



Eliciting Ethicality Requirements Using the Ontology-Based Requirements Engineering Method

Renata Guizzardi¹(✉), Glenda Amaral², Giancarlo Guizzardi^{1,2},
and John Mylopoulos³

¹ University of Twente, Enschede, The Netherlands
`r.guizzardi@utwente.nl`

² CORE/KRDB, Free University of Bozen-Bolzano, Bolzano, Italy
`{gmouraamaral,giancarlo.guizzardi}@unibz.it`

³ University of Toronto, Toronto, Canada
`jm@cs.toronto.edu`

Abstract. The advent of socio-technical, cyber-physical and Artificial Intelligence (AI) systems has broadened the scope of requirements engineering which must now deal with new classes of requirements, concerning ethics, privacy and trust. Unfortunately, requirements engineers cannot be expected to understand the qualities behind these new classes of systems so that they can conduct elicitation, analysis and operationalization. To address this issue, we propose a methodology for conducting requirements engineering which starts with the adoption of an ontology for a *quality domain*, such as ethicality, privacy or trustworthiness, populates the ontology for the system-to-be and conducts requirements analysis grounded on the populated ontology. We illustrate our proposal with ethicality requirements.

Keywords: Requirements elicitation and analysis · Foundational ontologies · Ethical requirements

1 Introduction

In a world where Artificial Intelligence (AI) is pervasive, controlling more services and systems everyday, humans may feel threatened or at risk by giving up control to machines. In this context, many of the potential issues are related to safety and ethics. For example, AI systems may be biased towards a group of people in detriment of others, they may lead to job loss and wealth inequality, and they may make mistakes and even go rogue, by acting against the interests of humanity [4].

Providing a global solution to these problems is a challenging endeavor, but one that has recently been recognized by different organizations, which have proposed guidelines and standards aimed at addressing this pressing matter. Among these, we may cite the IEEE Standard Model Process for Addressing Ethical Concerns during System Design [5] and the European Union Ethics Guidelines

for Trustworthy AI [3]. What these works have in common is proposing a set of principles to which the system-to-be must adhere to be considered ethical. The topic targeted in this paper is how to make sure that such principles may effectively guide system development.

We claim that Requirements Engineering (RE) is the Software Engineering area that may exert a bigger impact on developing ethical systems, emphasizing ethical principles in system development from the start. RE is not only responsible for producing the set of requirements that will conduct the design of the system-to-be, but also for validating if such requirements have been properly met, and for monitoring if they are still valid throughout the whole life cycle of the system, even after it becomes operational. However, proposing concepts, tools and techniques that support the incorporation of high-level societal concerns and goals (such as ethicality) into the software development processes as explicit requirements is still a challenge in the RE field.

The solution to the targeted issue involves a deep understanding of what the proposed ethical principles mean and how they can be converted in concrete system requirements, which can then guide system design, besides being validated and monitored. For that, we rely on an ontological approach, based on a novel Requirements Engineering method we name Ontology-based Requirements Engineering (ObRE). The ObRE methodology consists of three activities: 1) adopt or develop an ontology to conceptually clarify the meaning of a class of requirements (in this paper, ethicality requirements); 2) instantiate the ontology for a system-to-be, resulting in a domain model; and 3) use the domain model to guide analysis, resulting in requirements models, such as goal models, requirements tables, user stories etc. Besides presenting ObRE, this paper illustrates its use with an example drawn from the driverless car domain.

Ontological analysis provides a foundation for ObRE as it enables a deep account of the meaning of a particular domain. In turn, such analysis is based on a **foundational ontology** to offer a domain-agnostic set of concepts drawing ideas from Philosophy and Cognitive Science. Domain-agnostic identity and dependency, taxonomic relationships and mereology are examples of concepts offered by foundational ontologies. In turn, ontological analysis uses a foundational ontology to develop **domain ontologies**, i.e., a set of concepts and relationships for a specific domain to be shared by a community of users [19]. It is important to highlight that in our work, an ontology is a reference conceptual framework for conceptualizing a domain, rather than a mere logical specification to support automated reasoning. ObRE is based on the Unified Foundational Ontology (UFO) [20], extended with concepts from the Non-functional Requirements Ontology (NFRO) [23]. The use of NFRO is justified by the fact that the classes of requirements ObRE targets belong to quality domains (e.g. ethicality, trustworthiness and privacy) and, as such, fall into the category of non-functional requirements, as explained in detail in the next section of this paper. ObRE is intended to help a requirements analyst cope with non-functional requirements where the analysts literally doesn't know where to begin in conducting elicitation and analysis, which is the case of ethical requirements, the very focus of this paper. ObRE is intended to help by "semantically unpacking" requirements concepts thereby enabling requirements activities.

The remainder of this paper is structured as follows: Sect. 2 discusses what we mean by ontological analysis, and explains the ontological account of requirements adopted in this work; Sect. 3 presents the ObRE method; Sect. 4 illustrates the proposed method with an example from the ethical AI systems domain; Sect. 5 discusses related works; and finally, Sect. 6 presents final remarks.

2 Research Baseline

2.1 Ontological Analysis

The notions of *ontology* and *ontological analysis* adopted here are akin to their interpretations in philosophy [10]. In this view, the goals of *ontological analysis* are: (i) characterize what kinds of entities are assumed to exist by a given conceptualization of a set of phenomena in reality; (ii) the metaphysical nature of these kinds of entities. An *ontology*, in turn, is a system of categories and their ties (here represented as an artifact) that makes justice to what is uncovered by (i) and (ii).

In this sense, an ontology is neither merely a logical specification nor is it mainly concerned with making terminological and taxonomic distinctions. For example, in addressing the domains of risk, one is less concerned with what specific subtypes of risk exist (e.g., physical, biological, financial, electronic), but instead with what exactly *is* risk? (What kind of entity is it? What is its nature?). Is it an object? an event? a relationship? a complex property? If the latter, is a categorical or dispositional property? what is the bearer of such a property?, and so on.

Given the nature of this method of analysis, it must be supported by a domain-independent system comprising the most general categories, hence, crosscutting several domains (e.g., objects, events, relationships, dispositions, etc.), i.e., what is termed a foundational ontology (aka top-level or upper-level ontology). In this article, we adopt the Unified Foundational Ontology (UFO) given its successful track record of supporting the ontological analysis of complex notions such as *value*, *risk*, *service*, *trust*, *legal relations*, *money*, *decisions*, *economic preferences*, among many others [20, 22, 29].

2.2 An Ontology for Requirements (NFRO)

The Non-Functional Requirements Ontology (NFRO) is defined as an extension of UFO. As such, it adopts the UFO notion of *AGENT*, an entity having mental states such as belief, desire and intention and means to act accordingly. Also, the notion of *INTENTION* that refers to a situation (state-of-affairs) that the *AGENT* commits to bring about by pursuing goals and executing actions. It is also important to state that according to UFO, *AGENT* can be categorized into *HUMAN* (i.e. a person), *ARTIFICIAL* (i.e. artificial systems, such as information systems, cyber-physical systems, etc.) and *INSTITUTIONAL* (i.e. organization). A Stakeholder may be a *HUMAN* or an *INSTITUTIONAL* agent, while the system-to-be is an *ARTIFICIAL* one. For reasons of space, we do not include a figure showing this *AGENT* categorization, but we refer the reader to [21] (chap.3), for details.

Requirements can be functional and non-functional, but the latter are most relevant to ethical requirements, so we focus on them by adopting NFRO [23]. In NFRO, a requirement is a GOAL. Requirements are specialized into NFRs (aka QUALITY GOALS) and FUNCTIONAL REQUIREMENTS (FRs). FRs refer to a FUNCTION (a capability, capacity) that a system can manifest in particular SITUATIONS. NFRs refer to desired qualities taking QUALITY VALUES in particular QUALITY REGIONS. For example, a software system is considered to have good usability if the value associated to its “usability” quality maps to the “good” quality region in the “usability” quality space.

This ontological account delineates different kinds of requirements, and clarifies the nature of NFRs as qualities that map a system artifact into a quality region [23]. Figure 1¹ depicts a selected subset of the NFRO that is relevant here. For an in-depth discussion and formal characterization of QUALITIES, QUALITY UNIVERSALS, QUALITY REGIONS, and QUALITY SPACES, we refer the reader to [19].

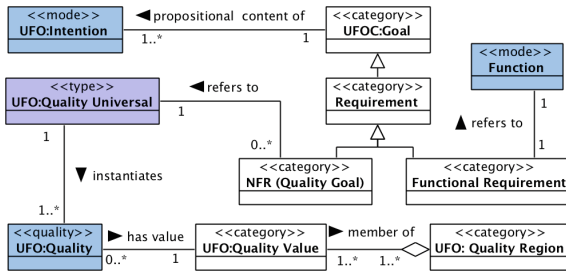


Fig. 1. A fragment of the ontology of non-functional requirements

3 ObRE Method

Figure 2 illustrates the process of the ObRE method, showing the three activities mentioned in Sect. 1.

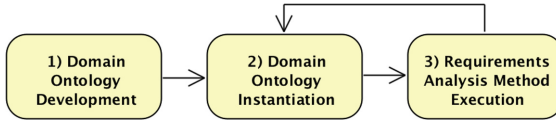


Fig. 2. The ObRE process

The process starts with **1) Domain Ontology Development**, requirements analysts and ontology engineers perform ontological analysis for a class

¹ In all OntoUML diagrams, we adopt the following color coding: types are represented in purple, objects in pink, modes in blue, events in yellow, and abstract entities such as numbers, sets and propositions in white.

of requirements. We emphasize that ObRE does not prescribe that the requirements engineer is versed in the use of ontological analysis concepts. For that, ObRE assumes the presence of an ontology engineer, and the requirements engineer plays a role of a domain expert in the ontology development process. The outcome of activity 1), is an ontology modelled in OntoUML². This activity is performed once for each class of requirements and doesn't need to be repeated for each new system development project. For example, in [7], we conducted ontological analysis the notions of *trust* and *trustworthiness* in order to unpack the meaning of trustworthiness requirements. According to the results of our analysis a system is trustworthy if it is believed to have the capability to perform its required functions (Capability belief) and its vulnerabilities will not prevent it from doing so (Vulnerability belief). Moreover, we define trustworthiness as a composition of three other qualities, namely *reliability* in performing its functions, *truthfulness* in presenting its credentials and *transparency* in its operations. To judge how reliable a system is, we must understand how much of the Stakeholder's Capability Belief is actually met by the system's operations. Note that reliability could have been defined in multiple other ways, for instance, it could have been related to accessibility, i.e., how often will the system be responsive to stakeholder needs; or inferred by the system possessing a specific reliability certificate. The trustworthiness ontology has been recently used in a real case study, reported in [8], showing promising results in defining and monitoring trustworthiness requirements for a particular system. In case a new trustworthy system needs to be developed, the same ontology can be fully reused, and instantiated for the new system-to-be.

Having the requirements explicitly defined and understood, the analyst may perform **2) Domain Ontology Instantiation**. Here, the analysts focus on a particular system and instantiate elements of the ontology. For a security ontology, this step would identify stakeholders, vulnerabilities, attack types, etc. for a particular system. This is intended to serve as domain model for conducting requirements analysis. We highlight the importance of this step, since the same class of requirements may lead to distinct concrete requirements for each system. Thus, instantiating the ontology created in 1) helps identify these particular requirements and opens the way for the system-to-be requirements analysis.

In activity **3) Requirements Analysis Method Execution**, analysts use the domain model to define and analyze system requirements. For instance, she may simply define a requirements table, listing the requirements instantiated with the help of the ontology. Or if she prefers a more sophisticated analysis methodology, she may use goal modeling, defining the contribution of different choices to accomplish a particular goal (i.e., requirement), and specifying how goals relate to each other, as well as to relevant stakeholders' resources and tasks. Or yet, she may create user stories based on the identified ontological instances. From this point on, the requirements analysis may progress as the

² OntoUML is an UML-based language developed to represent UFO's ontological categories, see [19,20] as well as the OntoUML Community Portal <https://ontouml.org/>.

chosen method prescribes, however, with the benefit of having the ontology and ontological instances as guides.

As depicted in Fig. 2, steps 2) and 3) are intended to be carried out iteratively, as with most RE methods. This supports the analyst in revisiting the previous activities while maturing the requirements elicitation and analysis.

4 Applying ObRE to Ethical Requirements for Intelligent Systems

We illustrate the application of ObRE to ethical requirements. But what is ethics after all? And what are the characteristics of an ethical system?

According to the Markkula Center for Applied Ethics “ethics refers to standards of behavior that tell us how human beings ought to act” while playing different roles, e.g. worker, driver, parent, citizen, friend etc. For each role, there are ethical codes of conduct that capture such standards of behaviour [1]. Ethicists and AI researchers have been studying the interplay of ethics and AI systems where the subjects of ethical codes are systems that play such roles, e.g., worker, driver. Floridi et al. [17] proposes five general principles that underlie ethical codes and are role-independent. These have been adopted by the European Commission in a document concerning trustworthy AI [9]. The principles are: *Autonomy* (respect human dignity), *Beneficence* (do good to others), *Non-maleficence* (do no harm to others), *Justice* (treat others fairly), and *Explicability* (explainability, transparently).

We can categorize ethical requirements for a system-to-be as types of *Ecological Requirements* [24], in that they are derived from the ecosystem within which the system-to-be is embedded. After all, it is that ecosystem that determine *values* and *risks*” that can lead to ethical behaviours by the system ([24], p. 253). In a nutshell, value and risk are both types of *dispositions* [29], which are properties that heavily dependent on contextual factors for their manifestation [26]. In fact, as mentioned in Sect. 1, the focus on ethics is motivated by the emerging feeling of risk brought by the use of recent technologies. And these risks must be accounted for and analyzed in contrast with the values delivered by systems and services applying such technologies. For the notions of value and risk, we rely on COVER [29] discussed in Subsect. 4.1³, while Subsects. 4.2 and 4.3 illustrate the application of the ObRE process, the former focusing on the first two ObRE process activities while the latter addresses the last activity.

³ Note that we present COVER here, as opposed to Sect. 2, because the selection of this ontology is attuned to the particular application of ObRE that we choose to illustrate in this section, and the role that *value* and *risk* play in unpacking some of the ethical requirements addressed. Had we chosen a different type of application, e.g., Run-Time Adaptability Requirements [15], a different ontology would have been chosen to play this role.

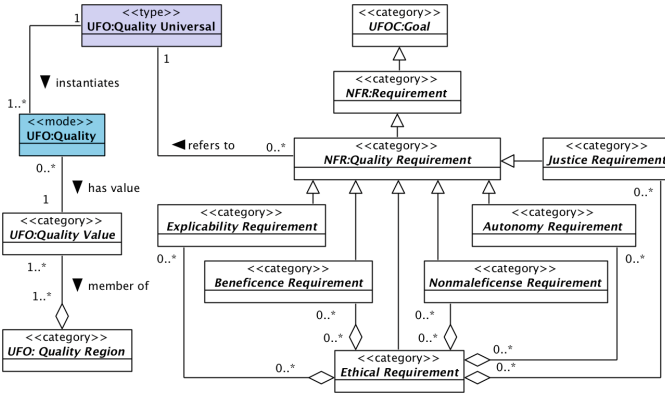


Fig. 4. Ethical requirements

Now, let us interpret ethical requirements in terms of value and risk. Value can be seen as a relational property, emerging from a set of relations between the intrinsic properties of a *value object* (or a value experience) and the goals of a *Value Subject* [29]. The **value** of an object (or experience) measures the degree to which the properties (*affordances*) of that object **positively** contribute (help, make) to the achievement of value subject goals. Mutatis Mutandis, risk is a relational property emerging from a set of relations between the intrinsic properties of an *Object at Risk* (vulnerabilities), as well as *Threat Objects* and *Risk Enablers* (capacities, intentions) and the goals of a *Risk Subject* [29]. The **risk** of an object at risk given threat objects and risk enablers amounts to the degree to which the properties of those entities can be enacted to **negatively** contribute to denting (hurt, break) the risk subject goals. Now, ontologically speaking, affordances, vulnerabilities, capacities, intentions are all types of *dispositions*, which are themselves ecological properties, i.e., properties that essentially depend on their environment for their manifestation [26].

For reasons of space, we are going to analyze two of these sub-requirements here, i.e., those of *beneficence* and *nonmaleficence*. This choice also allows us to contrast these two related NFRs. Considering the definition of beneficence as “doing good to others” [17], we can say that Beneficence Requirements are related to “**creating value**” to stakeholders in the ecosystem in which the system is included. It means that Beneficence Requirements can be seen as goals related to an intention of **positively** impacting the goals of stakeholders in this ecosystem. Analogously, considering the definition of nonmaleficence as “doing no harm to others” [17], we can say that Nonmaleficence Requirements are related to “**preventing risks**” to stakeholders. Consequently, Nonmaleficence Requirements can be seen as goals related to an intention of preventing the occurrence of events that may **negatively** impact stakeholders’ goals.

Events that impact agents’ goals, either positively or negatively, are defined in COVER [29] as Gain Events and Loss Events, respectively. In this sense,

Beneficence Requirements intend to create Gain Events, which positively impact stakeholders' goals. Similarly, Nonmaleficence Requirements intend to prevent the occurrence of Loss Events, which negatively impact stakeholders' goals. Figure 5 represents the OntoUML modeling of Beneficent and Nonmaleficent Requirements.

As presented in Fig. 4, REQUIREMENT is modeled as a GOAL, which is the propositional content of an INTENTION of a stakeholder. We use the notion of agent defined in UFO to model stakeholders. In UFO, agents are individuals that can perform actions, perceive events and bear mental aspects. A relevant type of mental aspect for our proposal is the intention. Intentions are desired state of affairs of which the agent commits to pursuing [11]. In the ontology, INTENTIONS are represented as modes (an externally dependent entity, which can only exist by inhering in other individuals [19]) that inhere in AGENTS. QUALITY REQUIREMENT is a type of Requirement. BENEFICENCE and NONMALEFICENCE REQUIREMENTS are types of QUALITY REQUIREMENTS, which are related to a BENEFICENCE INTENTION and a NONMALEFICENCE INTENTION, respectively. BENEFICENCE INTENTIONS are externally dependent on GAIN EVENTS as their focus of interest is the creation of such events. As previously mentioned, GAIN EVENTS are a type of IMPACT EVENT (as defined in COVER [29]) that positively impact AGENT's goals. NONMALEFICENCE INTENTIONS, in turn, are externally dependent on LOSS EVENTS as their focus of interest is to prevent the occurrence of such events. As aforementioned, LOSS EVENTS are a type of IMPACT EVENT that negatively impact AGENT's goals.

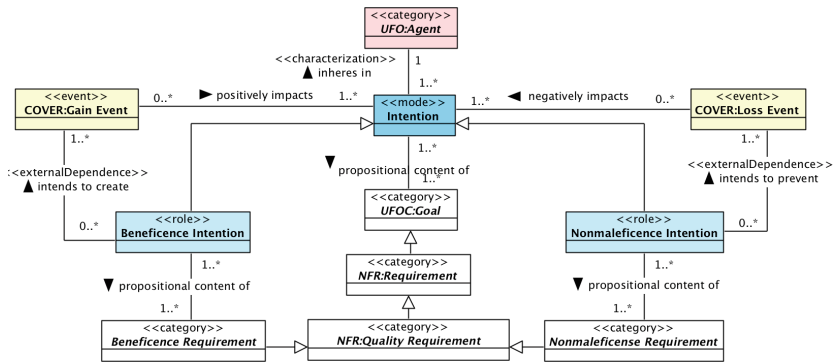


Fig. 5. Beneficence and nonmaleficence requirements

In the sequel, in Fig. 6, we instantiate the ontology with two examples (a Beneficence and a Nonmaleficence Requirement) in the context of driverless cars.

In the first example, the PASSENGER of a driverless car intends “not to be late”. In order to address this, we have the **BENEFACTENCE REQUIREMENT** that “the car should choose quicker route towards destination” related to the **INTENTION** that the “driverless car arrives on time at destination”, which is a **BENEFACTENCE INTENTION** that aims at creating a **GAIN EVENT**. The event “driverless car arrives on time at destination” is a **GAIN EVENT** that positively impact the PASSENGER’s goal of not being late.

In the second example, the PASSENGER intends “feel safe”. In order to address this, we have the **NONMALEFACTENCE REQUIREMENT** that “the car should adopt a defensive driving behavior” related to the **INTENTION** of “preventing aggressive direction”, which is a **NONMALEFACTENCE INTENTION** that aims at preventing the occurrence of a **LOSS EVENT**. The event “passenger feels nervous as the car drives aggressively” is a **LOSS EVENT** that negatively impact the PASSENGER’s goal of feeling safe.

4.3 Requirements Analysis Method Execution

We exemplify activity 3) of the ObRE process by analyzing the requirements of a driverless car faced with ethical dilemmas.

In particular, we present both a requirements table and a goal model for the driverless car case. We start by presenting Table 1, showing how a requirements table may be enriched with the inclusion of columns representing some of the ontological concepts described in the previous subsections. This facilitates requirements elicitation, by using the right concepts for a particular kind of requirement as guides. In the case of ethical requirements, concepts such as impact event (both positive and negative) and ethical principles. All words highlighted in boldface in Table 1 refer to ontological concepts analyzed in Sect. 4.2, while the ontological instances are written as non-emphasized text.

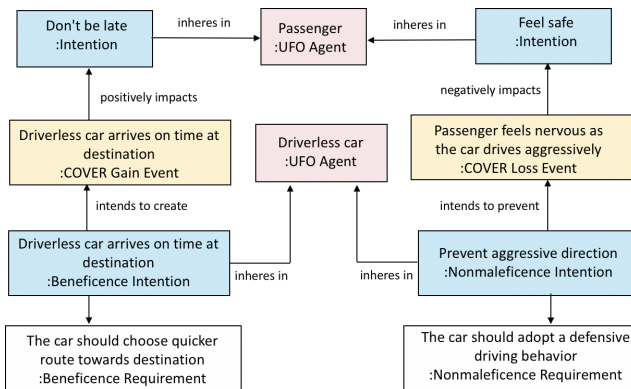


Fig. 6. Ontology instantiation

Note that the ontological analysis of Sect. 4.2 makes very explicit all involved ontological notions used in Table 1, thus supporting the communication and avoiding misunderstandings between the stakeholder and the requirements analyst. For example, having the concepts of GAIN EVENT or LOSS EVENT as well as the specialization of ETHICAL REQUIREMENTS may guide the analyst in asking the right questions during requirements elicitation. This is done by first capturing first the positive and negative impact events concerning Driverless Cars, then relating them with the ethical principles (Beneficence and Nonmaleficence, in this case), and finally coming up with particular requirements for the system-to-be to accomplish such principles. In particular, regarding the latter, these are requirements for the developing of functions and capacities that enable the manifestation of gain events, or that block the manifestation of loss events (e.g., by eliminating the vulnerabilities of the object at risk, or by changing either the intention or the threatening capacities of the threatening agent).

As an alternative, consider a requirements analysis for the Driverless car case using goal modeling. Figure 7 depicts a goal model for this case, using the i^* framework [13]⁴.

Due to space limitations, this model considers only three of the stakeholders referred to in Table 1, namely, *Passenger*, *Pedestrian* and *Nearby Car*. Moreover, the model depicts the dependency of each of these stakeholders and the *Driverless Car*. Many of the dependencies and goals depicted in this model have been already elicited by using the requirements table. For example, with respect to the Passenger, the *reaching destination on time* goal dependency

Table 1. The result of the application of the proposed process in the driverless car case

Stakeholder	ID	Impact event	Principle	Ethical requirement
Passenger	1	Arrive on time at destination (positive)	Beneficence	The car should choose quicker route towards destination
	2	Passenger feels nervous when the car drives aggressively (negative)	Nonmaleficence	The car should adopt a defensive driving behavior
Pedestrian	3	The car runs over a pedestrian (negative)	Nonmaleficence	The car should stop whenever a pedestrian is crossing the road
	4	Pedestrians waiting by a crossroad have priority to cross it (positive)	Beneficence	The car should stop before the crosswalk every time there is a pedestrian waiting to cross it
Bystander	5	Be splashed if the car passes by a puddle of water (negative)	Nonmaleficence	The car should slow down in case there is a puddle of water near a bystander
Nearby cars	6	Be hit (negative)	Nonmaleficence	The car should slow down when it gets around 20 m in the rear of a nearby car
			Nonmaleficence	The car should make enough distance when overtaking a car
Environment	7	Be polluted (negative)	Nonmaleficence	The car should turn off the motor every time it stops

⁴ The model was drawn using the piStar tool, available at <https://www.cin.ufpe.br/~jhcp/pistar/>.

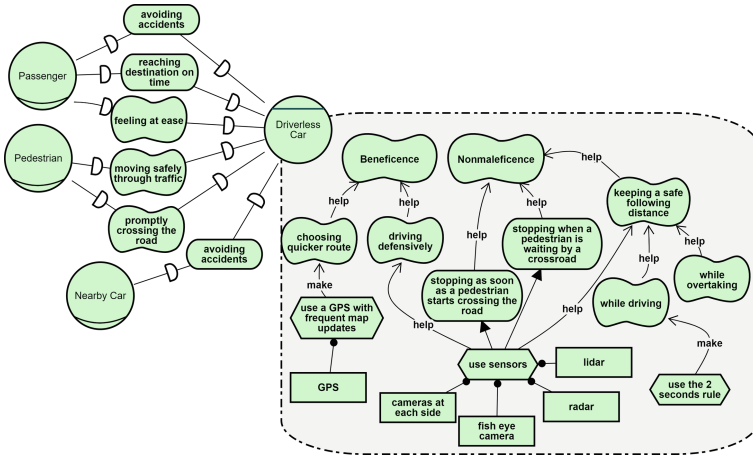


Fig. 7. The driverless car requirements model using i^*

relates to the positive impact event elicited to Passenger (see Table 1, first line), while the *feeling at ease* dependency relates to the negative impact captured for this same stakeholder (see Table 1, second line). Nevertheless, new dependencies have been added, for instance, when drawing the model, we realized that avoiding accidents dependency (previously only attributed to the Nearby Car stakeholder, see Table 1, line 6) is also relevant for the Passenger⁵.

Besides dependencies, the goal model of Fig. 7 depicts the internal perspective of the Driverless Car, assisting in the analysis of the system's requirements. Note that both ethical principles of Beneficence and Nonmaleficence are represented there by qualities (consistent with our ontological notion of NFR). Then, for each of these qualities, more specific goals and qualities are identified and related to them by contribution links. For instance, the *choosing quicker route* quality helps (i.e. partially contributes to) the achievement of Beneficence. Additionally, *choosing quicker route* may be indirectly related to the *reaching destination on time* goal dependency of the Passenger.

The goal model also allows the requirements analyst to progressively identify more concrete requirements and solutions and the resources needed to accomplish them. For example, the *use a GPS with frequent map updates* task makes (i.e. fully accomplishes) the *choosing quicker route* quality, and the *GPS* itself is a resource needed in this task.

⁵ We did not update our table on purpose, since although that would make both models more consistent, this is an interesting case in which the visualization of the goal model and its particular constructs (in this case, dependency, goals and qualities) helped us realized a missing requirement for one of the stakeholders. In this paper, the authors are playing the role of the requirements analyst, but cases such as this one may easily happen in practice.

Another task worth clarifying is *use the 2 second rule*. This is a well-known rule for maintaining a safe distance between vehicles. It is adopted in some countries as a good code for driver conduct for human drivers [2], and it can also be adopted as a requirement for driverless cars. Note that this task makes the *keeping a safe following distance while driving* quality. However, to accomplish the higher level *keeping a safe following distance* quality, other tasks and qualities are involved.

The reader may have noticed that each of the RE approaches has its advantages and limitations. For example, the relation between the impact events, principles and ethical requirements are easier to spot in the requirements table, much easier and fast to create in comparison with the goal model. The goal model, however, makes more explicit which intention (and thus which requirement) is related to each of the agents involved in our case. Moreover, it is visual and it allows a much more detailed requirements analysis, in terms of more and less abstract requirements, solutions and needed resources.

We emphasize that ObRE does not subscribe to a specific RE method, leaving this choice for the requirements analyst, based on their particular preference or skill. Another important point is that the choice for the RE approach may be taken based on the approach's underlying modeling languages. For instance, a language offering the concepts of threat, value etc. may be a preferred choice here.

5 Related Works

We examine related works in two directions. First, we take a look at ontology-based methods for RE, especially those targeting NFRs, as these kinds of requirements are the main focus of ObRE. Next, we investigate works that aim at embedding systems with ethics.

ElicitO [6] is an ontology-based tool aimed at providing guidance during requirements elicitation, conducting the requirements analyst in performing a precise specification of NFRs. Taking a similar direction, the work of Velela and Cysneiros [30] provides an ontology-based tool to help identify NFRs, making explicitly their interdependencies and possible conflicts. Hu et al. [25] also aim at detecting conflicts between NFRs, and conduct a trade-off analysis in case such conflicts arise. This is done by representing NFRs in a softgoal interdependency graph, which is formalized using an ontology. All these works follow a different path in comparison to ours, focusing much more on the automation of requirements analysis by the means of representing NFRs using OWL ontologies. Our work, on the other hand, uses reference ontologies to provide a deep understanding of NFRs whose semantics are usually subjective and complex, by interpreting these NFRs according to the particular domain of the system-to-be. And by the means of this interpretation, our work attempts to guide the requirements analyst in defining requirements that will support the analyzed NFRs.

Nowadays, many researchers have been busy trying to come up with frameworks and approaches targeting responsible AI and the development of systems

embedded with ethics. Interesting initiatives are those of Rashid, Moore, May-Chahal and Chitchyan [28], Peters, Vold, Robinson and Calvo [27], Etzioni and Etzioni [16], Dignum [14] and Floridi et al. [17]. The latter has been proposed by several specialists, and has served as basis for the European Union Ethics Guidelines for Trustworthy AI [3]. All these cited research works bring very relevant insight on how to develop ethical systems. However, their proposed frameworks and guidelines are still in an abstract level, and we believe that approaches specifically targeted at Requirements Engineering are still an open issue. Our proposal is proposed with the goal of filling in this gap.

6 Final Remarks

In this paper, we illustrate how a novel RE method named ObRE is able to elicit ethicality requirements. In particular, ObRE precisely defines the concepts that underlie a class of quality requirements (NFRs) through an ontology and offers these concepts for requirements analysis. ObRE aims to address the subjective and ambiguous nature of many classes of requirements, especially the ones that have become prominent recently with the advent of AI systems. As a result, ObRE facilitates the communication between analysts and stakeholders, besides assisting in the identification of requirements.

It is important to note that our approach does not prescribe a specific way to implement the analyzed requirements in the system, for example, by developing a rule-based system, or by having the requirements hardcoded. ObRE focuses solely on the RE activity, supporting the elicitation of requirements, which can then be analyzed, validated and monitored throughout the system's life cycle.

The success of RE activities largely depends on the creation of a shared understanding between stakeholders and analysts for a system-to-be [12, 18]. Werner, Li, Ernst and Damian [31] conducted an empirical study to find out why a shared understanding NFRs is so difficult in software organizations. Their study shows that two of the main problems were lack of domain knowledge and inadequate communication. They report on some interesting findings that we believe could be alleviated by the application of ObRE, e.g., i) some NFRs are considered complicated and out of the developers expertise; ii) there is no clear understanding of what particular NFRs mean; and iii) when two or more people are working simultaneously in the same system, even if they communicate, they end up approaching a given NFR in a different way. Ethical requirements, which are the focus of this paper, fit precisely into the situations just described.

Our agenda for the future includes, firstly, a full fledged implementation and validation of the ObRE method, by doing real case studies and having experts evaluate the results. Another interesting research direction is extending the ethical requirements ontology to deal with ethical conflicts. Many problems arise when intelligent systems face situations that involve ethical conflicts. For example, for the driverless car, what happens if the system needs to choose the lesser of two evils, such as either running over a bystander or a pedestrian? The principles we adopted so far in our ontological analysis do not seem to account alone for such cases, and we plan to address this limitation in the near future.

References

1. A framework for ethical decision making (2015). <https://www.scu.edu/ethics/ethics-resources/ethical-decision-making/a-framework-for-ethical-decision-making/>
2. The 2-second rule. In: Learn the Road Code (2016). <https://drive.govt.nz/get-your-learners/interactive-road-code/>
3. Ethics guidelines for trustworthy AI (2019). <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
4. The 7 Most Pressing Ethical Issues in Artificial Intelligence (2019). <https://kambria.io/blog/the-7-most-pressing-ethical-issues-in-artificial-intelligence/>
5. IEEE standard model process for addressing ethical concerns during system design. IEEE STD 7000-2021, pp. 1–82 (2021). <https://doi.org/10.1109/IEEESTD.2021.9536679>
6. Al Balushi, T., Sampaio, P., Dabhi, D., Loucopoulos, P.: ElicitO: a quality ontology-guided NFR elicitation tool. In: 13th International Working Conference on Requirement Engineering: Foundation for Software Quality, pp. 306–319 (2007)
7. Amaral, G., Guizzardi, R., Guizzardi, G., Mylopoulos, J.: Ontology-based modeling and analysis of trustworthiness requirements: preliminary results. In: Dobbie, G., Frank, U., Kappel, G., Liddle, S.W., Mayr, H.C. (eds.) ER 2020. LNCS, vol. 12400, pp. 342–352. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-62522-1_25
8. Amaral, G., Guizzardi, R., Guizzardi, G., Mylopoulos, J.: Ontology-based requirements engineering: the case of trustworthiness requirements. In: 40th International Conference on Conceptual Modeling, pp. 257–267 (2021)
9. European Commission High-Level Expert Group on Artificial Intelligence: Draft ethics guidelines for trustworthy AI, draft document (2018)
10. Berto, F., Plebani, M.: Ontology and Metaontology: A Contemporary Guide. Bloomsbury Publishing, London (2015)
11. Castelfranchi, C.: Commitments: from individual intentions to groups and organizations. In: ICMAS, vol. 95, pp. 41–48 (1995)
12. Charaf, M., Rosenkranz, C., Holten, R.: The emergence of shared understanding: applying functional pragmatics to study the requirements development process. Inf. Syst. J. **23**(2), 115–135 (2013)
13. Dalpiaz, F., Franch, X., Horkoff, J.: iStar 2.0 language guide. [arXiv:1605.07767](https://arxiv.org/abs/1605.07767) [cs.SE] (2016). <http://dalp-fran-hork-16-istar.pdf/>
14. Dignum, V.: Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way. Springer, Cham (2019). <https://doi.org/10.1007/978-3-030-30371-6>
15. Duarte, B.B., et al.: Ontological foundations for software requirements with a focus on requirements at runtime. Appl. Ontol. **13**(2), 73–105 (2018)
16. Etzioni, A., Etzioni, O.: Incorporating ethics into artificial intelligence. J. Ethics **12**, 403–418 (2017)
17. Floridi, L., et al.: AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. Minds Mach. **28**, 689–707 (2018)
18. Glinz, M., Fricker, S.: On shared understanding in software engineering: an essay. Comput. Sci. Res. Dev. **30**(3–4), 363–376 (2015)
19. Guizzardi, G.: Ontological foundations for structural conceptual models. Telematica Instituut/CTIT (2005)

20. Guizzardi, G., Wagner, G., Almeida, J.P.A., Guizzardi, R.S.S.: Towards ontological foundations for conceptual modeling: the Unified Foundational Ontology (UFO) story. *Appl. Ontol.* **10**(3–4), 259–271 (2015)
21. Guizzardi, R.: Agent-oriented Constructivist Knowledge Management. Ph.D. thesis, University of Twente, Netherlands (2006)
22. Guizzardi, R., Carneiro, B.G., Porello, D., Guizzardi, G.: A core ontology on decision making. In: *Proceedings of the 13th Seminar on Ontology Research in Brazil* (2020)
23. Guizzardi, R. et al.: An ontological interpretation of non-functional requirements. In: *Proceedings of FOIS*, vol. 14, pp. 344–357 (2014)
24. Guizzardi, R., et al.: Ethical requirements for AI systems. In: *Proceedings of Canadian AI 2020*, pp. 251–256 (2020)
25. Hu, H., et al.: Semantic modelling and automated reasoning of non-functional requirement conflicts in the context of softgoal interdependencies. *IET Softw.* **9**, 145–156 (2015)
26. Mumford, S.: *Dispositions*. Clarendon Press, Oxford (2003)
27. Peters, D., Vold, K., Robinson, D., Calvo, R.: Responsible AI: two frameworks for ethical design practice. *IEEE Trans. Technol. Soc.* **1**(1), 34–47 (2020)
28. Rashid, A., Moore, K., May-Chahal, C., Chitchyan, R.: Managing emergent ethical concerns for software engineering in society. In: *37th IEEE International Conference on Software Engineering*, pp. 523–526. IEEE (2015)
29. Sales, T.P., Baião, F., Guizzardi, G., Almeida, J.P.A., Guarino, N., Mylopoulos, J.: The common ontology of value and risk. In: Trujillo, J.C., et al. (eds.) *ER 2018*. LNCS, vol. 11157, pp. 121–135. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00847-5_11
30. Velela, R., Cysneiros, L.M.: Towards an ontology-based approach for eliciting possible solutions to non-functional requirements. In: Giorgini, P., Weber, B. (eds.) *CAiSE 2019*. LNCS, vol. 11483, pp. 145–161. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-21290-2_10
31. Werner, C., Li, Z., Ernst, N., Damian, D.: The lack of shared understanding of non-functional requirements in continuous software engineering: accidental or essential? In: *28th IEEE International Requirements Engineering Conference, RE 2020*, Zurich, Switzerland, 31 August–4 September 2020, pp. 90–101. IEEE (2020)