented [33]. Hence, evaluating a "proof-of-concept" in terms of its perceived usefulness, quality, and efficacy is taken to provide sufficient rigor for novel artifacts, while in-depth evaluation is to be expected in a research project focusing on the validation of an existing artifact [35]. In this research,

a multiple-case study was used to implement step 4, while surveys of key informants were used to support step 5. These two steps were intertwined to revisit step 3 and iteratively improve GoRIM as described in the next section. Since steps 4 and 5 were iteratively used to improve GoRIM and made use of proof-of-concept models, the evaluation performed in this research is considered to be formative [38].

3.2 Procedures for demonstrating, evaluating, and improving GoRIM

Several methods can be used to support the evaluation of a solution, including objective quantitative performance, satisfaction surveys, and client feedback [34]. This study relies on the involvement of key informants in order to evaluate GoRIM in terms of its applicability to real-life contexts and of its perceived usefulness for demonstrating the performance of regulations. *Key informants* are individuals considered to have specialist knowledge in a domain [39]. We recruited fourteen key informants playing varied managerial and analytical roles in three government agencies responsible for articulating, implementing, and managing regulations related to the environment and infrastructure domains. Each agency acted as a case study, leading to a multiple-case study research design that allows comparing cases to increase the generalizability of results [40, 41]. Moreover, applying GoRIM in the multiple-case study provided us with feedback that helped to refine GoRIM to its current state as described in Sect. 4. The study spanned a period of two years from July 2016 to July 2018 and each case is described in Sect. 5.

The following process was used to conduct the multiple-case study and key informant surveys, as well as to improve GoRIM:

(A) We met with each key informant at least once in person to understand each case's chosen regulation and regulatory initiative and to capture them as goal models. These initial meetings were followed by informal conversations over phone or email as needed to create representative models. We also asked key informants to provide us with sample data during this phase, which corresponds to the first and second steps of GoRIM (see Sect. 4.3).

(B) GoRIM's steps 3 to 6 were then applied. Hence, the models were populated using the sample data provided for each case, and analysed using an off-the-shelf data analytics software.

(C) Another set of meetings and interactions with key informants allowed us to validate the regulations and regulatory initiatives models, analysis results and insights obtained. This process was iterative, whereby the knowledge and learning we derived from one case study was applied to subsequent case studies to improve how

to create the model, populate the models with data, and carry out analysis to derive insights. The process of validating the models also allowed the refinement of GoRIM itself, in particular through an additional step as described in Sect. 4.3.

(D) Validated models were then provided to key informants as proof-of-concepts. At this time, key informants were asked to respond to a questionnaire on the applicability of GoRIM to represent of their regulatory context.

(E) Another set of meetings with key informants allowed the collection of additional data on the chosen regulation and regulatory initiative for their case, as well as the identification of performance-related questions of interest to key informants within each case. Models were analysed again using an off-the-shelf data analytics software to answer each set of questions.

(F) The resulting models and analysis results were then presented to key informants as a basis for answering a questionnaire on the perceived usefulness of GoRIM. This process was carried out concurrently across cases, as described in step C above.

3.3 Threats to validity

We used three of the four tests proposed by [41] to evaluate the findings we obtained from the qualitative data analysis of GoRIM's applicability and perceived usefulness in the three studies of our multiple-case study. The fourth test, internal validity, was not used since it does not apply to exploratory studies, which is the role played by key informants' feedback in this research. These tests were:

1. *Threats to construct validity* Construct validity refers to how well the studied parameter and their outcomes were relevant to the research questions addressed by our research [41–43]. In this research, this threat concerns whether we established correct measures of the concept we studied, namely the effectiveness of regulations. To mitigate this threat, we first demonstrated and validated GoRIM's ability to capture the regulatory context to the key informants before using GoRIM to address real regulatory performance issues. Hence, we were able to establish a chain of evidence on what GoRIM addressed. In addition, we provided a draft report of our presentations to the key informants to peruse and comment before any presentations were made to ensure that they were knowledgeable and in agreement.

2. *Threats to external validity* This refers to the ability of generalizing research findings obtained to other domains under different settings [41–43]. The external validity in this research reflects the extent to which we can generalize the research findings on GoRIM. To mitigate this threat, our research design involved applying GoRIM to a multiple-case study involving three studies. In addition, we applied analytical induction across all the studies by focusing on the review phase of the regulatory cycle where regulations are reviewed, and hence ensured that we had similar activities in all three studies [41, 44].

3. *Threats to reliability* Reliability refers to demonstrating that, if the operations of a study were repeated, the same results would be obtained [41]. When using qualitative data, the prerequisite for this criterion is to document case study research procedures in a transparent manner. To mitigate this threat, we created a research protocol and instruments ahead of data collection, including recruitment letters, consent letters, and an interview protocol approved by the authors' institutions' Research Ethics Board. Moreover, all collected data were kept in a case study database separate from the case study report, and each step of data analysis was documented. We provide a description of the thematic analysis used to generate insights from key informants' feedback (http://bit.ly/GoRIM-supp). Nevertheless, remaining limitations related to reliability are discussed in Sect. 7.2.

4 The goal-oriented regulatory intelligence method (GoRIM)

GoRIM is a model-driven method for assessing whether regulatory initiatives and regulations effectively support intended societal outcomes. GoRIM provides the procedures and tools that allow collecting, analysing, and reporting on behavioural and compliance outcomes for the purpose of monitoring and evolving regulations and related initiatives (see Fig. 2).

GoRIM makes use of goal modelling, using the Goal-oriented Requirement Language (GRL) and jUCMNav, a free and comprehensive Eclipse-based graphical editor for analysing and managing GRL models [45]. GRL allows analysing compliance and performance data in relation to the goals of a regulation, its supporting regulatory initiatives, and its intended societal outcomes with a common language. This unique capability helps to integrate varied compliance and performance data in a homogeneous dataset in order to further analyse the impact of a regulatory initiative on regulation compliance, and the impact of a regulation on intended societal goals. This analysis is carried out using off-the-shelf data analytics software that provides data exploration, visualisation, and reporting functionalities. The insights gained through data analytics can then be used to make decisions regarding the evolution and further evaluation of a regulation and its supporting regulatory initiatives. The next subsections describe GoRIM's artefacts, roles, and steps, which correspond to the main method concepts (respectively called work products, roles, and tasks) defined in OMG's *Software & Systems Process Engineering Meta-Model* (SPEM) specification [46].

Fig. 2 Overview of GoRIM, with steps for modellers (*M*), analysts (*A*), and regulators (*R*)
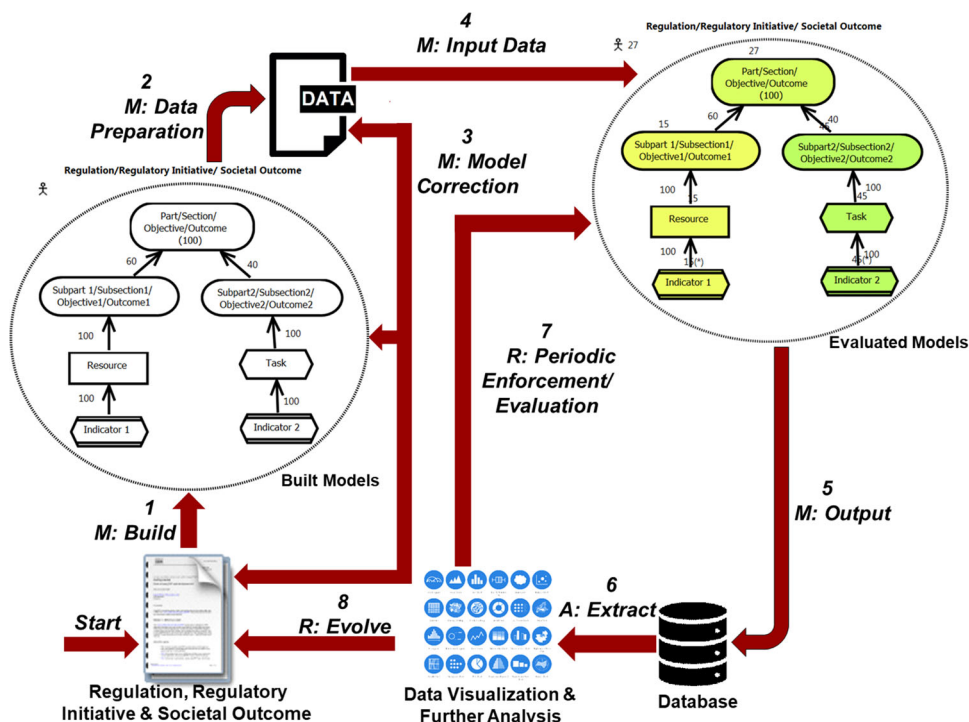
4.1 GoRIM's artefacts

GoRIM's main input artefacts are composed of descriptions of regulations, initiatives, and expected societal outcomes, supplemented by quantitative aspects usually not found in official regulatory documents. The second input is evidence needed to evaluate the models, especially data collected by regulators (e.g., through inspections or self-reported by regulated parties) or available publicly (such as national statistical organizations or open data [47]).

Intermediate output artefacts include three goal models (regulatory initiative, regulation, societal outcome), and their evaluation results based on evidence (e.g., per organisation at specific times for regulations), stored in a database. The models are *homomorphic* in nature, having the same views and created with the same language. This quality simplifies the learning of GoRIM and the integration of tools.

The main outputs of GoRIM are analysis results and visualisations, which can highlight performance along individual models (e.g., hot spots) together with correlations between models in support of the feedback loops in Fig. 1.

The metamodel capturing the essence of regulations and a mapping to GRL construct were provided by Shamsaei in her thesis [48] and are reused here to guide the creation of goal models for regulations. For regulatory initiatives, we created the metamodel in Fig. 3 based on the literature on the regulatory process on discussions with regulators. This metamodel describes the structural concepts and relationships of regulatory initiatives.

Table 1 describes the mappings between the concepts of this metamodel and GRL. Not all concepts need to be converted in our context; for instance, our analysis does not need to distinguish between immediate and long-term outcomes. This metamodel and this mapping help guide the manual transformation from textual initiatives to GRL models.

We have also observed from the literature and from discussions with regulators that the structure of societal outcomes is essentially identical to the structure of regulatory initiatives. Societal outcomes can also encompass activities, inputs, objectives, outcomes, outputs, regulatory initiatives, regulatory instruments, and stakeholders. This means that societal models can be created in GRL in the same way as regulatory initiatives.

4.2 GoRIM roles

Three different types of expertise are required to apply GoRIM: modelling, data analysis, and regulations. Each expertise corresponds to a GoRIM role, used to explain the method's steps in Sect. 4.3. These roles have been defined in line with the skills of personnel typically employed by regulatory organisations. Goal modelling expertise, however, would additionally be needed to implement the solution.

Modeller This role is responsible for creating and correcting the goal models of the regulations, regulatory initiatives, and intended societal outcomes (Steps 1 and 3). In addition to goal modelling skills, expertise in goal elicitation and negotiation is needed from discussions with regulators and
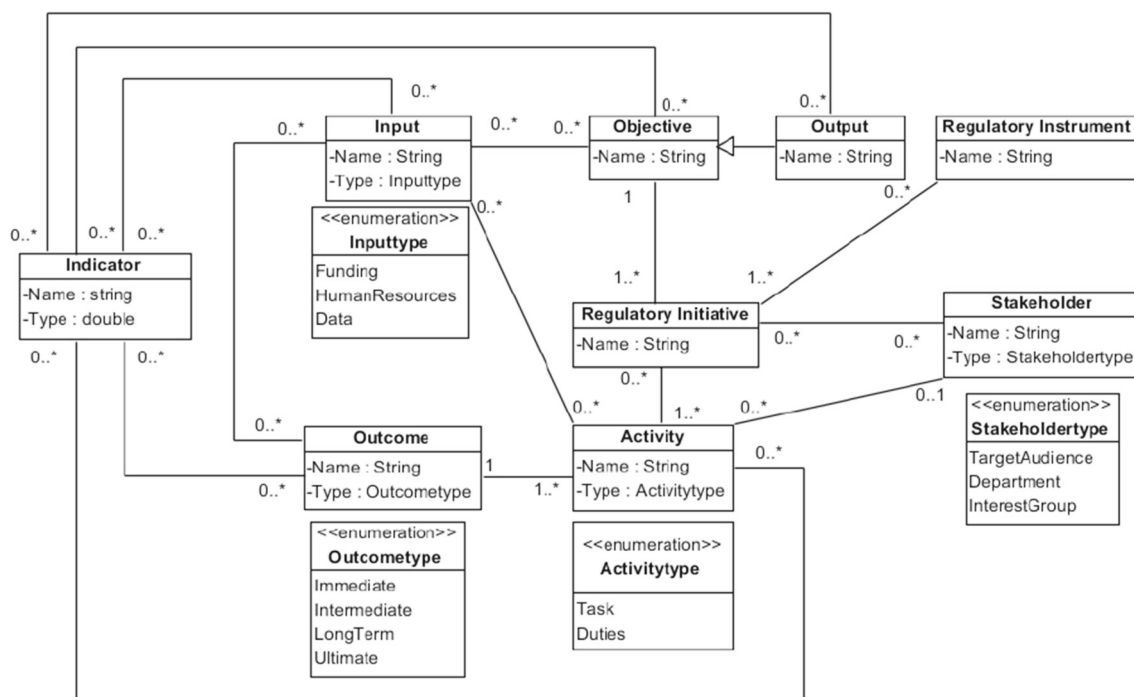
Fig. 3 Metamodel for regulatory initiatives and societal outcomes

other experts (e.g., how much a section of a regulation contributes to the satisfaction of its parent section, how indicators are defined, what are the societal objectives). The modeller also populates the models with data and exports the evaluations results (Steps 2, 4, 5), which requires familiarity with domain data and with databases. The modeller role can often be played by a Business Analyst or a Requirements Engineer.

Analyst This role is responsible for deriving different types of insights from the data that aggregates evaluations of models for multiple organisations and multiple initiatives, using data analytics software (Step 6). As such, this role requires analytics skills and expertise in the regulatory context. This role can be played by a Data Scientist or a Policy Analyst.

Regulator This role is responsible for making decisions on whether to change the period at which an organisation should be assessed for compliance (Step 7). For example, the next inspection or audit might be required earlier than usual due to non-compliance issues. The regulator can also evolve regulations, initiatives, and even societal objectives in case misalignment is discovered (Step 8). The regulator may also share best practices from a well-performing organisation with lower-performing organisations. The regulator role can be played by Policymakers and other people in charge of regulations and related initiatives.

4.3 GoRIM steps

We describe GoRIM's eight steps here and illustrate their application in the next section.

1. *Build* The modeller creates goal models of supporting regulatory initiatives and intended societal outcomes, using the metamodel in Fig. 3 and the mapping in Table 1. This is not automated and hence must be done manually, using the guidelines provided in [49, 50], inspired from Shamsaei's guidelines [48]. These goal models provide an abstraction mechanism that helps capture the parts of natural language documents that are important to support analysis. Many regulations are available in a structured format (e.g., in a database) that can be represented as a Comma-Separated Value (CSV) file importable by jUCMNav [51], which accelerates the creation of models. The resulting goal models are expressed in terms of intentional elements such as goals, tasks, and resources; key performance indicators (KPIs); and decomposition and contribution links between these elements (see Fig. 2). Quantitative information for these models, such as KPIs measuring compliance and the weight of their contribution links to goals capturing sections of the regulations/initiatives, must be elicited from regulators and other experts. Existing techniques for reaching consensus on such quantitative information in goal models [50, 52, 53] can be used here. Optionally,

Table 1 Mapping between regulatory initiatives (or societal outcomes) and GRL model elements

Term	Definition	GRL Element
Activity	Refers to an atomic process, i.e., the combination of actions that result in a certain set of products. This is the basic granularity used in the domain of regulatory initiatives	Task
Activitytype	Refers to the type of activity, e.g., task, duties	*Metadata on Task*
Indicator	Refers to a qualitative, quantitative or time factor used to measure	Indicator
Input	Refers to what is put in or operated on	Resource
Inputtype	Refers to the type of input, i.e., human resource, time, or money	*Metadata on Resource*
Objective	The purpose for which the regulatory initiative is required	Goal
Outcome	Refers to what an activity achieves	Goal
Outcometype	Refers to the type of outcome, i.e., immediate (short term), intermediate (medium-term), long-term or ultimate	*Metadata on Goal*
Output	Refers to what an activity produces directly or immediately. It also refers to a direct or immediate objective	Goal
Regulatory Initiative	Refers to actions, steps or things done that lead to actions on administering regulatory instruments	Task
Regulatory Instrument	Refers to tools Regulators use to change or modify behaviours in order to meet the objectives of each act or law, e.g., Regulation, Standard, or Code of Practice	Goal
Stakeholder	Party or person with an interest or concern in the objective or outcome of the regulatory initiative	Actor
Stakeholder-type	Refers to the type of stakeholder, i.e., department, target audience or interest group	*Metadata on Actor*

additional goal model validation techniques could also be used [54].

2. *Data preparation* Available data on regulation compliance, performance of the regulatory initiative and achievement of societal goals is identified, collected, and prepared by the modeller for input into the KPIs of the three goal models. Here also, three parameter values are defined for GRL KPIs; *Target* (corresponds to a full 100 satisfaction in GRL's standard [0…100] scale), *Threshold* (partial satisfaction of 50) and *Worst* (full dissatisfaction, i.e., 0).

3. *Model correction* The modeller uses sample data to evaluate the goal models to check whether the evaluations are as expected in chosen scenarios where the achievement of a given set of goals is known to the regulator. If the models are incorrect, then Steps 1 and 2 are revisited. In addition, models are checked for structural issues using well-formedness rules provided in jUCMNav [48]; the applicable OCL rules are listed in "Appendix A".

4. *Input data* When deemed correct, all prepared data are input into the models by the modeller. Using a GRL algorithm for computing goal satisfaction from KPIs, the satisfaction levels of each KPI is propagated to the other elements in the model. GRL models express satisfaction levels quantitatively on a [0.0…100] scale, as well as visually using colour-coding of intentional elements (greener when closer to the target value and redder when closer to the worst value). The resulting satisfaction values of model elements are the compliance levels for

the regulation model and the performance levels for the regulatory initiative and societal goal models. Such evaluations can be done within jUCMNav, or externally by exporting the GRL models to arithmetic functions usable in programs and spreadsheets [55]. The results of this stage provide a comprehensive view of regulatory compliance and performance to regulators. While some of this information would likely be known by a regulator today, discussions with key informants (who have specialist knowledge) confirmed that regulators rarely have access to all three views. Moreover, gaps in evaluated models resulting from missing data related to given KPIs can help regulators to identify data to be collected in the future.

5. *Output* Snapshots of different compliance and performance satisfaction values of model elements are exported and stored in a database. A snapshot refers to the evaluation of one model for one entity (group, regulated party, or target, as seen in Fig. 1) at a given time. The database then contains compliance and performance data structured along three dimensions (entities, time, and goal model elements).

6. *Extract* The analyst extracts relevant evaluated data to further explore it using an appropriate data analytics software, which provides data quality assessments and predictions (with confidence levels) together with a variety of visualizations. Data can be analysed for individual models (regulations, initiatives, societal goals) or for relationships (e.g., correlations) involving two models. New

insights such as parts of the regulation with good compliance, compliance trends, and performance levels for initiatives and societal goals can show whether there is a correlation between regulatory initiatives and compliance levels on one hand, and between compliance levels and societal outcomes on the other.

7. *Periodic enforcement/evaluation* Analytics results enable decision making by visualizing and demonstrating the relationships between and among elements. The regulator can exploit the above results to decide on short-term courses of action (e.g., warnings, penalties, or more/less frequent inspections of specific regulated parties) or to support sharing best practices between regulated organisations.

8. *Evolve* The regulator can also exploit the analytics results and new insights to justify the need for evolution (addition, modification, or repeal) of the regulation or related regulatory initiatives in order to better achieve intended societal outcomes. Specific scenarios can even be explored by evaluating tentative goal models that capture such evolutions against historical data [56].

For steps 6 and 7, it is required for the analytics software to be usable by non-scientists (e.g., regulators and policymakers), ideally with support for natural language queries (with automated proposals for candidate follow-up questions, to trigger relevant exploration of results) and default visualisations automatically selected (but modifiable). In addition to conventional filtering, slicing, and dicing capabilities, the software shall also be able to capture and reason about dimensions that reflect the structure of goal models (Fig. 1), which are not necessarily well balanced. The analytics software shall also be able to find correlations across datasets (e.g., for regulations, initiatives, and societal outcomes data), and not simply within datasets as the latter are often simple reflections of the goal models used to generate the datasets in the first place. IBM Watson Analytics (recently integrated to IBM Cognos Analytics [57]) was shown to meet these requirements [58], although other alternatives may exist or be developed. Watson Analytics is an online service that also runs many SPSS Modeler algorithms in the background during analysis, scores the models, and reports on those that performed the best, with statistical explanations of confidence levels. Note that correlations in Watson Analytics are based on the Pearson product-moment correlation index. In addition, algorithms and parameters are automatically selected and confidence levels computed based on the input data type and quality. Such capabilities are adequate for GoRIM's needs and for GoRIM users who are not data analysts.

More detailed and better illustrated guidelines are available online in [49].

5 Case studies

GoRIM was applied in three case studies related to Canadian regulatory agencies. This approach demonstrates the applicability of GoRIM in more than one type of regulatory domain: wildlife, mining, and safety. Each case study, presented in the next sub-sections, highlights a different use of GoRIM. The first case study focuses on model analysis using GoRIM. The second one focuses on GoRIM's use to model and analyse a regulatory initiative, in full compliance with the regulation. The third case study shows how GoRIM can be used in contexts with incomplete regulatory datasets. All three focus on the first loop of Fig. 1, between regulatory initiatives and regulations, without models of societal outcomes. These case studies served as proof of concepts for the application of GoRIM; their results were not intended to be used to influence actual decision-making by the regulatory agencies or to draw conclusions about the performance of the regulations or regulators.

5.1 Migratory bird regulation case study

In Canada, a permit is required to hunt birds considered as migratory birds. The *Migratory Birds Convention Act* (MBCA) [59] contains regulations that protects migratory birds, their eggs, and their nests from unauthorized hunting, trafficking, and commercialization [60]. The *Migratory Bird Regulations* (MBR) [61] are the regulatory instrument used to administer the MBCA. The Canadian Wildlife Service (CWS), a unit within Environment and Climate Change Canada (ECCC), administers the MBR using a regulatory initiative called *Migratory Birds Program* (MBP). While the objective of the hunting provisions of the MBR is to ensure alignment with the MBCA, the objectives of the MBP are to (1) *provide core information to ensure sound decision-making for setting conservation and protection goals*; (2) *enforce the MBR and the adoption of effective policies; and* (3) *champion actions to sustain healthy populations of migratory birds.*

This case study focuses on modelling the hunting provisions of the MBR and the activities that are involved in managing the MBP. In participating in the case study, the CWS was interested in identifying the contributions of the MBP to observed compliance with the MBR, i.e., whether the regulatory initiative actually improved compliance (compliance feedback loop in Fig. 1).

After creating the models, data on non-compliance with the hunting provisions of the MBR and data on performance of the MBP activities were fed to the models. Next, the data analysed through the models were exported into a data analytics tool, IBM Watson Analytics [57], to enable visualisation and further analysis of the effectiveness of the hunting provisions of the MBR and performance of the MBP.

Table 2 Ranking of enforcements activities to obtain contribution values

Rank	Enforcement activity	Average	Percentage
1	No action	0.0278	3
2	Warning	0.0556	6
3	Direction	0.0833	8
4	Ticket	0.1111	11
5	EPCO	0.1389	14
6	Ministerial order	0.1667	17
7	Arrest	0.1944	19
8	Prosecution	0.2222	22
Total		1	100

GoRIM steps were applied as follows:

Step 1 Model building: The hunting provisions sections of the MBR [62] includes General Prohibition: *Sub*sections 5*(4) and (11)*, Bag Limits: Sects. 7 and 8, Possession: *Subsections 10(1) and (2)*, Shipment: *Paragraphs 13(2)(a)* and *(c)*, Hunting Methods and Equipment: *Paragraphs 15(1)(c)* and *15.1(2)(a)* and *(b)*, Overabundant Species: *Subsections 23.1(2)* and *(3)* and *23.3(1)*, *subparagraph 23.3(2)(d)(iii), subsection 23.3(3)* and *subparagraph 23.3(4)(d)(ii)*. When the CWS enforces the hunting provisions of the MBR, the focus is on non-compliance rather than on compliance. Not all provisions in the hunting sections of the MBR are subject to enforcement actions, as some provisions are administrative in nature. Hence, to enforce compliance, the CWS uses the following enforcement activities to respond to incidents of non-compliance: *No Action*, *Warning*, *Direction*, *Ticket*, *Environmental Protection Compliance Order (EPCO)*, *Ministerial Order*, *Arrests*, and *Prosecution*. One or more enforcement activities can respond to a case of non-compliance. These enforcement activities were used as KPIs. To determine the contribution value of each KPI, with a go ahead from the key informants, the authors ranked the enforcement activities based on the perceived effort required to enforce non-compliance by the CWS as illustrated in Table 2.

To create a model of this regulation, the authors created a tabular representation of the regulation from its original textual format based on a Tabular Presentation Metamodel [51] and then transformed it into a goal model using jUCMNav. The goal model contained 156 intentional elements (35 goals, 1 resource, and 120 KPIs). Figure 4 illustrates the overview GRL model of the hunting provisions of the MBR and the "Bag Limits Section".

Following the same process, the authors manually created the GRL models of the MBP, the regulatory initiative used to administer the MBR. The MBP model includes *developing scientifically sound regulations through bird population surveys*, *banding of waterfowls and harvest surveys*

(also referred as "*Status Analysis*"). It also includes *compliance promotion,* an activity not carried out by the MBP. Interestingly, while the MBP is responsible for creating and administering the MBR and for planning and implementing compliance promotion activities, the Enforcement Branch of ECCC is responsible for enforcing compliance with the MBR. GoRIM proved useful in this scenario as it enabled the capture and linking of other related activities of the MBP in the GRL model. The resulting goal model contained two actors, 300 GRL intentional elements (7 goals, 7 tasks, 62 resources and 224 KPIs). This large model confirms GRL's ability to provide scalable and consistent representations of multiple views/diagrams of one model. This also accentuates jUCMNav as a robust tool for analysing goal models [63, 64]. Figure 5 gives an overview of the MBP goal model and related activities.

Step 2 Data preparation: The CWS provided data on non-compliance incidents recorded annually against the hunting provisions of the MBR for each enforcement activity from 2006 to 2016. These data are counts of violations and did not have any benchmark value set to identify what signifies a better/worse number of violations acceptable. The authors proposed using the average amount of violations recorded for each enforcement activity for the period of analysis (2006 to 2016) as a baseline, which was accepted by the key informants. This enabled determining the *target value* (value above the average amount of for the enforcement activity type)*, threshold value* (average amount for the enforcement activity type), and *worst value* (value below the average amount for the enforcement type) for the KPIs. Similarly, the CWS provided data on Harvest surveys (number of migratory birds of different species harvested), Population survey (number of migratory birds of different species counted), Waterfowl banding (number of migratory birds of different species banded), Compliance promotion (number of hunting summaries sold by province or territory), and Enforcement measures (number of occurrences of each enforcement activity) for the MBP and Enforcement Branch, from 2006 to 2016. There were also no benchmarks set to determine the value of these numbers. Hence, the authors also used averages as baselines. We defined the *target value* as the average plus an extremely small delta compared to the average, *threshold value* as the computed average, and the *worst* as 0.

Step 3 Model correction: To check the correctness of the models, data corresponding to a year were randomly selected from the provided data for input into the models of the MBR and MBP to check whether their evaluation was as expected. The models were checked against the URN well-formedness rules. The authors observed that the selected data contained some missing information. Whilst the jUCMNav tool does not allow the entry of blank or missing data, representing these missing data with "zero" values would influence the evaluated GRL strategies because zero is not the same as the
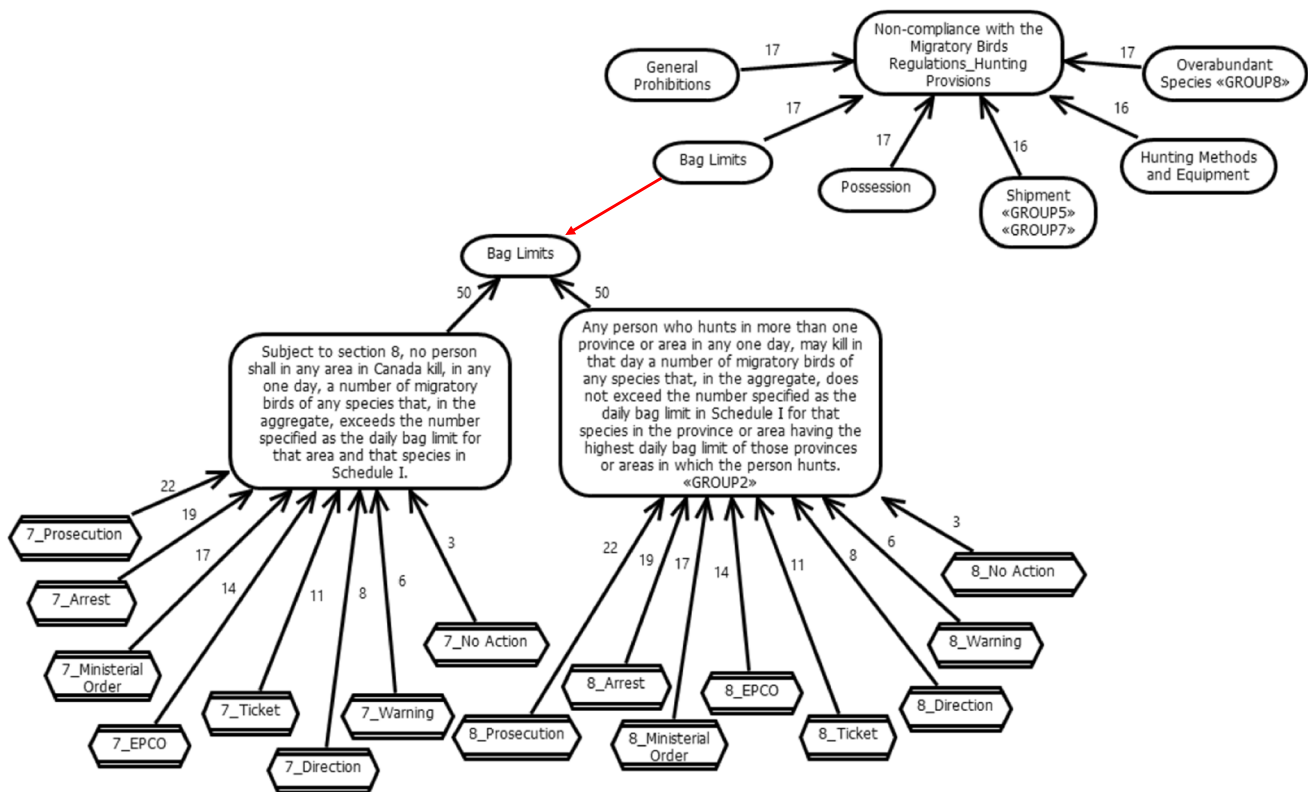
Fig. 4 Overview GRL model of the Hunting Provisions of the MBR and the Bag Limits Section
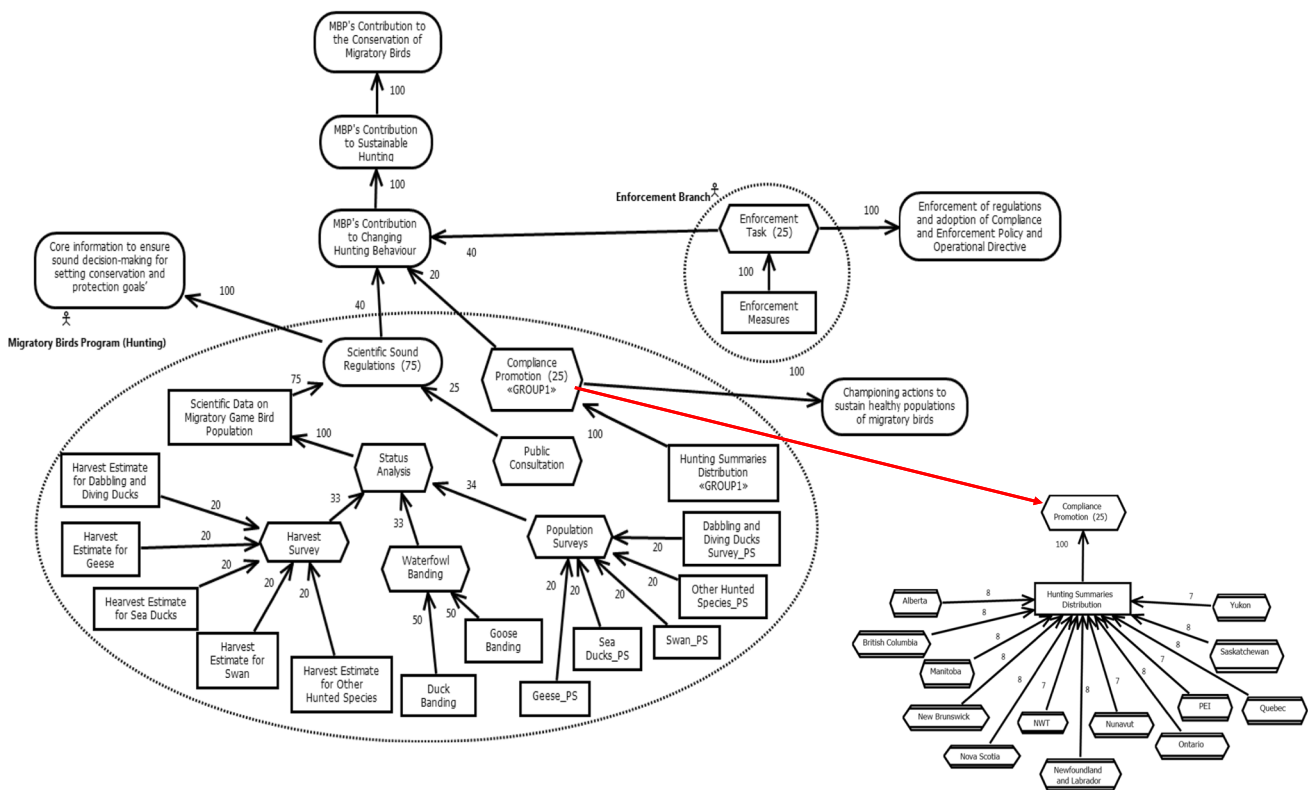

Fig. 5 Overview GRL model of the MBP, enforcement branch activities, and the compliance promotion activity
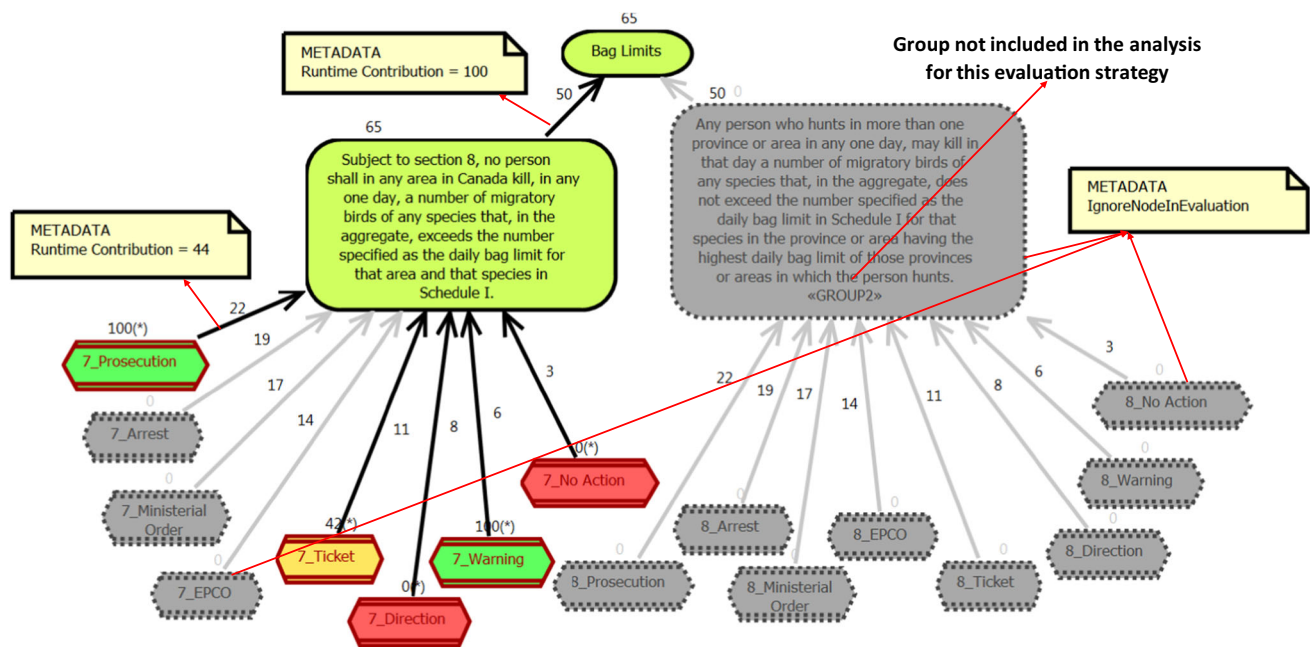
Fig. 6 Evaluated GRL model showing model elements automatically ignored for a given strategy

absence of a value. This scenario was addressed by ignoring the KPIs that have missing data from the GRL model during analysis, as if these KPIs were disconnected from the rest of the model, following Shamsaei's "Conditional GRL Algorithm" [48]. Occurrences of missing data (blanks) were replaced with the special character "#". This resulted in jUCMNav ignoring any corresponding GRL model element for that strategy and redistributing its contributions to remaining contributors dynamically during analysis. As illustrated in Fig. 6, we see model elements with the absence of values greyed out in the model; jUCMNav automatically distributed their contributions values to other model elements with data (see the Runtime Contribution comments on the left of the figure). This conditional algorithm enables the GRL models to produce reasonable analysis results in the absence of some data, a situation we came to realise is common in the regulatory context. After using this algorithm, it was observed that KPIs without data for a given year could have data in other years. As such, the metadata of the intentional elements these KPIs contribute to were tagged as a group (or type), and the GRL evaluation strategies were set not to include these intentional elements during the analysis. One such scenario is illustrated in Fig. 6.

Steps 4 and 5 Data input and output: After the models were deemed correct, all prepared data were fed to the models to produce yearly snapshots for 2006 to 2016 of the MBR and MBP models. Each of the eleven snapshots initializing the 120 KPIs in the MBR model and 224 KPIs in the MBP model were evaluated by jUCMNav to produce satisfaction

levels for each element of each goal model. These satisfaction levels were then exported as CSV files.

Steps 6 Extraction: The satisfaction values of the MBR and MBP models were imported into IBM Watson Analytics [57] for further analysis. This allowed answering the question of interested to the CWS, whether the MBP helped the MBR achieve compliance. The imported data was joined based on the "Year", which was common between both models [58]. The following is an example of the typical analyses requested by the regulator and enabled by GoRIM.

The annual hunting summaries set bag limits for each hunting season based on the status of migratory birds. Therefore, the regulator was interested in knowing whether status analysis (*developing scientifically sound regulations through bird population surveys*, *banding of waterfowls and harvest surveys*) has any effects on non-compliance with the bag limits provision of the MBR between 2006 and 2016, we queried "*Does status analysis have any effect on Bag Limits?*". The illustration in Fig. 7 enables putting this query in perspective using the GRL model.

The visualisation in Fig. 8, provided automatically by IBM Watson Analytics to answer the query, shows that the non-compliance level of the Bag Limits provision of the MBR ranges from 24 to 78 on a scale of 0 to 100. Within this range, when the average performance value of the Status Analysis activity is at its lowest (49), the level of non-compliance with the Bag Limit is low (32), although not the lowest. Similarly, when the average performance value of the Status Analysis activity is at its highest (66), the level of non-compliance with the Bag Limits is equally high (54 and 66), but not the
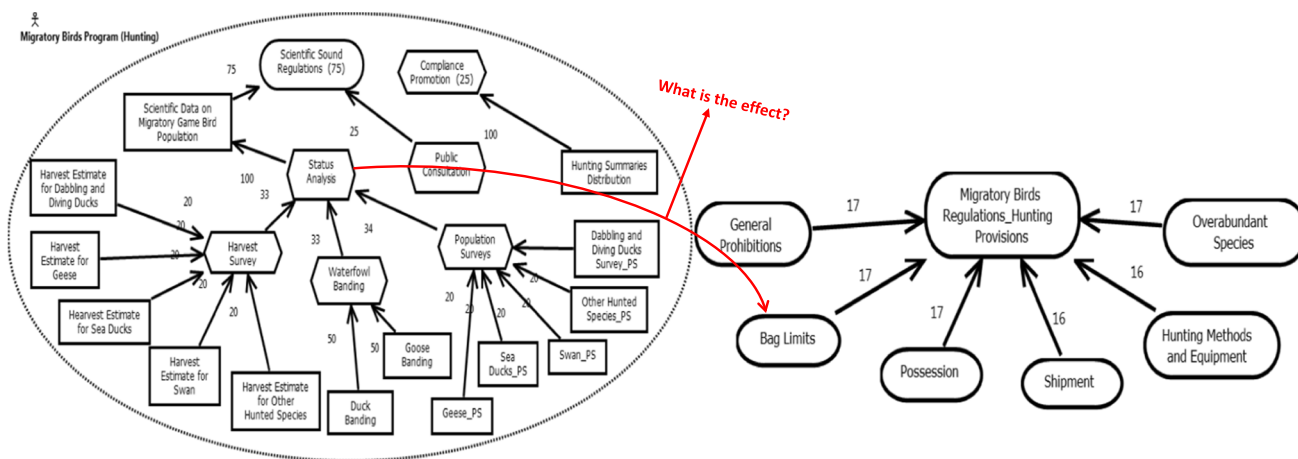
Fig. 7 Investigating relationships between the GRL models of the MBP Hunting activities (initiative) and of the Hunting provisions of the MBR (regulation)
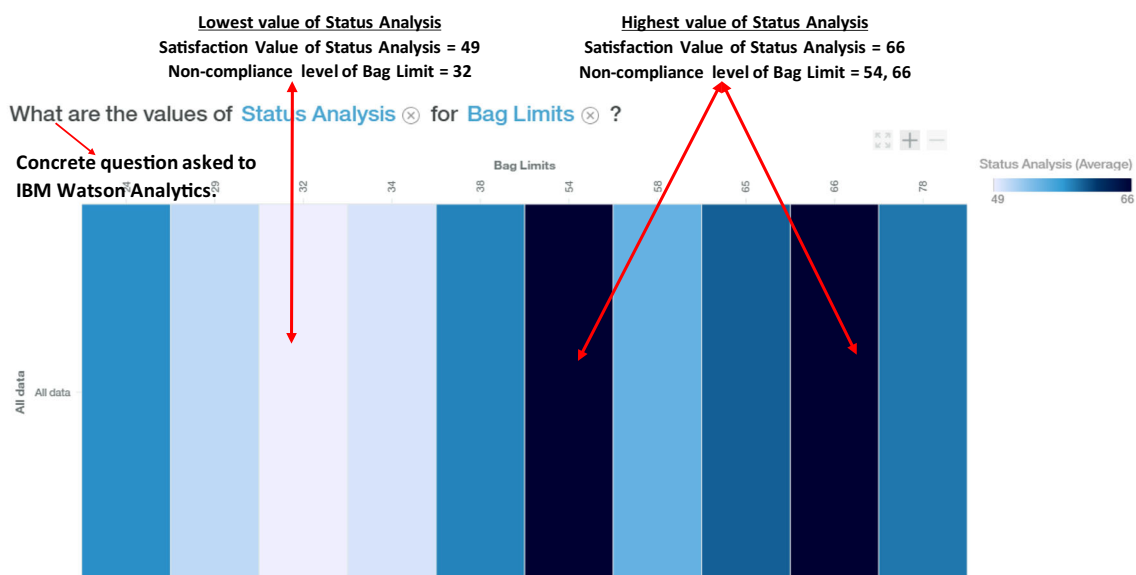

Fig. 8 Values of the Bag Limits provision of the MBR for MBP's Status Analysis activity

highest. This positive but weak correlation could imply that the Status Analysis activity of the MBP does not improve compliance with the Bag Limits provision of the MBR.

To further explore this insight, the performance of the MBP's Status Analysis activity and the level of non-compliance with the MBR's Bag Limits provisions between 2006 and 2016 were compared (Fig. 9). Results show a correlation where, on average, an increase in the performance of the Status Analysis activity in a year results in an increase in non-compliance with the Bag Limits provision in the following year. Thus, a positive correlation exists between the performance of the regulatory initiative in a year and the level of non-compliance with the regulation in the following year. In context, it implies that when the number of migratory birds is noticed to be high in a year after the Status Analysis

is done, the CWS increases the Bag Limit provisions of the MBR for the next year. This increase in Bag Limit provisions results in a decrease in non-compliance with the Bag Limit provisions in the following year (an increase in compliance).

The above analysis and others carried out using GoRIM provided the CWS with evidence confirming that the activities of the MBP are effective. Furthermore, the observed positive correlation provides data-driven evidence for the limits set in the Bag Limits provisions of the MBR.

5.2 Metal mining and effluent regulation case study

In the second case study, the analysis of the regulatory initiative using GoRIM enabled the Environment Protection Branch (EPB), a unit of the Industrial sectors, Chemicals and

Concrete question asked to IBM Watson Analytics.

Fig. 9 Comparison between values of the Status Analysis activity of the MBP and of the Bag Limits section of the Hunting provisions, by year

Waste (ICW) directorate, to identify drivers that they need to focus their resources upon to be effective. The ICW is the arm of Environment and Climate Change Canada (ECCC) responsible for the *Metal Mining and Effluent Regulations* (MMER) [65], which address Section 36 of the Fisheries Act [66]. The Fisheries Act prohibits the deposit of deleterious substances into waters in Canada frequented by fish, unless authorized by regulations. A deleterious substance is any substance that, if added to water, would degrade or alter its quality such that it could be harmful to fish, fish habitat, or the use of fish by people [66]. The *Environmental Effects Monitoring* (EEM) is the regulatory initiative the EPB uses to administer the MMER. The objectives of EEM are to (1) *assess the effects of effluent on fish, fish habitat, and the use of fisheries resources, and* (2) *evaluate the adequacy of the regulations to protect fish, fish habitat, and the use of fisheries resources.*

Despite compliance with the MMER by mines in Canada, mining effluents continue to have a negative impact on the receiving environments. In participating in the case study, the EPB was interested in a third-party assessment of the usefulness of the EEM to administer the MMER. The case study hence focused on modelling the sections of the MMER concerned with environmental effects monitoring and on reporting monitoring (i.e., monitoring whether results are provided to the government by the mines), as well as of tests involved in the EEM. After creating models of the MMER and EEM, data on compliance with the MMER, and EEM test results from mines regulated by the MMER were fed to the models. The analysed data from the models were

exported into IBM Watson Analytics to enable visualisation and further analysis of the compliance with the MMER and performance of the EEM tests.

Model of the *Environmental Effects Monitoring* (Sect. 7(1), (2) and (3)), and *Reporting Monitoring Result* (*Section 21(1), (3) and Section 22*) of the MMER were created. Key informants at the EPB provided KPIs for measuring compliance with these sections. The EEM stipulates that mines subject to the MMER must conduct monitoring tests to assess the impacts of effluent on a receiving environment. The monitoring tests strongly focus on biological monitoring studies (e.g., fish population survey, fish tissue analysis, and invertebrate communities' survey) and other analyses that bring vital information such as effluent and ambient water analyses, sub-lethal toxicity testing, and acute-lethality toxicity testing. EEM analyses the results of these monitoring tests to identify trends in effluent-related effects on fish, fish habitat, and/or use of fisheries resources, and to determine whether limits in the regulations are protective enough. "*The information obtained through the EEM can be used to determine the effectiveness of the MMER and provide a basis for determining the need for enhanced site-specific or national pollution prevention and control measures*" [65]. We used the indicators of the respective biological monitoring studies survey and the elements measured in the complementary measures studies as KPIs in the created GRL models of the EEM requirements.

Figure 10 shows the part of the MMER GRL model focusing on the Environmental Effects Monitoring sections and
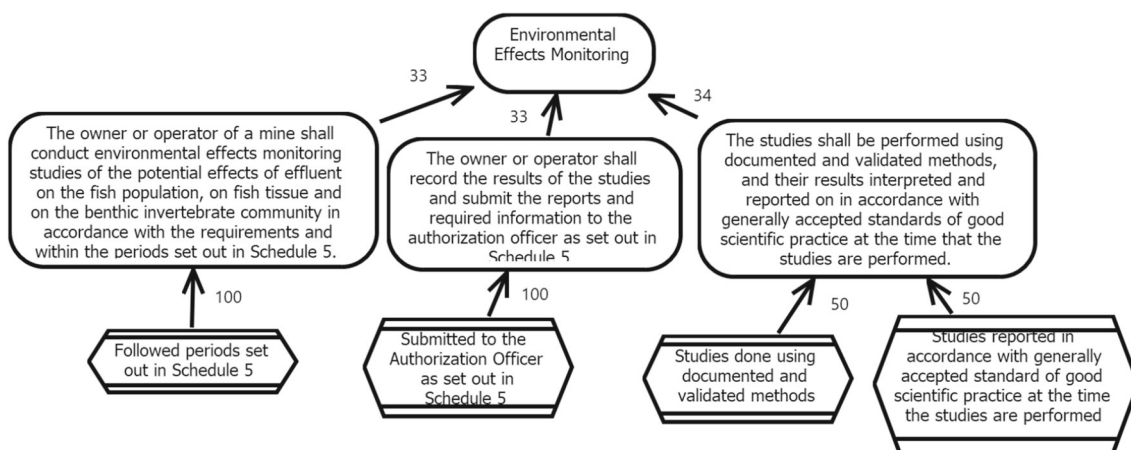
Fig. 10 GRL model of the Environmental Effects Monitoring sections of the MMER (regulation)
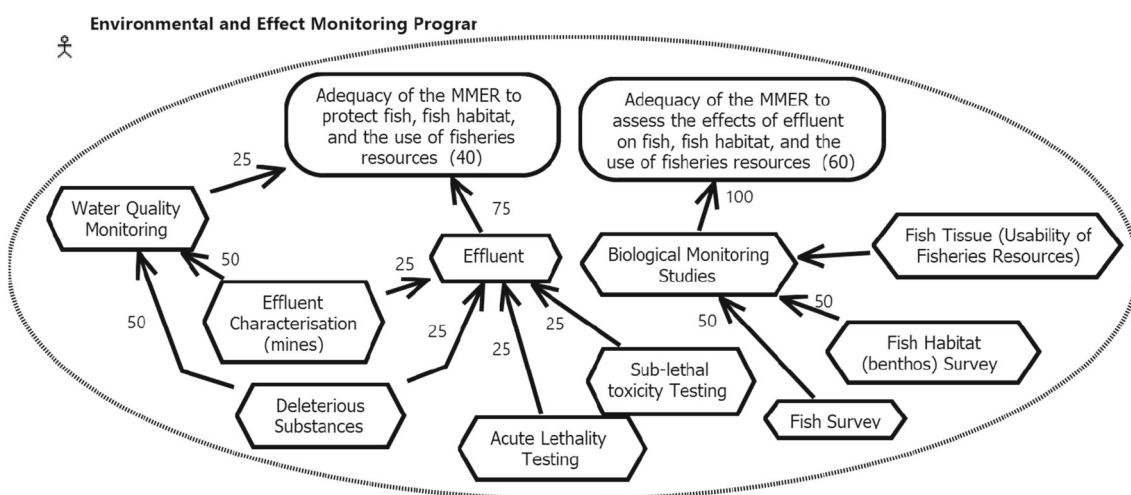

Fig. 11 Overview GRL model of the EEM requirements (initiative)

Fig. 11 gives an overview of the GRL model of the EEM requirements.

For the MMER sections, compliance was binary; mines were either in compliance or not. Hence, for the KPIs, the defined *target* and *threshold values* were the same. The *worst value* of each KPI was defined as not compliant, signifying that there was evidence that the mine did not meet the requirements of the sections. Regarding the EEM requirements, the different categorizations of test results were used to set the KPI's *target, threshold*, and *worst values*. Details on these studies and tests require an understanding of the water biology and are hence omitted here.

After defining the KPIs of the MMER and EEM, from the provided data of 25 mines, from 2002 and 2014, data for a year was selected at random, and fed to the models to check whether their evaluation was as expected. After correcting the models as necessary and when the models were deemed corrected, all prepared data were fed into the models to produce yearly snapshots for 2002 to 2012 of the MMER and EEM.

The GRL models in this case study were also large. There were 38 intentional elements (21 goals and 17 KPIs), and 201 GRL evaluation strategies spread across three diagrams in the MMER regulation's model. The EEM model had four diagrams, one for the complementary measures study and water quality study, one for the biological monitoring, one for the overview, and one for the conditions that address the absence of data. This EEM model had 57 intentional elements (2 goals, 15 tasks, 1 resource, and 39 KPIs), 210 evaluation strategies for the complementary measures study and water quality study, and 25 evaluation strategies for the biological monitoring study. Not all mines existed or had data for each year.

Analysis of the GRL models of the MMER showed full compliance. All 25 mines complied with the Environmental Effects Monitoring and Reporting Monitoring Result section of the MMER. To answer if the EEM was useful to administer the MMER as required by the EPB, we carried out the additional analysis explained below.
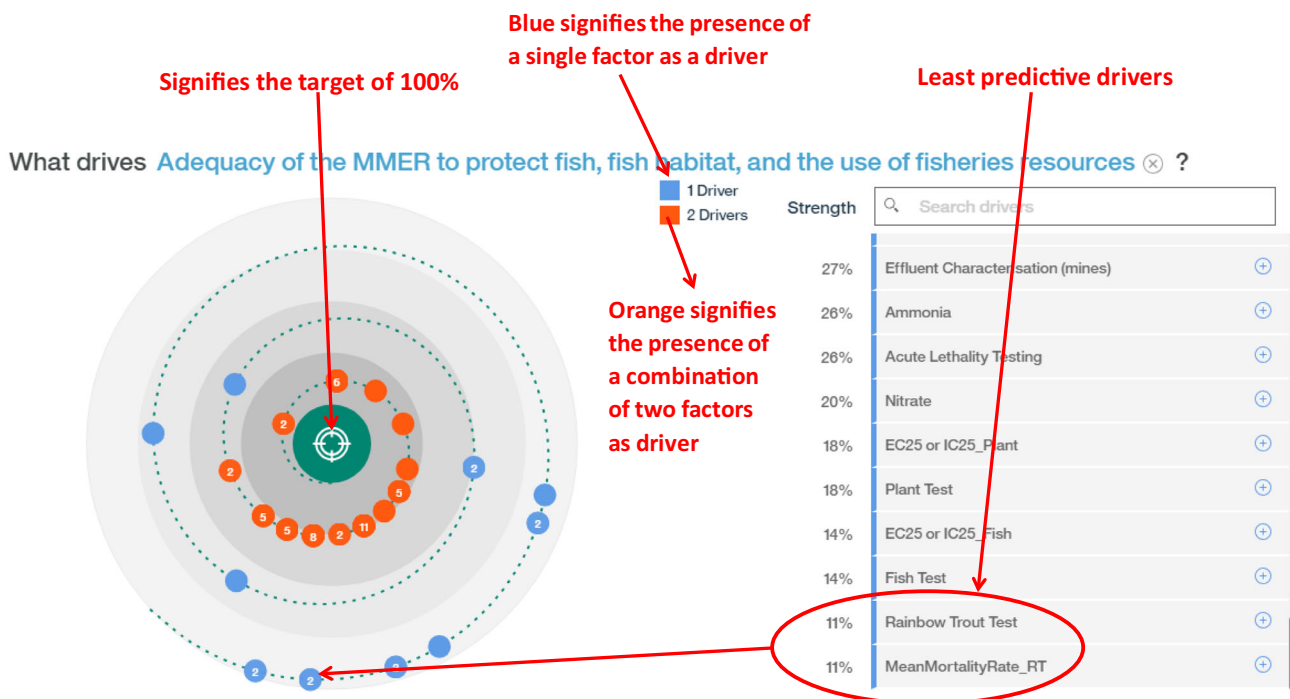
Fig. 12 Least-impactful drivers of the complementary measures study and water quality study from 2002 to 2014

The societal objective of the complementary measures studies and water quality studies is the "*adequacy of the MMER to protect fish, fish habitat, and the use of fisheries resources*". Based on the resources involved by mines in carrying out the required studies, it becomes useful to identify what contributes the least to these activities. Hence, our query for IBM Watson Analytics was "*What contributes the least to achieving the objective of the complementary measures studies and water quality studies?*". The visualisation selected by IBM Watson Analytics, illustrated in Fig. 12, shows that with a predictive strength of 11%, the "Rainbow Trout Test" and its KPI "MeanMortalityRate_RT", which assesses the mortality rate resulting from the exposure of rainbow trout to effluents, have the least predictive strength between 2002 and 2014. The values, while significant, are further out in the spiral graph, far away from the strongest predictive strength target of 100% found the middle of the spiral graph. This implies that they have the weakest correlation with achieving the objectives of the complementary measure studies and water quality studies.

Situating this result in the GRL model for the complementary measure studies and water quality studies, illustrated in Fig. 13, we can put this weak correlation of the "Rainbow Trout Test" in proper context in relation to the overall objective of the studies. This identification of "least-impactful drivers" is a feature derived from the evidence that triggered discussions among the EPB's key informants on the need to monitor these tests considering the time and resources

involved. Although the authors were not privy to the outcome of this discussion, this insight provided to the EPB could lead to potential cost-saving and time-saving opportunities by focusing on other indicators of higher impact.

5.3 Safety regulation case study

The final case study demonstrates the applicability of GoRIM in the safety domain with a government agency that prefers to remain anonymous. In this case, GORIM is applied in a situation characterized by very sparse data. The government agency oversees the safety of facility and infrastructure projects and as such, has twelve different regulations with which companies must comply. The agency ensures compliance through having companies self-report on their activities using five different tools. This self-reporting of non-compliance is called *Management Systems*. In addition, companies must report on (1) *Measures 1: A historic view of a company's performance in relation to reports of incidents (release of substances, injuries,* etc.*)*; and (2) *Measures 2: A prediction of a company's performance.* A mix of *Measures 1* and *2* provides an overview of a company's ability to meet the regulations. This mix, called the "*Safety, Security and Environmental Protection*" (SSEP) function, is the regulatory initiative employed by the government agency to administer its regulations. Consequently, this case study focused on modelling the Management Systems (self-reporting non-compliance with the twelve regulations) and the activities
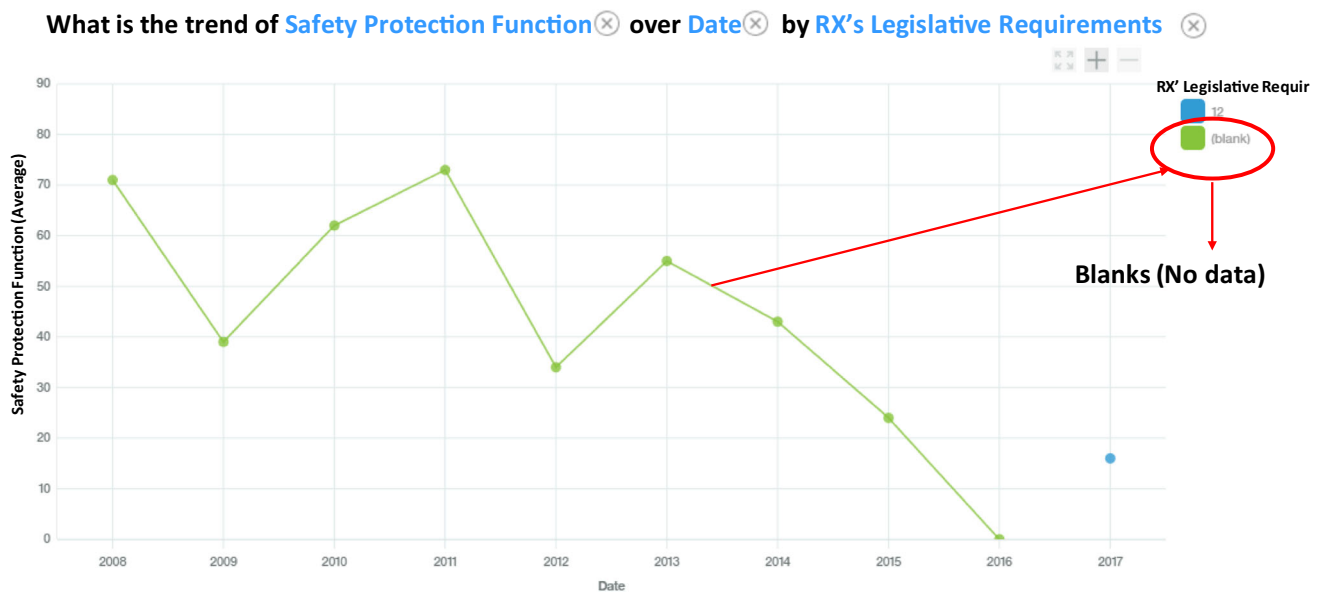
Fig. 13 GRL model showing the least-impactful drivers of the complementary measures study and water quality study from 2002 to 2014 in proper perspective

that constitute the SSEP function to determine if these activities help the Management Systems.

A GRL model of the twelve regulations was created. The five tools used by companies to self-report against each regulation were used as the KPIs in this model. As a norm, companies use these tools independently for each regulation. For example, the same tool can be used to report compliance to regulation 1 and regulation 8 independently. As in the first case study, KPIs were ranked based on their perceived importance in identifying non-compliance to derive their contribution values. A GRL model of the SSEP function, including its activities, was also created. The indicators already in use for *Measures 1* and *Measures 2* were used as KPIs in the SSEP model (with averages again used to determine the target, threshold, and worst value parameters for the KPIs).

Available data were prepared for input into the GRL models. Data were unfortunately missing for the twelve regulations from 2008 to 2015 and from 2015 to 2017 for *Measures 1* and *Measures 2*. As described in Sect. 5.1, occurrences of missing data were distinguished from zeros with the special character "#". Five groups in the model of the regulation and three groups in the model of the SSEP function, for conditions with no data, were created. The respective intentional elements in each of the evaluation strategies were tagged with these group labels. Shamsaei's "Conditional GRL Algorithm" [48] was then used to ignore all KPIs that do not have any data. The resulting GRL model of the twelve

regulations comprised 21 diagrams, 303 intentional elements (132 goals, 170 KPIs, and 1 resource), and 10 evaluation strategies. The GRL model of the SSEP function (the initiative) had 10 diagrams, 160 intentional elements (1 goal, 30 resources and 129 KPIs), and 10 evaluation strategies. Due to the presence of confidential information, these models cannot not be presented in this article.

Afterwards, the model satisfaction values for the *Management Systems* and performance of the *Safety, Security and Environmental Protection* function were exported. Unfortunately, owing to the absence of much data, results from analyses conducted with IBM Watson Analytics provided limited insights. For example, our query to IBM Watson Analytics to determine the effects of the SSEP function on the Management Systems resulted in the visualisation in Fig. 14. This figure shows that there appears to be a trend in the SSEP function resulting from the level of effort required to identify non-compliance with the regulations (the Management Systems) between 2008 and 2016. Despite the presence of multiple blanks (missing data), IBM Watson Analytics made this projection based on the whole dataset. Hence, it was impossible to ascertain what the trend really meant since the trend projections were based on the presence of multiple blanks (missing data) in the analysed data.

To address the issue of missing data and provide more meaningful insights on whether the self-reporting initiative helped to comply with the regulations, *synthetic data* were used. Some data were randomly generated under assump-

What is the trend of Safety Protection Function⊗ over Date⊗ by RX's Legislative Requirements ⊗

Fig. 14 Sample visualisation with sparse or no data in IBM Watson Analytics

tions discussed with the key informants, such as perceptions on self-reporting for the tools by companies. The synthetic data were combined with the sparse data and fed to the GRL models of the Management Systems and of the SSEP function, and the new model satisfaction values exported. Although analysis of the new satisfaction values provided some interesting insights, they provided relationships that required further analysis to provide a clearer picture. The key informants decided not to pursue the analysis further but were pleased with the potential of the approach provided enough data becomes available.

In conclusion, this case study gave us the opportunity to address a real-life scenario with very sparse data, but with the same requirement of supporting decision-making regarding the evolution of regulations. Our use of synthetic data allowed us to determine how this could be accomplished with GoRIM. The use of the synthetic data was informative to the government agency on the type of data they should collect to enable them to make decisions regarding the evolution of their regulations, and this proved to be of high interest to them as they saw an opportunity to propose benchmarks and simulate whether responses from companies were sufficient or not. Furthermore, our use of GoRIM in this case study also motivated the agency to start collecting data more rigorously on compliance, on the regulatory initiatives, and on societal objectives as they saw practical usefulness of what the data allowed them to assess. The agency also became better aware of technological opportunities and needs for processing and reporting on this data, which guided them in expressing better requirements for future software acquisitions.

6 Evaluation

6.1 Results of surveys on the applicability and perceived usefulness of GoRIM

This section presents the results of the two sets of surveys conducted to evaluate GoRIM, as presented in Sect. 3.2. These surveys (see http://bit.ly/GoRIM-supp for details) included questions on demography (i.e., responsibility level of the key informants, years of experience in the level, and their duties), Likert-scale format questions to measure responses about agreement or disagreement (strongly agree + + , agree + , neutral 0, disagree −, or strongly disagree –), and open ended questions. The key informant responses to demography and Likert scale questions (for both iterations) are shown in Table 3. Note that a subset of eleven key informants could answer the first survey, and only five answered the second survey (no survey results were discarded). The very last row indicates how many of these five key informants changed their mind towards more negative or positive scores.

The questions on the applicability of GoRIM were administered after GoRIM was used to capture and derive insights on the regulator's individual regulatory context. As illustrated in Table 3, there were two questions on the perception of the key informants on whether their respective agencies were monitoring and reporting on regulations (Q3) and were doing formal reviews of regulations (Q4). In both cases, the participants had a similar or *lower* score after the second iteration of GoRIM, suggesting that some people better understood what was actually feasible and required in these important regulatory activities and their own maturity in these areas.

Table 3 Key informants' responses to questions on the first and second questionnaires

Key Informant	Level	Yrs. In Level	Q3 1st	Q3 2nd	Q4 1st	Q4 2nd	Q6 1st	Q6 2nd	Q7 1st	Q7 2nd	Q8 1st	Q8 2nd	Q9 1st	Q9 2nd	Q10 1st	Q10 2nd	Q11 1st	Q11 2nd	Q12 1st	Q12 2nd
CWS01	Regulation	≥15	+	0	+	0	0	+	+	+	+	+	0	0	0	−	0	−	0	0
CWS02	Initiative	<5	+	+	+	0	+	+	+	+	+	+	−	−	0	−	−	0	−	0
CWS03	Regulation	<10	++		++		0		0		0		0		0		0		0	
EPB01	Initiative	≥15	+		+		+		+		+		0		+		+		+	
EPB02	Initiative	<10	+		+		+		+		+		0		+		0		0	
EPB03	Regulation	<15	+		+		+		+		+		0		+		0		0	
EPB04	Initiative	<5	++		++		−		−		−		−−		−−		−−		−−	
RX01	Both	<15	0	−	0	−	0	+	0	0	0	+	−	0	−	+	−	0	−	+
RX02	Both	<15	0	0	−	0	−	+	+	+	0	+	0	+	0	+	0	+	0	+
RX03	Both	<10	+	0	+	0	0	+	0	+	−	0	0	+	0	+	0	0	0	0
RX04	Initiative	<10	+		+		0		+		0		0		0		0		0	
Change counts			3−		4−, 1+		4+		1+		1−, 2+		3+		2−, 3+		1−, 3+		3+	

Two questions targeted the accuracy of the GoRIM models of the regulations (Q6) and regulatory initiatives (Q7), including their structure and KPIs. Most key informants were satisfied with both types of models, especially after the second iteration. A similar positive level of agreement (especially after the second iteration) was observed for Q8, on the extent to which GoRIM models show the objective of the regulations and the resulting relationship between the regulations and supporting regulatory initiatives. These answers hence support that GoRIM's GRL models are good and accurate representations of regulations and their related initiatives.

For Q9 on whether GoRIM is useful for measuring compliance levels of the modelled regulations, after the second iteration, two of the six key informants responsible for regulations (in the "Level" column) agreed while the other four were neutral. Question Q10 assessed whether GoRIM is useful for assessing how measured compliance levels translate to the objective of the regulation. Table 3 shows that five key informants changed their views after the second iteration of GoRIM: three were more positive and two were more negative, suggesting mixed feelings about the latest version of GoRIM for that aspect. One potential issue here is the fact that the societal objectives (which indirectly provide regulation objectives) were not modelled separately in the case studies. Similarly, question Q11 focused on whether GoRIM is useful for assessing how well the regulations meet their objectives. Four key informants changed their views after the second iteration of GoRIM; three had a more positive view and one had a more negative view, resulting in a fairly neutral level of agreement overall.

The last question (Q12) assessed whether GoRIM is useful for monitoring and reporting on the performance of the regulation. Table 3 shows that three key informants changed their views after the second iteration of GoRIM for a more positive answer. Overall, three key informants agreed that GoRIM is

useful here while seven were neutral and one strongly disagreed. There is hence a slightly positive agreement about the perceived usefulness of GoRIM for performance assessment, especially among the more senior participants.

The questions on the perceived usefulness of GoRIM were administered after GoRIM had been applied to answer performance questions of interest to key informants, and elicited responses on GoRIM's ability to demonstrate whether the regulation and regulatory initiatives were effective. These were open-ended questions. We used the thematic analysis method of Braun and Clarke [67] to identify themes in the feedback obtained from the key informants on the perceived usefulness of GoRIM questionnaires, with details of the analysis available at http://bit.ly/GoRIM-supp. The feedback and suggestions were used to improve GoRIM. In particular, the Data Preparation and Model Correction steps (i.e., Steps 1 and 2 in Fig. 2) were explicitly added. Modified models and analysis results were then sent to key informants to assess their perception of GoRIM's usefulness. We discuss this feedback in Sect. 6.2.

6.2 Key informant feedback

Key informants' feedback focused largely on the capabilities provided by the method and its accompanying tools, particularly in relation to the visual representations it generates and the analytical support it provides. The informants also identified perceived challenges to using GoRIM in professional contexts, including the lack of sufficient data available for analytics. We present a synthesis of key informants' feedback below, including supporting quotes where appropriate.

Capabilities provided by GoRIM

- *Comprehension and communication* A recurring comment made by key informants was the way in which GRL models, as visual representations, supported the description,

comprehension, and communication of the regulations they managed. The fact that models showed all regulatory components (related to regulations, regulatory initiatives, and intended societal goals) side-by-side was often mentioned as a source of comprehension, with one participant (RX01 in Table 3) mentioning that "This makes [the] range of the impact of the regulation easier to see and understand". Evaluated models generated through Step 4 were seen as a tool that could help various units within a regulation agency to communicate and align their efforts and goals.

- *Understanding of relationships among regulatory components and objectives* Several key informants discussed the visual capabilities of evaluated GRL models in terms of the visual and granular links that they established between regulatory activities and regulatory objectives. One participant (CWS02) stated "I think the model accurately describes the program and its various components in a useful and helpful visual.", while another (RX04) stated, "…the process has forced us to look at clear connections between regulatory provisions and our desired outcomes." These depictions, as well as the process of creating the models, were perceived to support thinking and reflection, and to encourage further analysis. This capability provides important support to the evolution of regulations and regulatory activities: "By modelling the connection between regulation, expected outcome, and compliance, it encourages/allows regulators to be more thoughtful about whether certain activities or regulations are actually useful or the best way to achieve an outcome", as participant (RX03) mentioned.

- *Identification of where to target efforts* Several key informants stated that GoRIM could help them identify areas for improvements in the regulations and regulatory activities that they manage, hence helping them to target their efforts towards these specific areas. While one participant (EPB01) mentioned that "… The potential value that I see is the possibility of better identifying the factors that could best be followed to indicate the degree to which the objectives of the regulations, those being minimization of effects on fisheries resources, are being achieved.", another (RX02) stated that "This information could help us in determining where to focus our compliance verification efforts and where to focus our regulatory improvement efforts". The notion of targeting regulators' efforts extends to data collection, as some key informants noted that identifying areas on which to focus implied identifying specific areas on which to collect more data. The analysis of the relations between regulatory components and objectives (not just their representation) was perceived as providing this capability. In particular, the ability to join varied datasets and to summarize them through evaluated KPIs was emphasized as key to targeting efforts.

Challenges to using GoRIM in professional contexts

Insufficient data Many key informants stated that a challenge to using GoRIM was the lack of sufficient data available for analytics purposes at their organisation, for example: one participant (EPB03) noted that "GoRIM would support [the process of obtaining feedback on the performance of regulatory initiatives] if the model is finalized and if a sufficient amount of data is used to run GoRIM". Several key informants disclosed that the amount of data currently available to their organisations were not sufficient for a valid analysis. However, this situation was not perceived as a final state, since "Knowing how [GoRIM] operates could be very useful as we make changes to our compliance tools going forward", as observed by one participant (RX03).

Model complexity Several key informants mentioned the size and complexity of GoRIM models as a deterrent to their use in a professional environment, e.g.: "I think the presentation [model] overall could be made much more palatable and understandable.", (CWS02), "Model seems complex" (EPB02), and "The model is large so won't be easily accessible/digestible by many Staff" (RX02). One suggestion to address this concern was to select and focus on the KPIs that are most directly related to a regulation's requirements. However, this suggestion leads back to the first challenge identified above, since the identification of these KPIs should rely on the analysis of a large dataset of historical data able to determine, for example, which KPIs have the most predictive power. As a counterpoint to these comments, some key informants felt that models were not complete enough and should capture additional information such as external variables that could impact regulations and the level of enforcement efforts dedicated to compliance with a given regulation. These comments reflect the models developed through interactions with key informants, since GoRIM does not preclude the inclusion of such information. Nevertheless, the amount of information captured by GoRIM models and the complexity of models presented to users may best be understood as a trade-off to be addressed on a case-by-case basis.

Difficulty of correctly weighting contributions to model goals While developing the models, some key informants mentioned that the importance given to varied regulatory components in relation to goal achievement (i.e., weights assigned to contribution links) was incorrect. Indeed, as stated in Step 1, while GoRIM supports the partially automated development of models for regulatory texts, these documents typically do not contain quantified information on the relative importance of sub-components. Determining that importance thus requires domain expertise. While regulators themselves can often provide such knowledge, one key informant noted the need to involve

regulatory program experts in assigning weights to factors to ensure useful results. While this comment implies the need to achieve consensus among stakeholders, different categories of stakeholders may have different views on weights. GRL models can accommodate the resolution of conflicting views using for example AHP [68], as well as different sets of weights to remember conflicting perspectives and analyse them separately [45, 50].

While these findings are exploratory in nature and would benefit from being validated more formally, for example through a large-scale survey, they indicate the relevance of using GoRIM within regulatory organisations. Indeed, the capabilities that key informants identified as stemming from GoRIM—intra-agency communication, visualization at different levels of granularity, and support to data collection planning—can act as enablers of timely and evidence-based decision making about regulations. Moreover, they suggest the feasibility of using GoRIM since the three challenges identified—insufficient data, model complexity, and difficulty of weighing contributions—can all be mitigated. Improved data collection planning should, over time, help to address data needs; existing modelling expertise within regulatory agency could be drawn upon to mitigate models' complexity; and, improved inter-agency communication using models could help achieve consensus about correct weighting of contributions.

Additionally, the DSR methodology we employed in the creation and evaluation of GoRIM afforded us the opportunity to correct and improve GoRIM to its current state. Feedback from key informants enabled the iteration of the steps of GoRIM from its initial format in [58] to its current state described in Sect. 4. We removed the "Select" step where questions to be answered are selected and added the "Data Preparation" step where sample data are tested on the data. Also, the "Model Correction" step, where the models are checked against well-formedness rules and sample strategies from the regulator and then corrected to ensure the analysis behaviour is aligned with the given regulatory context, was also added based on feedback from the key informants.

7 Discussion

The development and evaluation of GoRIM highlighted key benefits to regulators, including the possibility to monitor effectiveness at a much higher frequency than what is currently done in practice nowadays, potentially leading to faster regulatory review and evolution processes, and many other points already discussed in the previous section.

This section discusses further opportunities of applications of GoRIM, while identifying limitations of the method and general threats to the validity of this research.

7.1 Opportunities

As previously observed by Braun et al. [69], many regulations are not written in ways that support simple compliance and performance measurements. For example, indicator definitions and data considerations usually come as an afterthought, once regulatory text or regulatory initiative projects are already produced. GoRIM-related discussions with our key informants made them realise that indicator definitions and availability of data should also be taken into consideration while drafting regulatory text and initiative projects, so alignment between data, indicators, and regulations/initiatives is optimized. GoRIM could hence have an indirect impact on how regulations and initiatives are actually drafted. The increasing availability of large sets of open data should also be considered as an opportunity at that level, especially as they have been shown to be useful in other contexts (e.g., sustainable business ecosystems) where KPIs, goal models, and business intelligence were used [47].

Opportunities surrounding the use of GoRIM also exist for software developers. The results of the previous section already highlighted the enhanced understanding that using GoRIM brings in relation to the data that must be collected to feed indicators in a regulatory intelligence context. They also showed that GoRIM could be used to identify data and indicators that could stop being collected and computed without affecting the quality of the decision-making process. We adapted existing tools to support GoRIM, but there are still other opportunities for better integration of goal models in business intelligence/analytics environment, for the proper visualisation and navigation of large goal models by non-experts (especially for their validation prior to feeding evaluated model data to a database), for the generation of alerts and sharable reports, and maybe for dedicated GoRIM tools.

Although GoRIM has focused mainly on regulatory intelligence, similar problems exist for software and system certification, where compliance to standards (instead of regulations) is the main objective. Although many existing approaches help software and systems engineers demonstrate compliance to standards for certification purpose [70], to our knowledge, there is no approach that attempts to assess the performance of the standards themselves, across multiple certification exercises, and enable evidence-based decision-making for their evolution. GoRIM could likely be adapted to such context. This is even more important for certification authorities in emerging and quickly evolving application areas such as autonomous vehicles [71, 72], where lives are at stake, and also financial technologies (FinTech), including

blockchain-based systems, where critical digital assets are being managed [73].

7.2 Limitations

Despite its contributions, this research suffers from several limitations. The first limitation is related to the creation of the GRL models used in the three case studies. All GRL models were indeed created by one of the authors of this article, in collaboration with co-authors, because the key informants did not have the required modelling skills. This could raise questions on the applicability of GoRIM since other regulators may also lack the required expertise to create and analyse GRL models. This limitation is partly overcome by describing the varied roles that are required to apply GoRIM in Sect. 3.2, which include a "Modeller" with goal modelling skills. The choice of IBM Watson Analytics to support further analysis in the case studies was also a deliberate choice in this regard, since this off-the-shelf tool supports requests in natural language and is easy to use by people familiar with other analytics tools. Note that GoRIM was also recently used with another commercial tool (IBM Cognos Analytics 11 [57]) and that other analytics tools that meet the requirements started at the end of Sect. 4.3 could also be used.

A second limitation is related to the application domains. The case studies in which GoRIM was applied are not reflective of all domains that regulators cover. However, the case studies present three different domains with different regulatory challenges and regulations. They also represent different levels of data variability and quality, and each provides a varied mix of roles related to the monitoring of regulations and initiatives. Nevertheless, applying GoRIM to other domains and jurisdictions would confirm the scope of applicability of the method and its supporting tools.

The third limitation results from a partial evaluation of GoRIM's steps. Although we demonstrated the use of GoRIM through three case studies involving real regulators from different contexts, our evaluation mainly addressed the effectiveness of regulations and regulatory outcomes. Societal outcomes were not modelled separately, and their indicators were not populated with data. However, as seen in Fig. 1, the dimensions of societal outcomes data are very similar to that of compliance data and regulatory initiatives data. Moreover, models for societal outcomes are homomorphic to models for regulatory initiatives, as they are both instances of the same conceptual metamodel. Hence, the approach taken to model and explore relationships between regulatory initiatives data and compliance data should be applicable to the modelling and exploration of relationships between social outcome data and compliance data. Moreover, there is no foreseen conceptual challenge in adding additional dimensions for analysis purposes. Yet, practical challenges might exist. For example, in relation with the creation of the soci-

etal outcome models (as there are fewer formal documents on outcomes than on regulations), the provision of data (which likely come from sources external to the regulator) and to the analysis (as it might be difficult to isolate the contributions of a given regulation on large-scale societal outcomes). Finally, the evaluation of GoRIM presented in this paper relies on a small number of key informants and their subjective perspectives, rather than on objective measures such a time and costs comparisons with the methods they currently used, to assess if the use of GoRIM would increase the efficiency of regulatory monitoring.

8 Related work

Previous research in goal-oriented modelling has provided a number of approaches and tools that served as a basis for GoRIM. As the domain of regulation shifted its focus in the past decade to measuring the achievement of regulatory goals rather than on prescriptive ways of achieving these goals, GRL has been used to develop methods and tools for performance modelling of regulations [18, 51]. These methods allow transforming regulatory text into goal models and defining KPIs for the regulation's sub-parts in order to evaluate goal achievement, hence compliance to the regulation. Such goal models can then be used as analytical dimensions in business intelligence tools to evaluate the performance of the regulation based on collected compliance data. GoRIM extends these concepts to models of regulatory initiatives and societal outcomes.

Other approaches have been derived from GRLaw and Legal-GRL, which extend GRL to model legal requirements that can be linked to organisational goal models in order to assess the impact of regulations on organisational goals [74–76]. While this research provides the grounding for GoRIM's model creation step and its use of regulatory data for business intelligence, it is not organized as a comprehensive method that can be used by practitioners for the purpose of regulatory intelligence, especially from a regulator's perspective.

Similarly, goal modelling was used in various software and system development activities (e.g., to better comply with legal security requirements [77]) but also certification activities [70]. However, this was mainly done in contexts where systems must comply with different laws or standards, from the developer's perspective, without measuring whether the laws and standards themselves were effective and whether they should be modified based on data from multiple certifications.

Another related body of work focuses on business process compliance, hence supporting the analysis of the compliance of organisations and their processes to regulations [78, 79]. Research on business process compliance has been pursued

in several fields, including Requirements Engineering, Natural Language Processing, Formal Methods, and Artificial Intelligence (e.g., by Jiang et al. [80]). Of particular interest is the Legal-URN framework, which extends URN to allow the comparison of different interpretations of a regulation on the compliance level of business processes [81]. The Legal-URN framework has also been combined with Eunomos, a legal knowledge and document management system, to establish and manage compliance and hence assist decision-making and reporting [82]. Another modelling language for regulatory compliance, Nòmos 3, focuses on evaluating if a software system complies with relevant laws by modelling the law itself, the requirements of the system, and the roles that are responsible for complying with the law [83]. This body of work is thus focused on organisational or system compliance to regulations, rather than on measuring regulations' impacts (or lack thereof) and whether regulations should evolve.

Goal-oriented modelling has also been used to drive business intelligence (BI) activities in general [56], from the design of data warehouses [84], BI architectures [85], and BI systems [86, 87], to the use of goal models in the selection of appropriate dashboard visualisations [88]. These combined applications of goal modelling with BI are orthogonal to the regulation-based and performance-oriented method presented in this paper but could also be used to address complementary implementation concerns.

Finally, many non-goal-oriented approaches were used to reason about legal requirements [24], but most again focus on the needs of regulated parties rather than on those of regulators, and few are actually driven by data. One notable exception is the use of UML and OCL by Soltana et al. [89] for simulating the potential impact of modifications to legal policies. Their work is however not targeting the continuous monitoring of the performance of existing regulations, which is essential in a regulatory intelligence context.

9 Conclusion and future work

The practical problem addressed by this research is the lack of support provided to regulators for ongoing evaluation of regulations' performance and the evolution of regulations based on evaluation results, hence for reviewing regulations. For regulators, demonstrating the performance of the regulations they administer is also desirable to build trust in the regulatory system and to demonstrate that they operate with transparency and responsiveness to all stakeholders in the regulatory ecosystem. Current methods for evaluating regulations either have a scope that is too narrow for this purpose (e.g., focus on costs) [4, 13], or use impact analysis methods that imply significant delays and costs [1].

GoRIM provides a solution to regulators' needs by providing modelling guidance and tools for leveraging existing and future datasets in a manner that may support timely, evidence-based decision-making. While GoRIM builds on existing research on the use of goal modelling for evaluating regulated parties' compliance and existing business intelligence practices and tools, it provides novel capabilities for understanding and analysing a regulatory ecosystem by supporting a transparent capture and analysis of regulations, multi-dimensional data, and performance indicators. GoRIM is thus aligned with current regulatory practices of identifying and linking performance indicators and regulatory activities through logic models [90], but provides much-needed, actionable support for ongoing evaluation and evolution.

Indeed, logic models graphically depict the shared relationships among the resources, activities, outputs, outcomes, and impact for the program, but do not support the quantitative analysis and understanding of these relationships in terms of success of a regulatory program and, indirectly, of the performance of the regulation. In contrast, the results of data analysed by following GoRIM steps facilitate conversations on the relevance of indicators that measure each dimension and, hence, inform the regulator where data collection could be minimized. They also provide opportunities for traceability of regulatory effectiveness as the models can be monitored and analysed continually to show how well (or not) a regulation is performing. Moreover, the goal models created through GoRIM can help regulators identify which types of data to collect in order to support a comprehensive analysis of a regulation's effectiveness.

The model-driven nature of GoRIM is hence core to its contribution. GoRIM relies on goal-oriented modelling and data analytics, provides a structure that enables early and continuous intervention in the regulatory ecosystem and supports decision-making related to the evolution of regulations and associated regulatory initiatives. As such, GoRIM provides a systematic model-driven approach, with related tools, for regulatory intelligence. Of particular interest are (1) the use of a requirements modelling standard (GRL) to formalize and visualise initiatives, regulations, and societal models, (2) the use of these structures as dimensions for analysis and reporting, and (3) the use of off-the-shelf analytics tools to detect correlations between these complementary views.

Our three case studies have highlighted a good applicability of GoRIM to different regulatory contexts, and our evaluation further indicates good perceived usefulness. Feedback from regulators showed that, despite some perceived challenges in using GoRIM in professional contexts, such as the lack of sufficient data and the complexity of the goal models, the method provides capabilities that can be important to professional regulator practices, such as improved understanding of regulations and relationships among regulatory activities and objectives, improved communication among

varied regulatory stakeholders, and support to data collection planning.

In the future, we intend to apply GoRIM to more regulatory contexts as a way of further evaluating and improving GoRIM as a method to show the effectiveness of regulations. Selected case studies would ideally provide access to data on all three dimensions—regulatory initiatives, regulations, and societal outcomes—and be longitudinal in nature to evaluate GoRIM's ability to support an ongoing monitoring and evidence-based decision-making. A special attention needs to be paid to the second loop between regulations and societal outcome models in Fig. 1. We also plan to address some of the limitations identified in the previous section, and explore some of the opportunities identified, including the adaptation of GoRIM to the software/systems engineering standardization and certification contexts. Taking these different contexts explicitly into account would also help identify and better formalize method fragments in GoRIM that are only relevant in or need to be tailored to particular situations (e.g., what consensus-building approach to select for quantifying contribution weights in goal models). In particular, situational method engineering [91] could help refine GoRIM more systematically here. There are also opportunities to better relate our metamodel (Fig. 3) to existing legal ontologies and benefit from their conceptual foundations, and also to improve the visualization of GRL models for regulatory intelligence purposes, along the lines proposed by Griffo et al. for legal models [92]. In addition, although our work in showing how data analytics supports the decision-making process is an addition to how current methods provide frameworks for better leveraging data analytics tools, limitations in terms of data availability is still an issue. We intend to validate our approach to the absence of data discussed in Sect. 5.3 as well as explore further approaches in this regard.

Availability of data and material The interview questionnaires used to validate GoRIM, together with the protocol and the thematic analysis of answers, are available at http://bit.ly/GoRIM-supp. Other data collected as part of this research cannot be made available due to constraints imposed by our institution's Research Ethics Board.

Code availability jUCMNav's code, including OCL well-formedness rules, are available online at https://github.com/JUCMNAV/.

Declarations

Conflict of interest The authors declare no conflicts of interests.

Appendix A: Selected well-formedness rules for GRL models

In step 3 of our method (Sect. 4.3), the modeller checks the GRL models against well-formedness rules, listed in Table 4. These 19 rules are user-selectable OCL constraints supported by jUCMNav [45]. Violations to these rules are reported automatically by jUCMNav, which then highlights the violating model elements. Satisfying these rules helps ensure that the input models satisfy specific static properties that go beyond standard URN.

Table 4 Mapping between regulatory initiatives (or societal outcomes) and GRL model elements

#	Rule Name	Description
1	GRLactorEmpty	GRL actor is empty
2	GRLactorNoCLASSTYPE	ST_CLASSTYPE stereotypes cannot be used on actors
3	GRLactorNoCycle	GRL actor must not be part of a containment cycle
4	GRLactorWithoutRef	GRL actor definition without reference
5	GRLconditionDependeeOfIE	A GRL resource with a ST_CONDITIONTYPE metadata must be a dependee of an intentional element
6	GRLconditionNotADepender	A GRL resource with a ST_CONDITIONTYPE metadata cannot depend on anything else
7	GRLcontributionRange	GRL quantitative importance must be between 0 and 100
8	GRLdependencyEvaluation-Consistency	The depender of a GRL dependency should not have an evaluation value higher than its dependee's
9	GRLgraphEmpty	GRL graph is empty
10	GRLimportanceRange	GRL quantitative importance must be between 0 and 100
11	GRLincomingContributions-NotMore Than100	Contributions to an intentional element must not sum up to a value higher than 100
12	GRLindicatorThreshold-Consistency	Indicator threshold value must be between the target and worst values
13	GRLintentionalElementIn-ManyActors	GRL intentional element must not be bound to more than one actor
14	GRLintentionalElemWithoutRef	GRL intentional element definition without any reference
15	GRLintentionalLinkWithoutRef	GRL intentional link definition without any reference
16	GRLintentionElemInsideBut-Unbound	GRL intentional element is visually contained by actor but is not bound to it
17	GRLnoOverlappingActors	GRL actor boundary must not overlap with the boundary of another actor
18	GRLnoOverlappingIEonActor	GRL intentional element must not overlap with the boundary of an actor
19	GRLstrategyWithIncorrect-XORinit	GRL strategy initializes more than one child of an XOR decomposition

References

1. Coglianese, C.: Measuring Regulatory Performance: Evaluating the Impact of Regulation and Regulatory Policy. OECD Publishing, Paris (2012). https://www.oecd.org/gov/regulatory-policy/1_coglianese%20web.pdf

2. Plantin, G.: When Insurers Go Bust: An Economic Analysis of the Role and Design of Prudential Regulation. Princeton University Press, USA (2016)

3. OECD: OECD best practice principles for regulatory policy. OCSD iLibrary (2014). https://doi.org/10.1787/23116013

4. Ellig, J., Broughel, J.: Regulation: What's the Problem? George Mason University, USA (2011). https://www.mercatus.org/publication/regulation-whats-problem

5. Crain, W.M., Crain, N.V.: The Cost of Federal Regulation to the US Economy, Manufacturing and Small Business. National Association of Manufacturers (2014). https://bit.ly/30GB3iB

6. Head, B.W.: Three lenses of evidence-based policy. Aust. J. Public Adm. **67**(1), 1–11 (2008). https://doi.org/10.1111/j.1467-8500.2007.00564.x

7. Parker, D., Kirkpatrick, C.: Measuring Regulatory Performance: The Economic Impact of Regulatory Policy: A Literature Review of Quantitative Evidence. OECD Publishing, Paris (2012). https://www.oecd.org/gov/regulatory-policy/3_Kirkpatrick%20Parker%20web.pdf

8. Akhigbe, O., Amyot, D., Richards, G.: Information Technology Artifacts in the Regulatory Compliance of Business Processes: A Meta-Analysis, pp. 89–104. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-17957-5_6

9. Gasmi, F., Noumba, P., Recuero Virto, L.: The role of institutional design in the conduct of infrastructure industry reforms—an illustration through telecommunications in developing countries. S. Afr. J. Inf. Commun. **200**(9), 4–16 (2008). https://doi.org/10.23962/10539/19776

10. Gasmi, F., Noumba, P., Recuero Virto, L.: Does political accountability matter for infrastructure regulation? The case of telecommunications. In: Infrastructure Regulation: What Works, Why and How Do We Know? pp. 61–82. World Scientific (2011). https://doi.org/10.1142/9789814335744_0003

11. Berg, S.V.: Developments in best-practice regulation: principles, processes, and performance. Electr. J. **13**(6), 11–18 (2000). https://doi.org/10.1016/S1040-6190(00)00120-2

12. Berg, S.V.: Seven elements affecting governance and performance in the water sector. Util. Policy **43**, 4–13 (2016). https://doi.org/10.1016/j.jup.2016.04.013

13. Hahn, R., Hird, J.: The costs and benefits of regulation: review and synthesis. Yale J. Regul. **8**(1), 233 (1991)

14. Savaya, R., Waysman, M.: The logic model: a tool for incorporating theory in development and evaluation of programs. Adm. Soc. Work **29**(2), 85–103 (2005). https://doi.org/10.1300/J147v29n02_06

15. Knowlton, L.W., Phillips, C.C.: The Logic Model Guidebook: Better Strategies for Great Results. SAGE, London (2012)
16. Felgate, T.: What is Regulatory Intelligence? (definitions) (2013). https://regulatory-intelligence.blogspot.com/2013/02/what-is-regulatory-intelligence.html
17. Hynes, C.: Regulatory Intelligence: Implications for product development. TOPRA (2014). http://bit.ly/2pr5UiY
18. Badreddin, O., Mussbacher, G., Amyot, D., Behnam, S.A., Rashidi-Tabrizi, R., Braun, E., Richards, G.: Regulation-based dimensional modeling for regulatory intelligence. RELAW **2013**, 1–10 (2013). https://doi.org/10.1109/RELAW.2013.6671340
19. Amyot, D., Mussbacher, G.: User requirements notation: the first ten years, the next ten years (invited paper). J. Softw. **6**(5), 747–768 (2011)
20. International Telecommunication Union.: Rec. Z.151 (10/18): User Requirements Notation (URN)-Language definition (2018). https://www.itu.int/rec/T-REC-Z.151/en
21. Ghanavati, S., Amyot, D., Peyton, L.: A requirements management framework for privacy compliance. In: WER'07, pp. 149–159 (2007). http://www.inf.puc-rio.br/~wer/WERpapers/artigos/artigos_WER07/Qwer07-ghanavati.pdf
22. Tawhid, R., Braun, E., Cartwright, N., Alhaj, M., Mussbacher, G., Shamsaei, A., Richards, G.: Towards outcome-based regulatory compliance in aviation security. In: 20th IEEE International Requirements Engineering Conference (RE), pp. 267–272 (2012). https://doi.org/10.1109/RE.2012.6345813
23. Pourshahid, A., Amyot, D., Peyton, L., Ghanavati, S., Chen, P., Weiss, M., Forster, A.J.: Business process management with the user requirements notation. Electron. Commer. Res. **9**(4), 269–316 (2009). https://doi.org/10.1007/s10660-009-9039-z
24. Akhigbe, O., Amyot, D., Richards, G.: A systematic literature mapping of goal and non-goal modelling methods for legal and regulatory compliance. Requir. Eng. **24**(4), 459–481 (2019). https://doi.org/10.1007/s00766-018-0294-1
25. Aparicio, M., Costa, C.J.: Data visualization. Commun. Des. Q. **3**(1), 7–11 (2015). https://doi.org/10.1145/2721882.2721883
26. Radaelli, C., Fritsch, O.: Measuring regulatory performance: evaluating regulatory management tools and programmes. OECD Expert Paper No. 2 (2012). https://www.oecd.org/gov/regulatory-policy/2_Radaelli%20web.pdf
27. Akhigbe, O., Amyot, D., Mylopoulos, J., Richards, G.: What can information systems do for regulators? A review of the state-of-practice in Canada. In: IEEE Eleventh International Conference on Research Challenges in Information Science (RCIS), pp. 57–65 (2017). https://doi.org/10.1109/RCIS.2017.7956518
28. Nielsen, V., Parker, C.: Is it Possible to Measure Compliance? SSRN 935988 (2006). https://papers.ssrn.com/abstract=935988
29. Robertson, A.S., Reisin Miller, A., Dolz, F.: Supporting a data-driven approach to regulatory intelligence. Nat. Rev. Drug Discov. (2020). https://doi.org/10.1038/d41573-020-00101-4
30. Hevner, A., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Quart. **28**(1), 75–105 (2004)
31. Wieringa, R.J.: Design Science Methodology for Information Systems and Software Engineering. Springer, Berlin (2014). https://doi.org/10.1007/978-3-662-43839-8_1
32. Engström, E., Storey, M.A., Runeson, P., Höst, M., Baldassarre, M.T.: How software engineering research aligns with design science: a review. Empir. Softw. Eng. **25**(4), 2630–2660 (2020). https://doi.org/10.1007/s10664-020-09818-7
33. Bider, I., Perjons, E.: Design science in action: developing a modeling technique for eliciting requirements on business process management (BPM) tools. Softw. Syst. Model. **14**, 1159–1188 (2015). https://doi.org/10.1007/s10270-014-0412-6
34. Peffers, K., Tuunanen, T., Rothenberger, M., Chatterjee, S.: A design science research methodology for information systems research. J. Manag. Inf. Syst. **24**(3), 45–77 (2007). https://doi.org/10.2753/MIS0742-1222240302
35. Gregor, S., Hevner, A.: Positioning and presenting design science research for maximum impact. MIS Quart. **37**(2), 337–356 (2013). https://doi.org/10.25300/MISQ/2013/37.2.01
36. Braatz, B., Brandt, C.: A framework for families of domain-specific modelling languages. Softw. Syst. Model. **13**, 109–132 (2014). https://doi.org/10.1007/s10270-012-0271-y
37. van Brocke, J., Winter, R., Hevner, A., Maedche, A.: Special issue editorial—accumulation and evolution of design knowledge in design science research: a journey through time and space. J. Assoc. Inf. Syst. **21**(3), 9 (2020). https://doi.org/10.17705/1jais.00611
38. Venable, J., Pries-Heje, J., Baskerville, R.: FEDS: a framework for evaluation in design science research. Eur. J. Inf. Syst. **25**(1), 77–89 (2016). https://doi.org/10.1057/ejis.2014.36
39. Payne, G., Payne, J.: Key Concepts in Social Research. SAGE, London (2004)
40. Mills, A., Durepos, G., Wiebe, E.: Encyclopedia of Case Study Research, Thousand Oaks, California (2010). https://doi.org/10.4135/9781412957397
41. Yin, R.K.: Case Study Research: Design and Methods, 5th edn. SAGE, New York (2013)
42. Feldt, R., Magazinius, A.: Validity threats in empirical software engineering research—an initial survey. In: Proceedings of the 22nd International Conference on Software Engineering and Knowledge Engineering, pp. 374–379 (2010).
43. Perry, D.E., Porter, A.A., Votta, L.G.: Empirical studies of software engineering: a roadmap. In: Proceedings of the Conference on the Future of Software Engineering, pp. 345–355 (2000). https://doi.org/10.1145/336512.336586
44. Wieringa, R., Daneva, M.: Six strategies for generalizing software engineering theories. Sci. Comput. Program. **101**, 136–152 (2015)
45. Amyot, D., Shamsaei, A., Kealey, J., Tremblay, E., Miga, A., Mussbacher, G., Alhaj, M., Tawhid, R., Braun, E., Cartwright, N.: Towards advanced goal model analysis with jUCMNav. In: Advances in Conceptual Modeling (ER 2012), LNCS, vol. 7518, pp. 201–210 (2012). https://doi.org/10.1007/978-3-642-33999-8_25
46. OMG.: Software & Systems Process Engineering Metamodel (SPEM), Version 2.0, formal/2008-04-01 (2008). https://www.omg.org/spec/SPEM/
47. Kampars, J., Zdravkovic, J., Stirna, J., Grabis, J.: Extending organizational capabilities with Open Data to support sustainable and dynamic business ecosystems. Softw. Syst. Model. **19**, 371–398 (2020). https://doi.org/10.1007/s10270-019-00756-7
48. Shamsaei, A.: Indicator-based Policy Compliance of Business Processes. Doctoral thesis, University of Ottawa, Canada (2012). https://doi.org/10.20381/ruor-6171
49. Akhigbe, O.: A Goal-Oriented Method for Regulatory Intelligence. Doctoral thesis, University of Ottawa, Canada (2018). https://doi.org/10.20381/ruor-22507
50. Akhigbe, O., Alhaj, M., Amyot, D., Badreddin, O., Braun, E., Cartwright, N., Richards, G., Mussbacher, G.: Creating quantitative goal models: governmental experience. In: 33rd International Conference on Conceptual Modeling (ER 2014), LNCS, vol. 8824, pp. 466–473 (2014). https://doi.org/10.1007/978-3-319-12206-9_40
51. Rashidi-Tabrizi, R., Mussbacher, G., Amyot, D.: Transforming regulations into performance models in the context of reasoning for outcome-based compliance. RELAW **2013**, 34–43 (2013). https://doi.org/10.1109/RELAW.2013.6671344
52. Liaskos, S., Jalman, R., Aranda, J.: On eliciting contribution measures in goal models. In: 20th IEEE International Requirements Engineering Conference (RE), pp. 221–230 (2012). https://doi.org/10.1109/RE.2012.6345808

53. Trinkenreich, B., Santos, G., Perini Barcellos, M.: SINIS: a GQM+strategies-based approach for identifying goals, strategies and indicators for IT services. Inf. Softw. Technol. **100**, 147–164 (2018). https://doi.org/10.1016/j.infsof.2018.04.006

54. Hassine, J., Amyot, D.: An empirical approach toward the resolution of conflicts in goal-oriented models. Softw. Syst. Model. **16**(1), 279–306 (2017). https://doi.org/10.1007/s10270-015-0460-6

55. Fan, Y., Anda, A.A., Amyot, D.: An arithmetic semantics for GRL goal models with function generation. In: SAM 2018, LNCS, vol. 11150, pp. 144–162 (2018). https://doi.org/10.1007/978-3-030-01042-3_9

56. Horkoff, J., Barone, D., Jiang, L., Yu, E., Amyot, D., Borgida, A., Mylopoulos, J.: Strategic business modeling: representation and reasoning. Softw. Syst. Model. **13**(3), 1015–1041 (2014). https://doi.org/10.1007/s10270-012-0290-8

57. IBM.: IBM Cognos Analytics (2016). https://www.ibm.com/ca-en/products/cognos-analytics

58. Akhigbe, O., Heap, S., Islam, S., Amyot, D., Mylopoulos, J.: Goal-oriented regulatory intelligence: how can watson analytics help? In: 36th International Conference on Conceptual Modeling (ER 2017), LNCS, vol. 10650, pp. 77–91 (2017). https://doi.org/10.1007/978-3-319-69904-2_7

59. Government of Canada.: Migratory Birds Convention Act, 1994 (S.C. 1994, c. 22) (1994). http://laws-lois.justice.gc.ca/eng/acts/M-7.01/

60. Environment and Climate Change Canada.: Birds Protected Under the Migratory Birds Convention Act (2011). https://www.canada.ca/en/environment-climate-change/services/migratory-birds-legal-protection/convention-act.html

61. Government of Canada.: Migratory Birds Regulations (C.R.C., c. 1035) (2016). http://laws-lois.justice.gc.ca/eng/regulations/C.R.C.,_c._1035/index.html

62. Government of Canada.: Migratory Birds Regulations (C.R.C., c. 1035) - Game Birds (2016). http://laws-lois.justice.gc.ca/eng/regulations/C.R.C.,_c._1035/page-3.html

63. Amyot, D., Horkoff, J., Gross, D., Mussbacher, G.: A lightweight GRL profile for *i** modeling. In: Advances in Conceptual Modeling-Challenging Perspectives (ER 2009), LNCS, vol. 5833, pp. 254–264 (2009). https://doi.org/10.1007/978-3-642-04947-7_31

64. Amyot, D., Rashidi-Tabrizi, R., Mussbacher, G., Kealey, J., Tremblay, E., Horkoff, J.: Improved GRL Modeling and Analysis with jUCMNav 5. In: 6th International i* Workshop (iStar 2013), CEUR-WS, vol. 978, pp. 137–139 (2013). http://ceur-ws.org/Vol-978/paper_26.pdf

65. Government of Canada.: Metal and Diamond Mining Effluent Regulations (SOR/2002-222) (2002). http://laws-lois.justice.gc.ca/eng/regulations/SOR-2002-222/PITIndex.html

66. Government of Canada.: Fisheries Act (R.S.C., 1985, c. F-14) (2007). http://laws-lois.justice.gc.ca/eng/acts/F-14/

67. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. **3**(2), 77–101 (2006). https://doi.org/10.1191/1478088706qp063oa

68. Saaty, T.L.: A scaling method for priorities in hierarchical structures. J. Math. Psychol. **15**(3), 234–281 (1977). https://doi.org/10.1016/0022-2496(77)90033-5

69. Braun, E., Cartwright, N., Shamsaei, A., Behnam, S.A., Richards, G., Mussbacher, G., Alhaj, M., Tawhid, R.: Drafting and modeling of regulations: is it being done backwards? In: RELAW 2012, pp. 1–6. https://doi.org/10.1109/RELAW.2012.6347802

70. Nair, C., de la Vara, J.L., Sabetzadeh, M., Briand, L.: An extended systematic literature review on provision of evidence for safety certification. Inf. Softw. Technol. **56**(7), 689–717 (2014). https://doi.org/10.1016/j.infsof.2014.03.001

71. Johnsen, S.O., Hoem, Å., Stålhane, T., Jenssen, G., Moen, T.: Risk-based regulation and certification of autonomous transport systems. In: Safety and Reliability—Safe Societies in a Changing World (ESREL 2018), pp. 1791–1799 (2018). https://doi.org/10.1201/9781351174664

72. de la Vara, J.L., Ruiz, A., Attwood, K., Espinoza, H., Panesar-Walawege, R.K., López, Á., del Río, I., Kelly, T.: Model-based specification of safety compliance needs for critical systems: a holistic generic metamodel. Inf. Softw. Technol. **72**, 16–30 (2016). https://doi.org/10.1016/j.infsof.2015.11.008

73. Sharifi, S., McLaughlin, P., Amyot, D., Mylopoulos, J.: Goal modeling for fintech certification. In: iStar 2020, CEUR-WS 2641, pp. 73–78 (2020). http://ceur-ws.org/Vol-2641/paper_13.pdf

74. Ghanavati, S., Amyot, D., Peyton, L., Siena, A., Perini, A., Susi, A.: Integrating business strategies with requirement models of legal compliance. Int. J. Electron. Bus. **8**(3), 260–280 (2010). https://doi.org/10.1504/IJEB.2010.034171

75. Ghanavati, S., Humphreys, L., Boella, G., Di Caro, L., Robaldo, L., van der Torre, L.W.N.: Compliance with multiple regulations. In: ER 2014: Conceptual Modeling, pp. 415–422 (2014). https://doi.org/10.1007/978-3-319-12206-9_35

76. Rabinia, A., Ghanavati, S., Humphreys, L., Hahmann, T.: A methodology for implementing the formal legal-grl framework: a research preview. In: REFSQ 2020, LNCS, vol. 12045, pp. 124–131 (2020). https://doi.org/10.1007/978-3-030-44429-7_9

77. Islam, S., Mouratidis, H., Jürjens, J.: A framework to support alignment of secure software engineering with legal regulations. Softw. Syst. Model. **10**, 369–394 (2011). https://doi.org/10.1007/s10270-010-0154-z

78. Elgammal, A., Turetken, O., van den Heuvel, W., Papazoglou, M.: Formalizing and applying compliance patterns for business process compliance. Softw. Syst. Model. **15**, 119–146 (2016). https://doi.org/10.1007/s10270-014-0395-3

79. El Kharbili, M.: Business process regulatory compliance management solution frameworks: a comparative evaluation. In: APCCM'12, pp. 23–32 (2012). https://doi.org/10.5555/2523782.2523786

80. Jiang, J., Aldewereld, H., Dignum, V., Wang, S., Baida, X.: Regulatory compliance of business processes. AI Soc. **30**(3), 393–402 (2015). https://doi.org/10.1007/s00146-014-0536-9

81. Ghanavati, S., Hulstijn, J.: Impact of legal interpretation in business process compliance. TELERISE **2015**, 26–31 (2015). https://doi.org/10.1109/TELERISE.2015.13

82. Boella, G., Tosatto, S.C., Ghanavati, S., Hulstijn, J., Humphreys, L., Muthuri, R., Rifaut, A., van der Torre, L.W.N.: Integrating legal-URN and eunomos: towards a comprehensive compliance management solution. AICOL **2013**, 130–144 (2013). https://doi.org/10.1007/978-3-662-45960-7_10

83. Ingolfo, S., Jureta, I., Siena, A., Perini, A., Susi, A.: Nòmos 3: Legal compliance of roles and requirements. In: ER 2014: Conceptual Modeling, pp. 275–288 (2014). https://doi.org/10.1007/978-3-319-12206-9_22

84. Giorgini, P., Rizzi, S., Garzetti, M.: GRAnD: a goal-oriented approach to requirement analysis in data warehouses. Decis. Support Syst. **45**(1), 4–21 (2008). https://doi.org/10.1016/j.dss.2006.12.001

85. Fekete, D, Vossen, G.: The GOBIA method: towards goal-oriented business intelligence architectures. In: FGDB 2015, CEUR-WS 1458, pp. 409–418 (2015). http://ceur-ws.org/Vol-1458/H03_CRC44_Fekete.pdf

86. Barone, D., Topaloglou, T., Mylopoulos, J.: Business intelligence modeling in action: a hospital case study. In: CAISE 2012, LNCS, vol. 7328, pp. 502–517 (2012). https://doi.org/10.1007/978-3-642-31095-9_33

87. Burnay, C., Jureta, I.J., Linden, I., Faulkner, S.: A framework for the operationalization of monitoring in business intelligence requirements engineering. Softw. Syst. Model. **15**, 531–552 (2016). https://doi.org/10.1007/s10270-014-0417-1

88. Lavalle, A., Maté, A., Trujillo, J., Rizzi, S.: Visualization requirements for business intelligence analytics: a goal-based, iterative framework. In: IEEE 27th International Requirements Engineering Conference (RE), pp. 109–119 (2019). https://doi.org/10.1109/RE.2019.00022

89. Soltana, G., Sannier, N., Sabetzadeh, M., Briand, L.: Model-based simulation of legal policies: framework, tool support, and validation. Softw. Syst. Model. **17**, 851–883 (2018). https://doi.org/10.1007/s10270-016-0542-0

90. Mazur, E.: Outcome Performance Measures of Environmental Compliance Assurance: Current Practices, Constraints and Ways Forward. OECD Environment Working Papers, No. 18. OECD Publishing, Paris (2010). https://doi.org/10.1787/5kmd9j75cf44-en

91. Henderson-Sellers, B., Ralyté, J., Ågerfalk, P., Rossi, M.: Situational Method Engineering. Springer (2014). https://doi.org/10.1007/978-3-642-41467-1

92. Griffo, C., da Silva Teixeira, M.G., Almeida, J.P.A., Gailly, F., Guizzardi, G.: LawV: Towards an ontology-based visual modeling language in the legal domain. In: ONTOBRAS 2020, CEUR-WS 2728, pp. 75–88 (2020). http://ceur-ws.org/Vol-2728/paper6.pdf

Okhaide Akhigbe is a Part-time Professor at the University of Ottawa and Business Consultant with the Government of Canada. His research focuses on the use of digital technologies to achieve data-driven evidence in the regulatory domain.

Daniel Amyot is Professor at the School of Electrical Engineering and Computer Science of the University of Ottawa. His research interests include requirements engineering, software engineering, goal and process modeling, regulatory compliance, smart contracts, and healthcare informatics. Daniel led the standardization of the User Requirements Notation at the International Telecommunication Union from 2002 to 2013. He was general chair of the Requirements Engineering conference in 2015 and program co-chair in 2018. Daniel is on the editorial boards of SoSyM and the Requirements Engineering Journal. More information can be found here: https://www.site.uottawa.ca/~damyot/.

Gregory Richards is an Adjunct Professor at the University of Ottawa's Telfer School of Management. His research focuses on the use of advanced analytics and knowledge utilization in organizations. More information can be found here: https://telfer.uottawa.ca/en/directory/gregory-richards/.

Lysanne Lessard is Associate Professor at the Telfer School of Management, University of Ottawa. Her research centers on the study, modeling, architecture and design of digitally-enabled services. More information on her work can be found at: https://telfer.uottawa.ca/en/directory/lysanne-lessard/.