Development and Evaluation of Ontology for Intelligent Decision Support in Medical Emergency Management for Mass Gatherings

Pari Delir Haghighi\textsuperscript{a,b}, Frada Burstein\textsuperscript{a}, Arkady Zaslavsky\textsuperscript{b} and Paul Arbon\textsuperscript{c}

\textsuperscript{a}Monash University, Australia
\textsuperscript{b}CSIRO, Australia
\textsuperscript{c}Flinders University, Australia

Abstract
Conducting a safe and successful major event highly depends on the effective provision of medical emergency services that are often offered by different public and private agencies. Poor communication and coordination between these agencies and teams can result in delays in decision-making and duplication of efforts. Another related issue is that emergency decisions are usually made based on individual experience and domain knowledge of relevant managerial personnel. For sustainable knowledge management and more intelligent decision support it is beneficial to collect, consolidate, store and share these experiences in a form of a knowledge base or domain ontology. State-of-the-art surveys identify this gap that there is no common ontology describing the domain knowledge for planning and managing medical services in mass gatherings. Part of the reason is that the process of construction of such an ontology is not a trivial task. In this paper, we describe the process of developing and evaluating a Domain Ontology for Mass Gatherings (DO4MG) with a focus on medical emergency management. As part of the evaluation, we illustrate the application of DO4MG for implementing a case-based reasoning decision support for emergency medical management in mass gatherings. Such an implementation demonstrates the potential of using ontology for resolving terminology inconsistencies and their usefulness for supporting communication between medical emergency personnel in mass gatherings. We also illustrate how this ontology can be applied to different stages of medical emergency management as part of a system architecture. The

\textsuperscript{1} Corresponding Author: Faculty of Information Technology, Monash University, 900 Dandenong Rd, Caulfield East, 3145, Victoria, Tel: +613 9903 2355, Fax: +613 9903 1077, Australia, Pari.DelirHaghighi@monash.edu
lessons learnt from building DO4MG for this domain could be beneficial in general to the theory and practice of intelligent decision support and knowledge management in complex problem domains.

**Keywords:** Intelligent Decision Support, Ontology Development and Evaluation, Medical Emergency Management, Mass Gatherings, Case-Based Reasoning, Knowledge Management

1. **Introduction**

   Medical emergency decision making could be a challenging task, particularly during mass gathering events [2, 75]. In mass gatherings, when a crisis occurs, medical emergency decisions are usually made under time pressure [3]. Further, these events typically involve participation of various emergency medical agencies that frequently use different terminology to represent the same concepts. This inconsistency can complicate communication between different emergency teams as well as integration and management of data recorded in these events.

   Intelligent decision support systems aim to provide decision makers with timely, useful and valid information based on some pre-coded domain knowledge [12]; and Medical Emergency Services (MES) in mass gatherings can benefit from access to such systems [26, 74, 75]. The underlying condition for successful development of such a system is creation of a reliable mechanism for collection, representation, and storage of domain knowledge. Literature shows that there is no common ontology describing the domain knowledge for planning and managing medical services in mass gatherings.

   We suggest that better knowledge management for decision support can benefit from standardization of mass gathering’s Medical Emergency Management (MEM) terminology using domain ontology [64]. The importance of incorporating an ontology into a system architecture has been well recognized, in the context of intelligent decision support, as the means of knowledge representation and management and to assist decision makers with complex problem-solving [15, 21, 36]. With regards to mass gatherings, ontologies can improve coordination and interaction between different emergency agencies and facilitate data capture, storage, integration and querying of recorded data in the events. Moreover, the use of a common and unified domain ontology can improve the
decision making process where most of the emergency decisions are dependent on individual experiences and domain knowledge of relevant managerial personnel.

The need for further research in knowledge management and decision support in medical emergency management is well recognised [2, 26, 71]. A domain ontology for medical emergency management is a mechanism for providing a consistent view on the problem domain that can be used by all concerned stakeholders. Delir Haghighi et al. [19] introduced a domain ontology, named DO4MG (Domain Ontology for Mass Gatherings), and discussed how it could improve decision making in the field of medical emergency management by providing a unified and common knowledge base for intelligent decision support. However, one of the major requirements to adopt a domain-specific ontology is to determine its fitness and suitability over other existing ontologies in the same field and to evaluate the given ontology against certain criteria and a set of standards [10, 70]. Ontology evaluation requires the use of explicit and formal criteria that match ontology construction objectives.

During the years, many evaluation approaches and criteria have been proposed to analyse and validate ontologies [10]. These evaluation methods normally focused on specific domain problems, and were tested for validating a certain type of ontology. These past studies confirmed that it is important to select an evaluation approach which fits the given ontology and its application domain. In this paper we present an approach for creating and evaluating the domain ontology for medical emergency management. We undertake the task of an extensive review of the existing approaches for ontology evaluation and propose a systematic process suitable for the mass gathering ontology. This includes assessment of the proposed ontology by domain experts and from its application perspective. We describe the application of the ontology within the overall mass gathering management as an integral part of knowledge management for running a safe event [17]. An example of a Case-Based Reasoning (CBR) intelligent decision support implementation of this architecture is also presented.

In this paper, we make two main contributions as follows.

- First, we present a thorough review of current ontology evaluation methods on the basis of multiple criteria and describe the process of selecting the ones applicable for validation of DO4MG. These include the process of empirical testing of the evaluation method with the
domain experts. In doing so, we also report on the lessons learned and therefore provide a refined understanding of ontology evaluation methods.

- Second we describe an ontology-based system architecture for medical emergency management in mass gatherings that incorporates the DO4MG ontology, which illustrates effectiveness of the proposed approach for ontology construction and evaluation. As an example, a CBR prototype decision support system (DSS), which can be used during the pre-event and post-event stages of mass gathering events for respectively training, workload estimation and data collection and integration is described. This DSS uses the ontology to provide a unified and standard vocabulary of the domain and to overcome any problems that can arise from inconsistencies in terminology and discrepancies of data collection in emergency management for mass gatherings.

Both contributions, although validated in a specific problem domain, can provide a useful insight for researchers and practitioners in dealing with complex decision situations, which involve multiple agencies and a wealth of expert knowledge. Through the use of domain ontologies, these complex situations can be better managed and their underlying domain knowledge can be better understood and managed. The development and evaluation of such domain ontologies could follow a systematic approach as described in this paper.

The rest of this paper is organized as follows. Section 2 defines ontologies and presents the benefits of ontologies for intelligent decision support systems. Section 3 provides an overview of decision support systems for MEM in general and then discusses the application of ontologies to medical management emergency in mass gatherings. Section 4 describes the development process that we propose for building DO4MG including pre-development and design stages. Section 5 presents an overview of the current ontology evaluation approaches and provides justification for selected evaluation methods for DO4MG. Section 6 describes the application of the evaluation and refinement methods for verifying the contents and testing usability of DO4MG. These include criteria-based and application-based evaluations. The criteria-based evaluation includes the refinement of DO4MG according to the domain experts’ feedback. The application-specific evaluation of DO4MG details the overall architecture for intelligent decision support in mass gatherings and an illustration of the case-
based reasoning prototype developed to test its usability. Finally, Section 7 concludes the paper and suggests some directions for further research.

2. The Role of Ontologies in DSS

An ontology presents ‘a shared and common understanding of the knowledge domain’ [15:786] using major concepts and terms applied in that domain and identifies the relationships between these concepts. Ontologies enable aggregation and use of knowledge items and sub-processes and provide a way to move from a document-oriented view of knowledge management to a content-oriented view [63]. An ontology provides a world view and the shared understanding of a given domain which can be used as a unifying framework to address the domain problems [66]. As Gruber suggests “an ontology is an explicit specification of a conceptualization” [29:1].

The core of any decision support system is knowledge from which, and of which, decisions are made [12]. To provide a structured and formal representation of knowledge, ontologies have been applied to a number of DSS [15, 48, 51]. For example, Musen et al., [51] proposed EON architecture as a decision support system for protocol-based therapy. EON integrates an ontology that represents clinical protocols such as drug therapy to benefit from a shared and computer-based representation of all the common data elements in a precise and consistent structure. EUEDE (End-User Enabled Design Environment) [48] for dairy farm management applied semantic ontology tools to achieve effective decision systems, which are context sensitive to end-user factors and provided a generic knowledge model applicable across rural industries.

OntoWEDSS (Ontology-based Wastewater Environmental Decision-Support System) [15] uses ontologies to solve complex problems related to environmental science and engineering. The inclusion of ontologies in that study allowed improved modeling of wastewater treatment processes and facilitated the communication among different components of the environmental DSS. The advisory system for crime investigation processes proposed by Dzemydiene and Kazemikaitiene [21], used ontologies to ensure that the crime information was extracted and represented in a structured model format appropriate for decision-making in crime investigation.
Most of these examples do not report on the ontology evaluation process, neither provide enough
details on general principles that can be applied to developing and evaluating other ontologies. At the
same time prior research conducted by Sujanto, et al. [64] for example, demonstrated the importance
of a rigorous process of ontology development and the lack of consistency in current approaches to
ontology evaluation. Sujanto et al [65] proposed a framework for development and evaluation of
medical emergency management ontology from the generic design science principles.

In the work presented in this paper, we focus on qualitatively extending the approach introduced by
Sujanto et al [64] for medical decision support in the mass gathering context. We also propose a
systematic validation methodology which includes formal methods and uses expert feedback for
validation and refinement of the ontology for decision support.

The next section introduces the context for our research and discusses the importance of intelligent
decision support in medical emergency decision making.

3. Intelligent Decision Support for Medical Emergency Management

Decision support systems are in high demand when users need to make informed decisions
especially during emergency situations [12, 74]. Emergency situations are time-constrained and
dynamic environments that change rapidly [9]. Access to up to date information and data is critical in
emergency management decision support, in particular to determine the priorities for operation and
resource allocation and management [9, 23, 74]. Examples of such systems include the Australian
NSW Fire Brigades I-Zone planning system [14] and Gold Coast DSS for flood emergency
management [50].

A common approach to decision support is employing analytical models that usually result in
generation of potential solutions, evaluation/comparisons of these options based on some parameters
(criteria) and selecting the best option accordingly [12, 17]. In emergency situations, due to the time
pressure and uncertainty, it is argued that intuitive decision processes such as Recognition-Primed
Decision (RPD) model can result in higher performance compared to analytical models [37, 57]. The
RPD approach describes how experienced people make decisions in uncertain and time-critical
environments by extracting a course of action that matches the occurring situation and implementing
it. Typically emergency incidents are not entirely identical to each other but the knowledge of past incidents enables emergency personnel and commanders to recognise a similar situation and tailor their strategies accordingly by taking a course of action that experience has shown is effective and successful [41]. The experience can be gained through attendance at many different emergency situations and, to a lesser extent through studying procedures and past incidents [16].

Case-based reasoning [39] is one of the approaches used for intelligent decision support [6], which can benefit decision making in emergency situations. CBR is used in decision making that produces solutions by providing access to past events in a case base and retrieving the experiences (i.e. cases) that are similar to the ‘problem case’ and providing solutions that were successfully applied in similar situations in the past [16, 41]. The effectiveness of CBR can be further improved by the application of ontologies as a mechanism for reasoning about the domain concepts and dealing with the inconsistencies that can arise in the applied vocabulary when multiple agencies are involved. In Section 6 we describe an overall architecture for intelligent decision support in emergency management and use CBR as part of the application-based validation of the ontology.

The next section describes the knowledge management needs in mass gatherings, as well as the context in which the decisions are typically made.

3.1 The Need for Ontology-Based Intelligent Decision Support for MEM in Mass Gatherings

Mass gatherings are defined as a temporary collection of large numbers of people at one location or over multiple sites. Examples include music concerts, sporting events, cultural gatherings, parades and etc. During emergencies in mass gathering events, the medical emergency teams and other emergency personnel need to make complex and time-critical decisions. These decisions can relate to timely treatment of injured or ill spectators, more advanced levels of medical care, which requires rapid evacuation of patients to nearby hospitals, requests for external and additional resources and maintaining the safety of the crowd [2, 73]. Making such emergency decisions under time pressure can be facilitated by using appropriate decision support systems that cater for mass gathering emergency problems. Ontologies can also be utilized in developing DSS for mass gathering to further assist emergency teams and services with the decision-making process.
Planning and managing a safe mass gathering event requires involvement of various stakeholders such as emergency medical services, police, security personnel, ambulance services and first aiders. Different agencies and services frequently use different terminology and employ emergency plans that may be established independent of other agencies. This fact increases the complexity of coordination and interaction between the involved teams during the emergencies. Inter-organisational collaboration and coordination is essential to ensure the provision of appropriate and timely medical care and to maintain a safe working environment [72]. A mass gathering ontology can be applied and shared across all events and facilitate coordination and communication between different teams [45, 58].

The role of ontologies in supporting knowledge sharing activities has been emphasized by many researchers [29, 32]. This strength of ontologies has been particularly recognized in crisis and emergency systems where it is imperative to share knowledge in order to establish effective coordination between stakeholders and reduce ambiguities during decision making [45]. Examples of systems that apply ontologies include rescue operation management [40], emergency evacuation planning [42], emergency alerts [45] and emergency response [58].

An ontology for mass gathering provides a common understanding of mass gatherings, their characteristics, and the relationships between them. Its inclusion in medical emergency decision support systems enables dealing with the inconsistencies and discrepancies that could arise from data modeling and management by various researchers and agencies. Since to the best of our knowledge, there is no standard ontology for this domain, we undertook a systematic review of the relevant approaches from other domains. In the next section we present an overview of the ontology development process, which was applied to construct and evaluate Domain Ontology for Mass Gatherings - DO4MG.

4. Development of DO4MG

Ontologies can be built from scratch or they can reuse existing ontologies. Holsapple and Joshi [34] classify ontology design approaches into five categories. These include 1) inspiration approach that is based on individual creativities and personal views, 2) induction approach where the ontology is created by observation and analysis of a particular case in that domain, 3) deduction technique applies
general principles and adapts them according to a specific case, 4) synthesis approach first identifies a set of ontologies and then synthesizes them with other related concepts and 5) collaboration approach is concerned with a joint effort and using the group members experience and opinions to build the ontology. Our ontology design can be considered as a hybrid approach that combines inspiration, induction and collaboration techniques as described in this section.

Methodologies provide a systemic and repeatable guideline for building ontologies and enable sharing, reuse and extension of the ontologies by others [24, 25, 31, 34]. Some example methodologies introduced for specific domains are: i) enterprise modelling processes ontology proposed by Uschold and King [66], ii) Methontology for the domain of chemicals proposed by Fernández-López et al. [24, 25], iii) the generic guidelines proposed in Ontology 101 [52], iv) the knowledge metaprocess introduced by Staab et al. [63] that targets knowledge management applications (implemented), v) the OntoClean methodology for validation of ‘the adequacy and logical consistency of taxonomic relationships’ [32:201], and vi) the methodology for creating business ontology supporting semantic interoperability [53].

Development of a domain ontology for medical emergency management in mass gatherings requires a rigorous and inclusive ontology development approach as well as an ontology evaluation method involving domain experts. Studying the above-mentioned methodologies reveals that there are certain steps in these methodologies that are relatively common. The four overlapping stages that we have identified include: i) specifying the ontology scope and objectives, ii) knowledge acquisition and identifying key concepts, iii) building/coding, and iv) evaluation. We have followed similar steps in DO4MG ontology development.

Fig.1 depicts the main steps in the process of ontology development and evaluation. These steps will be discussed and exemplified in the context of a domain ontology for medical emergency management in mass gatherings in this paper. In addition to these steps, some other approaches such as those proposed by Bernaras et al. [8] and Staab et al. [63] explicitly specify a stage for refinement of the ontology based on expert feedback to improve the final product. Taking into account the context of our ontology, emergency management, it was very important to also consider the
refinement phase to ensure the best quality outcome. This stage was incorporated as part of the evaluation phase and is described in Section 6.

Fig. 1. An overview of step-by-step process of ontology development and evaluation.

The next subsections discuss the proposed ontology development stages that have been followed for development of DO4MG.

4.1 Pre-Development Stages of DO4MG

Identifying the Scope and Objectives

As a first step we have identified the scope and objectives of the DO4MG. The ontology focuses on MEM for mass gathering events and therefore its users include emergency medical services such as first aiders, nurses, ambulance services and other involved teams in mass gathering events. The primary purpose of the ontology is to provide a better understanding of these events and enable access to domain knowledge for study and analysis. For instance, the ontology can be used for developing knowledge management applications (including an online knowledge base) to enable stakeholders to study, utilize and extend the ontology collectively. The DO4MG offers a standard knowledge structure and model (common terminology) that can be shared and used between different stakeholders, applicable to any type of emergency or mass gathering event. It can be utilized as a
unified knowledge base for improving communication and interaction between different emergency
teams and for assisting decision makers with event planning and resource allocation. Moreover, it can
facilitate integration of mass gathering data and its management (i.e. storage and querying).

Knowledge Acquisition

After identifying the ontology’s scope and objectives, the second stage, i.e. knowledge acquisition,
involves discovering, eliciting and extracting knowledge from the domain of interest. The resources
that we have used at this stage included primary journals and conference papers for emergency and
crisis management (e.g. conference proceedings for the Integrative and Analytical Approaches to
Crisis Response and Emergency Management Information Systems (ISCRAM) and the Australian
Journal of Emergency Management (AJEM) were used as primary starting points). In addition, public
reports, and government manuals were used such as the Emergency Management Australia (EMA)
manuals series for mass gatherings, and the ‘Compendium of Mass Gatherings’ that includes a
collection of papers mainly from the Prehospital and Disaster Medicine journal (PDM), an Official
Publication of the World Association for Disaster and Emergency Medicine, collected by its
President, Professor Arbon (2009). Delir Haghighi et al [18] presented an overview of mass gathering
concepts extracted from the literature (see Table 1 below). Each concept was further discussed with
the experts, defined, and documented including the reference sources.

Table 1 Mass gathering concepts and characteristics extracted from the current literature.

<table>
<thead>
<tr>
<th>Reference source</th>
<th>Concepts and Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milsten et al. [49]</td>
<td>Weather, attendance, event duration, whether event occurring indoors/outdoors, seated or mobile, event type, crowd mood, alcohol or drugs, crowd density, locale/physical plant, age</td>
</tr>
<tr>
<td>Arbon [2, 3]</td>
<td>Psychological domain (i.e. crowd behavior and mood, individual motivation and behavior, crowd interests and culture, attendance reason, duration, use of alcohol or drugs); Biomedical domain (i.e. health status, latent potential for illness/injury, age, heat or cold-</td>
</tr>
</tbody>
</table>
related physiology, alcohol or drug-related physiology); Environmental domain (i.e. crowd attendance/density, venue, event type, outdoor/indoor weather, availability of alcohol or drugs)

Parrillo [54] Structure and location, nature of event, crowd size and demographics, environmental factors, transportation, equipment and involved agencies such as police


Zeitz et al. [73] Crowd mood: passive, active and energetic

De Lorenzo [20] Weather, duration, mobility, crowd mood, crowd density, and alcohol and drugs

AEM, Manual 2, [4] Venue, crowd movement, hazards, event type, legal issues, police and security, defence assistance, safety issues (i.e. emergency response plan, indoors/outdoors, load capacity, seating, emergency tools, fire safety, communication systems, OHS), crowd control (e.g. entrance and exits, barriers, seating, alcohol, drugs and weapons), public health (e.g. food safety, water, infection control), medical care (e.g. ambulance, medical teams and equipment), psychological dimension, high risk events

In addition to the above-mentioned resources, we also gathered the domain knowledge and key concepts and elements from domain experts and researchers working in the field of EM during individual interviews (with three experts) and a focus group with 10 participants (discussed in Section 6.1).

4.2 Design and Implementation of DO4MG
The DO4MG ontology is implemented in Protégé 4.0, which supports OWL (Web Ontology Language), a common ontology language used to define and describe the concepts (classes), subclasses, properties, and associated relationships of the domain of interest. The core of DO4MG is the concept of Mass Gathering. There are five main key concepts, in the DO4MG ontology, which define every mass gathering event. These include CrowdFeatures, EventVenue, GatheringType, EnvironmentalFactors and MassGatheringPlan (shown in the left hand side of Fig. 2). The second level of the ontology includes 38 subclasses, i.e. “children” or “leaf classes”, which are broken into further subclasses. The total number of classes considering all the levels is 234. The encoding of the elements includes definitions for most of the domain concepts and provides the references for sources of its origin. The slots/properties, instances and relations are defined for the concepts for better understanding and to enable querying and reasoning about the concepts. Fig. 2 shows an example of using protégé to define relationships, instances and properties for the DO4MG ontology.

![An example of properties and instances used in DO4MG.](image)

**Fig. 2.** An example of properties and instances used in DO4MG.

The PersonalBrain5.5 system has been used to visualize and present the entire ontology. The software has extra features which enable to attach files of the referenced papers used in defining the concepts, which is useful for the ontology long-term management and documentation. Fig. 3 depicts the use of the PersonalBrain for codification and visualization of DO4MG. The figure

---


shows the main classes of the ontology and some of their subclasses. The software provides an easy and flexible interface to visualise and traverse the classes.

Fig. 3. An overview of the DO4MG using PersonalBrain system.

This concludes the process of ontology development. The next section provides an overview of ontology evaluation approaches and justification for the selection of evaluation methods for DO4MG.

5. DO4MG Ontology Evaluation

Due to the increasing availability of ontologies, it is imperative to evaluate existing ontologies as the basis of designing new, and to determine whether the ontology is suitable for representing certain domains and applications within it [69]. Ontology evaluation requires use of proper and formal evaluation criteria and methodologies. Therefore, it is important to select and apply an appropriate evaluation approach which is fitting to the given ontology and its application domain. The next subsection provides a critical review of ontology evaluation approaches and derives the one, which we consider suitable and applied it for DO4MG evaluation.

5.1 An Overview of Ontology Evaluation Approaches

During the years, researchers have proposed a variety of approaches for evaluating ontologies. Brank et al. [10] categorizes these approaches into four main classes as follows:
1. **The ‘gold standard’ evaluation** - This approach compares the ontology to high-level and ‘golden’ standards which can be an ontology itself [44]. For example, Abramowicz et al. [1] describe the gold standard built ontology based on the input obtained from an expert group during a series of workshops. With regard to the gold standard evaluation, in some scenarios, access to such standards (or ontology) or its provision may not be possible. Furthermore, the evaluation results can suffer from the flaws in the applied comparison methodology or from inappropriateness of the gold standard [11].

2. **Data driven evaluation** - This assessment method compares the ontology with a source of data such as a corpus [11]. The data-driven approach compares the ontology to a corpus, for example, by performing automated term extraction on the corpus and counting the number of terms that overlap between the ontology and the corpus [11]. If the terms used in the ontology are not present in the corpus or vice versa, the ontology is penalised. Such an evaluation approach is not suitable for assessing the correctness, clarity, or usefulness/applicability of an ontology and is more fitting for measuring coverage of the ontology. The coverage has been also considered as one of the criterion in the criteria-based evaluation [69, 70] that will be discussed later.

3. **Evaluation by humans** - This approach uses a set of pre-defined criteria. Lozano-Tello and Gómez-Pérez [43] proposed Ontometric, a method to “quantify the suitability of these ontologies for the system” [pg.3], which is an example of a human-based evaluation approach. Ontometric is a multilevel framework of characteristics that allows the users to measure the suitability of existing ontologies considering the requirements of a system. The measurements are based on the five main dimensions of tools, language, content, methodology, and costs. However, this approach exhibits limited support for building an ontology from scratch and evaluating it.

4. **Application-based evaluation** - The application-based evaluation approach first uses the ontology in an application and then evaluates the results. This approach is very useful to assess the capabilities of the developed ontology to meet its objectives, e.g. decision support or knowledge management. Yet it does not validate the quality of the content and design of the ontology.

Yu et al. [70], in comparison, suggests three main categories for ontology evaluation as follows:

1. **The gold standard evaluation** - It can be seen that the first one, the “gold standard evaluation” is overlapping with the approach proposed by Brank et al. [10].
2. **Task-based evaluation** - The task-based evaluation, similar to the application-based evaluation proposed by Brank et al. [10], assesses the ontology according to its competency in achieving target tasks by measuring its performance within the context of the application. Yu et al. [70] suggest this approach to be conducted for each task separately, because evaluation results for different applications and tasks may not be comparable with each other.

3. **Criteria-based evaluation** - The criteria-based evaluation uses a set of proposed criteria for evaluating the ontology [30]. This approach has some similarity with the third, evaluation by humans, category specified by Brank et al. [10], but is more generic and flexible. The criteria-based evaluation can verify the content and design of the ontology and the application-based evaluation can assess the usability and applicability of the ontology in its application domain. The criteria-based evaluation requires the use of appropriate and effective attributes.

   In the literature, there are various criteria specified for evaluation of ontologies. For example, to measure the quality of ontology design, Wand and Weber [67] consider the *Clarity* and *Completeness* of ontology constructs and propose the criteria of Construct Overload, Construct Redundancy, Construct Excess, and Construct Deficit for examining them. They argue that these criteria measure the deficiencies which could undermine the usefulness of the ontology. This evaluation model has been later extended and modified by researchers to cater for additional requirements [56]. We have synthesised the criteria proposed by these authors and Table 1 presents the summary and comparison of commonly used evaluation criteria. Out of these, Yu at al. [69, 70] integrate most of the existing criteria and include in addition an attribute of *Coverage*.

**Table 2** Examples of various criteria to evaluate ontologies.

<table>
<thead>
<tr>
<th></th>
<th>Gruber [29]</th>
<th>Gomez-Perez [27]</th>
<th>Yu et al. [69, 70]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarity</td>
<td>-</td>
<td>Clarity</td>
<td></td>
</tr>
<tr>
<td>Coherence</td>
<td>Consistency</td>
<td>Consistency/Coherence</td>
<td></td>
</tr>
<tr>
<td>Extendibility</td>
<td>Expandability</td>
<td>Expandability/Extendibility</td>
<td></td>
</tr>
<tr>
<td>Minimal Encoding Bios</td>
<td>-</td>
<td>Minimal Encoding Bios</td>
<td></td>
</tr>
<tr>
<td>Minimal Ontological Commitments</td>
<td>-</td>
<td>Minimal Ontological Commitments</td>
<td></td>
</tr>
</tbody>
</table>
The above review demonstrates that there is no one approach which will perfectly fit all the objectives of ontology evaluation, hence it is often the case that a combination of the above or their variation is used. In the next section we propose and argue for a selection of approaches and criteria adopted for DO4MG evaluation.

### 5.2 Selection of Evaluation Approach for DO4MG

Having reviewed different ontology evaluation approaches, based on the DO4MG’s objectives and the limitations of each approach, such as the lack of gold standards as well as of an existing ontology for mass gathering, we have selected the following two approaches for DO4MG evaluation:

- Criteria-based evaluation; and
- Application-based evaluation.

Considering the evaluation criteria discussed in section 5.1.1, we selected eight criteria which match the objectives of DO4MG. These criteria include:

1) clarity; 2) consistency/coherence; 3) conciseness; 4) expandability/extendibility; 5) correctness; 6) completeness; 7) minimal ontological commitment; and 8) coverage.

The two criteria that we have not considered are sensitiveness and minimal encoding bias. The sensitiveness of the ontology [27] refers to how minor changes in the definitions can modify other well-defined properties that have already been guaranteed. Using this criterion requires a standard measuring tool that can determine the level of sensitivity made by changes. However, since we assume a regular review and update of the ontology performed in collaboration with human experts, any changes of definitions, either minor, or major will be performed accordingly. The relationships between the concepts would also capture the need for such modifications.
The criterion of minimal encoding bias suggests that the ontology representation should not be limited to a certain symbol-level encoding due to the convenience of implementation. DO4MG is implemented in Protégé-OWL which provides the adequate representation choices in order to meet our objectives.

The following section reports on the results of the criteria and application-based evaluation.

6. Results of Evaluation and Refinement of DO4MG

To evaluate the ontology, we reviewed each concept individually based on the selected criteria and then conducted a focus group with domain experts who looked at the proposed ontology and evaluated its content in particular, for clarity, completeness, consistency, correctness. As the result of this session the expandability of it was also demonstrated. The details of the focus group are presented in the following subsection.

6.1 Data Collection for Ontology Evaluation

There are different techniques for data collection in a group setting. Examples of two well-known methods are focus groups and Delphi method. Holsapple and Joshi [34] propose the Delphi approach as an effective collaborative technique for ontology design. Delphi is a systematic and iterative process to assist a group of experts to arrive at a consensus [33, 68]. This approach generally requires the use of a questionnaire, and responses are typically collected anonymously and without the need for a face-to-face meeting. Alternatively, a focus group technique can be used to collect rich qualitative data when access to the group can be arranged [60]. This technique is similar to a group interview and is coordinated by a facilitator who has strong personal skills. We selected the focus group technique since there was an opportunity for a face-to-face meeting with a group of highly qualified emergency management domain experts to assist with our data collection and ontology validation. This allowed us to avoid the need for a long and complex questionnaire about more than 200 ontology concepts. Moreover, one of the group members who had previous experience as a facilitator had the role of the moderator in our focus group.

The focus group to validate and refine the DO4MG involved 10 participants from different emergency management related organisations, including the World Association for Disaster and
Emergency Medicine, St John Ambulance Australia, Metropolitan Ambulance Victoria, and Flinders University, Adelaide, Australia. The selection of the participants was based on their research, domain knowledge, publications and professional experience in the area of mass gathering medical management in Australia. Our aim was to use their advice to validate the draft domain ontology and improve and refine it based on their feedback.

First the facilitator presented an overview of DO4MG ontology to the participants. Then he showed them every main concept of the ontology and its subclasses and collected feedback. Every ontology concept, its subclasses, and the relationships between them were discussed and suggested changes were recorded. This included deletion of some concepts and addition of new concepts.

In the next subsection we discuss examples of the domain expert feedback to refine the ontology as part of our criteria-based evaluation.

### 6.2 Criteria-based Evaluation of DO4MG

In this section, we describe how the data that was collected from focus group was used to conduct the evaluation according to the selected criteria.

**Clarity** – Gruber [29] states three requirements for clarity that include the following: i) the ontology terms should be defined formally without subjectivity; ii) the ontology needs to be documented with natural language, and iii) and the terms must convey ‘the intended meaning’ with regard to the requirements of social situations and computation rather than their context. In the context of formal definition of the ontology terms, since we have mainly extracted these terms from the domain-related publications, formal definitions are available for most of the terms. For example, the term ‘Crowd Catalyst’ is defined as factors that “contribute to or trigger a crowd from being one that is managed to one that needs to be controlled” [7:245]. We have documented these terms using natural language. With regard to the third requirement, as Yu et al. [70] suggest, that lack of measurement methods for clarity makes assessing clarity a difficult task. Having access to domain experts was important and provided an efficient way to verify the clarity of the ontology and refine it based on their feedback.
**Examples** - In the DO4GM ontology, we had ‘CrowdControl’ as a key concept that included several subclasses such as CrowdCatalyst and CrowdMood. During our focus group with domain experts, it was pointed out that the term did not communicate the intended meaning and it was replaced with ‘CrowdFeatures’ to improve clarity.

We also had a subclass of ‘Pollution’ under the ‘EnvironmentalFactors’ concept. Based on the feedback, this was replaced with ‘AirQuality’ with instances of ‘Dust’, ‘Pollen’, ‘Smoke’, etc.

**Consistency/coherence** - The concepts and elements of ontology should have a logical consistency and avoid contradictions or ambiguity. The ontology “should sanction inferences that are not consistent with the definitions” [29:3]. The consistency check produced several examples where changes were required based on experts’ advice.

**Examples** - Initially the ‘GatheringType’ was broken into two subclasses of ‘GeneralGathering’ such as political or cultural events and ‘HighRiskGathering’ like fireworks or motor racing; according to Emergency Management Australia, Manual 2, Safe and Healthy Mass Gatherings [22]. However, the inferences were considered inconsistent and contradictory with the defined concepts because in the event types such as a firework gathering the preparedness measures are increased and this reduces the risk. Thus a political gathering (i.e. a general gathering event) might have a higher risk than a firework gathering. Based on the feedback, the concepts of ‘general’ and ‘high risk’ gatherings were removed and all the gathering types were grouped under the ‘GatheringType’.

We had included the subclass ‘Insurance’ under the ‘MassGatheringPlan’. In the focus group, we found out that the insurance topic was carrying some ambiguity. There could be different types of insurance policies involved in planning a mass gathering event. These policies are considered to be, generally, the responsibility of the event organisers rather than emergency medical services. Since the DO4MG ontology targets the MEM, this term was removed to avoid the ambiguity and to maintain consistency.

**Conciseness** – The conciseness criterion means that an ontology should not include unnecessary concepts or redundancies [28, 69]. This aspect has been carefully considered during our ontology development and validation.
An example - the only redundant term used in the ontology was ‘Parking’. This term was a subclass of ‘Internal Structure’ (also subclass of ‘Event Venue’) and an instance of ‘Event Management’ (also a subclass of CrowdCatalyst belonging to the class ‘CrowdFeatures’). However, since the ‘Parking’ under ‘Event Management’ refers to the lack of adequate parking spaces for the crowd, it was changed into ‘LackOfParking’.

Expendability/extendibility – This criterion refers to the ability of ontology to extend further or to be applied to a specific application domain. DOEM has been built such that it provides for the reuse and extension of the different parts of the ontology.

An example - In [13], the proposed ontology for emergency transportation in mass gatherings is an extension to DO4MG. The left hand side of Fig. 4 shows the concept of ‘EmergencyTransportation’ in DO4MG, and the right graph shows the extension of DO4MG for ‘Transportation’.

![Fig. 4 Extending DO4MG for emergency transportation.](image)

Correctness – Correctness means that the ontology represents the correct modelling of the real-world concepts [69]. The correctness of DO4MG has been the main focus of our evaluation. The feedback provided by domain experts has greatly assisted in verifying this criterion.

Minimal ontological commitment – This criterion refers to allowing more flexibility and freedom in the specialization of the ontology by minimizing the claims about the modelled world [29]. Yu et al. [69] examines this attribute with regard to supporting multiple views for the same information and flexibility in classifying items. We have evaluated this feature by developing applications that target different activities of the domain.

Completeness – This criterion applies to completeness of the individual definitions of the ontology [63, 69]. As Yu et al. [69] suggest this attribute can be evaluated by using competency questions which include the queries and requirements that the ontology must be able to answer [31, 63]. Because of the size of the DO4MG, we present only two examples of the competency questions in Table 3.

Table 3 An example of the competency questions used to evaluate completeness.

<table>
<thead>
<tr>
<th>Competency questions</th>
<th>Concept</th>
<th>Relations</th>
</tr>
</thead>
<tbody>
<tr>
<td>What demographic factors of the crowd need to be considered?</td>
<td>Age</td>
<td>hasDemographicsOf</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ethnicity</td>
<td></td>
</tr>
</tbody>
</table>

Coverage is one of our selected criteria for ontology evaluation. To examine the DO4MG coverage, we employed a different approach. We conducted a simple experiment using a text mining software system, Leximancer [61, 62] and compared the results with those included in the DO4MG ontology. The next subsection describes this evaluation.

6.2.1 Coverage

Coverage can be defined as the completeness and coverage of terms and concepts to represent an information domain [69]. The criterion of coverage is more suitable for the data-driven evaluation where the ontology is compared to a corpus. The concepts of DO4MG were mainly extracted manually from current publications on the medical emergency management for mass gatherings or added by domain experts.
Leximancer is a computer-assisted text analysis application that uses a machine-learning technique for conceptual analysis and relational/semantic analysis [62]. Leximancer enables discovering new information from text-based resources, finding patterns, and generating context maps and statistical outputs that provides an insight into a large amount of data corpus and a set of documents [61, 62]. We used Leximancer to extract the main concepts and terms from the two main sets of documents that cover the domain of mass gathering. These two resources are the ‘Mass Gathering Compendium’ collected by Arbon and the Emergency Management Australia Manual ‘Safe and Healthy Mass Gatherings’, 1999, Part III, Vol. 2, Manual 2. Fig. 6 depicts the visualization and analysis output which has been automatically created.

![Concept Map Produced by Leximancer for the mass gathering corpus.](image)

**Fig. 6** Concept Map Produced by Leximancer for the mass gathering corpus.

The themes are large circles that represent the main groupings of concepts (clusters) within the above-mentioned documents. We did the test using the default Leximancer parameters without any configuration and that resulted in extracting some of the noise words and concepts such as Nornberge, data, information, area, public, system, on-site, people, support, staff, etc. The figure also shows some overlapping/repeating clusters like Medical, medical and Disaster Medicine, and Crowd and Spectators.
Comparing the figure with our ontology, all the extracted concepts are included in the DO4MG but with a different hierarchical structure, except for the noise words which are either incomplete, too general or out of context. The DO4MG has five main concepts (first level) that include EventVenue, GatheringType, CrowdFeatures, MassGatheringPlan (including MedicalResponse, PatientPresentation and Injury subclasses), and EnvironmentalFactors (including WaterCourse and Weather subclasses). The extracted Crowd and Spectator correspond to CrowdFeatures, Event to Gathering Type, and Medical, Disaster Medicine and Patients to the MedicalResponse subclass under the MassGatheringPlan. The Alcohol, Water, and Access can be mapped to the subclasses of AlcoholSale, AlcoholUse, Watercourse, AccessEgress under the EnvironmentalFactors class. The Mass cluster which is too broad and incomplete can be matched to the MassGatheringPlan class.

The other extracted terms inside the clusters (ignoring the noise words) are also included in the ontology like food, first-aid, police, ambulance, time yet with a relatively different structure. This experiment is useful to match and validate the concepts of an ontology against a related corpus and identify the overlapping or absent concepts. However, it also shows the limitations of Leximancer to discover the accurate hierarchy and relationships of these concepts in each cluster. Based on our findings we suggest that this stage requires significant refinement by domain experts.

6.3 Application-specific Evaluation of DO4MG

To perform application-specific evaluation of DO4MG, we have developed generic architecture for intelligent decision support and knowledge management in mass gatherings that integrates DO4MG (shown in Fig. 7). The aim was to verify our ontology development and evaluation approach and also validate the usability of the DO4MG ontology for resolving problems which require the clear structure of the problem domain. These problems include terminological differences and semantic conflicts that may arise from applying the different terminology used by medical emergency services to express the same concept.

Organizing a successful mass gathering is complex and includes several stages and a variety of tasks which require participation of different agencies and services. In general mass gathering activities and tasks can be grouped under three phases of pre-event, during-the-event and post-event
phases. The pre-event phase typically involves training, workload estimation and planning the event. The during-the-event phase is the operational stage where medical response and treatment are provided and there is a need for real-time DSS. The final stage of mass gathering usually entails recording data, debriefing, evaluation of the event. Using a common ontology in all the stages of mass gatherings provides consistency and effectiveness in all the activities, and facilitates data entry, management, filtering and integration, avoids discrepancies [17, 64, 65].

*Fig. 7* A comprehensive architecture for mass gatherings integrating ontologies (adapted from [17]).

The case-based reasoning application can be used during the Pre-Event stage of mass gathering (see Fig. 7) for predicting workload and assisting trainees and during the Post-Event stage for improving data collection, integration and storage.

Case-based reasoning (CBR) [39] is a method used in decision making that provides solutions to problems using similar past cases, CBR process includes four main stages Retrieve, Revise, Reuse and Retain [6, 39]. The CBR tool in our context can assist medical emergency personnel with resolving inconsistencies in the domain vocabulary, for example, when searching for similar events in event resource planning context.
The data repository used in this application includes 201 records of different mass gathering events which occurred over a period of 12 months. The data was collected by St John Ambulance Australia personnel (St.John, 2010) and has previously been used for building a predictive model [2] and for identifying the important attributes when predicting injury/illness rates (i.e. Patient Presentation Rate) [59]. The CBR prototype was implemented using the jCOLIBRI2 framework\(^1\) in Java. jCOLIBRI2 is an open source tool that provides supports for development of different CBR applications including ontology-based CBR systems by using the Onto-Bridge\(^2\) libraries [55]. The application enables the user to enter the details of a future event (Querying Stage) such as the event type and location, the number of attendees, environmental values like temperature and humidity, etc (as shown in Fig.8).

![Fig. 8. The querying stage of CBR using DO4MG ontology.](image)

As Fig.8 shows some of the attributes like Gathering Type allows the user to directly select the input from the ontology. After defining a future event, the user will enter two important values which are Patient Presentation Rate and Transportation to Hospital Rate. These two attributes are used to estimate the workload of medical emergency services. For example, a larger number of attendees or higher temperature and humidity may result in a higher rate of patients and transportation to hospital.

\(^{1}\) http://gaia.fdi.ucm.es/projects/jcolibri/

The CBR application enables the user to compare their estimated Patient Presentation Rate (PPR) to the similar events in the past. This comparison provides users with better understanding of such events and assists them with decision making about the expected workload and required resources. The usability of DO4MG is mainly assessed during the Retrieve stage. During this stage, a ‘problem case’ entered by the user is compared to the case base (i.e. stored past events) by using the selected similarity functions, and the most similar cases are retrieved. At this stage we demonstrate how the DO4MG ontology deals with terminology conflicts.

Reviewing the literature for mass gathering shows that different researchers or agencies use different attributes/terms to record the same concept [35, 38, 46, 49, 71]. For example, there are terms that represent the same or a very similar concept to ‘motor racing’. These terms include ‘auto racing’ [49], ‘motor sports’ [38] ‘motor race’ [35], ‘automobile races’ [46].

When the data from several events need to be integrated or used collectively, the inconsistencies and discrepancies between different data sets can complicate data integration and management (i.e. storage, retrieve, query, etc). In the context of case-based reasoning, this can result in retrieval of incorrect cases. If the recorded data and the query entered by the user use different terms/attributes but have the same meaning (i.e. they are synonyms), they will not be matched correctly. This will reduce the accuracy and effectiveness of the case-based reasoning and decision support. To address this issue, ontologies provide an elegant solution by supporting synonyms and providing richer queries that could be performed using concept based similarity functions [55]. The DO4MG is developed in Protege which provides several options to create synonyms. We have used annotations (i.e. labels) to create different synonyms for a concept. For ‘motor racing’ we have included all the above-mentioned synonyms as labels in DO4MG ontology.

The data set used for case-based reasoning can store different terms/words to express the same concept which can lead to inaccurate retrieval of similar cases. In our prototype, the database stores the term ‘MotorSports’ to express the concept MotorRacing. Through using the ontology and defined synonyms, the developed application is able to match the terms specified in the query correctly to the stored cases such that all the synonyms will be considered at the Retrieve stage. Thus the accuracy of results and generated solutions was improved. The retrieved results will include similar cases that use
the synonyms to represent the same concept even though they are different words (compared to the query terms). In our example, the ‘MotorSports’ attribute used in the data set has been matched correctly to the MotorRacing concept in the DO4MG ontology because it was defined as its synonym. This will retrieve the correct cases from the data set and resolves the issue of inconsistency and terminology conflicts.

The developed prototype demonstrates feasibility of applying DO4MG to the decision support system component to deal with complexity and inconsistency of decisions in medical emergency management, thus satisfying the requirements of the application-based evaluation of our ontology.

7. Conclusion

Medical emergency decision making is a challenging and difficult task, particularly in the context of mass gathering events. Success of running a major event highly depends on the effective provision of medical emergency services that are often offered by different agencies. Hence, poor communication, coordination and knowledge sharing between these agencies and teams can result in delays and duplication of efforts. Another related issue is that emergency decisions are usually made based on individual experience and domain knowledge of the relevant managerial personnel, which could be beneficial to collect, consolidate, store and share as part of knowledge management for future decision support. The complexity of decision making with regard to MEM in mass gatherings arises when different agencies and emergency services use different terminology to refer to the same concepts. Such inconsistencies are usually dealt with naturally by humans; however, they create problems when computerised decision support systems are introduced. Ontologies as a unified representation of a domain of interest provide a common basis to deal with such inconsistencies, and are often recommended as part of knowledge management for intelligent decision support systems [59, 63, 64]. The use of such generic and unified knowledge structure can also improve the coordination and communication between different emergency teams.

Despite an increasing number of ontologies, there is still no generic agreement on how they can be constructed in the most efficient way to facilitate better knowledge management and decision-making. When using ontologies for decision support in the risk-intensive context, such as emergency
management in mass gatherings, it is imperative to determine whether they provide a valid representation of the application domain. In this paper, we describe the process of construction and evaluation of DO4MG (Domain Ontology for Mass Gatherings) based on a systematic review of the literature, analysis and synthesis of existing methods of ontology construction. The resulting ontology was evaluated based on two main approaches; criteria-based and application-specific evaluation. As part of the validation, we illustrate the application of the DO4MG for implementation of a cased-based reasoning decision support for medical emergency management in mass gatherings. Such implementation demonstrates the potential benefits of using ontologies in resolving terminology inconsistencies and conflicts, and its usefulness to increase the efficiency of communication between emergency medical personnel in mass gatherings.

In future we intend to use the constructed ontology and developed evaluation methodology in other decision support approaches and study the outcomes involving field studies and domain experts. We aim to offer this ontology to others as a validated robust description of this important problem domain. We also intend to target mobile applications and implement ontology-based decisions support for MEM in mass gatherings that can operate on mobile devices as extension to the proposed approach [17, 18]. The potential of transferability of the approaches used in this research onto other areas of ontology creation would be worthwhile empirical testing.

Acknowledgements

This research is funded by Australian Research Council (ARC) Linkage grant LP0774834. The team include CIs: Prof Frada Burstein, A/Prof Shonali Krishnaswami, Prof Paul Arbon, PI: Prof Arkady Zaslavsly; Research Fellow Dr Pari Delir Haghighi. We acknowledge support of Jamie Ranse in conducting the focus group as part of data collection for this paper.

References


and Education. (Eds.) A. Caplinskas, O. Vasilecas, W. Woitkowsky, S. Wrycza et. Al. Kluwer

Healthy Mass Gatherings, Emergency Management Australia, Canberra: Commonwealth of
Australia, Accessed 11 September 2011,

in emergency situations. CSCWD 2011, pp 669-676.

ontology using methontology and the ontology design environment. IEEE Intelligent


[26] M. Gaynor, M., Seltzer, S. Moulton, J. Freedman, A Dynamic, Data-Driven, Decision
Support System for Emergency Medical Services. International Conference on Computational

[27] A. Gómez-Pérez. Towards a framework to verify knowledge sharing technology. Expert

391–409.


[31] M. Gruninger and J. Lee, Ontology Applications and Design, Special issue in


[41] A. Lewis RIMSAT DSS Project: Integrating Model-Based and Case- Based reasoning, DSSResources.COM, 2004, Available: 
http://dssresources.com/papers/features/lewis/lewis04052004.html


[64] F. Sujanto, F. Burstein, A. S. Ceglowski, L. Churilov, Application of domain ontology for
decision support in medical emergency coordination, Proceedings of the 14th Americas

comprehensive collaborative emergency management, in Collaborative Decision Making:
Perspectives and Challenges, P. Zarate, J. Belaud, G. Camilleri and F. Ravat (eds), IOS Press,
Amsterdam Netherlands, 2008, pp. 127-138


[67] Y. Wand, R. Weber, Research commentary: Information systems and conceptual modeling a
research agenda. Information Systems Research, 13 (2002), 363-376


[69] J. Yu, J. A. Thom, and A. Tam. Evaluating ontology criteria for requirements in a geographic
classical travel domain. In Proc. of Intl. Conf. on Ontologies, DataBases and Applications of
Semantics, 2005.

[70] J. Yu, J. A. Thom, A. Tam. Ontology evaluation using wikipedia categories for browsing,
Proceedings of the sixteenth ACM conference on Conference on information and knowledge

[71] K. Zeitz, D. Schneider, D. Jarrett, C. Zeitz, Mass gathering events: Retrospective analysis of
patient presentations over seven years at an agricultural and horticultural show. Prehosp and


[74] A. Zerger, D. I. Smith, Impediments to using GIS for real-time disaster decision support.