BalOnSe: Ballet Ontology for Annotating and Searching Video performances

Katerina El Raheb¹, Nicolas Papapetrou², Vivi Katifori³, Yannis Ioannidis⁴
³,⁴ Athena Research and Innovation Center
¹,²,⁴ Department of informatics-University of Athens
Athens, Greece
{kelraheb, papapetrou, vivi, yannis}@di.uoa.gr

ABSTRACT
In this paper we present BalOnSe (named after the ballet step balance), an ontology-based web interface that allows the user to annotate classical ballet videos, with a hierarchical domain specific vocabulary and provides an archival system for videos of dance. The interface integrates a hierarchical vocabulary based on classical ballet syllabus terminology (Ballet.owl) implemented as an OWL-2 ontology. BalOnSe supports the search and browsing of the multimedia content using metadata (title, dancer featured, etc.), and also implements the functionality of "searching by movement concepts", i.e., filtering the videos that are associated with particular required terms of the vocabulary, based on previous submitted annotations. In the paper, we present the ballet.owl ontology, and its structure, explaining the conceptual modeling decisions. We highlight the main functionality of the system and finally, we present how the manual ontology guided annotation allows the user to search the content through the vocabularies and also view statistics in the form of tag clouds.

Author Keywords
Video Annotation; Ontology; Movement Representation; Semantics; Ballet, Interface.

ACM Classification Keywords
H.5.5. Information interfaces and presentation (e.g., HCI): Sound and Music Computing—Systems. J.5. Arts and Humanities: Performing arts (e.g. dance, music).

INTRODUCTION
It is no question that dance videos of every kind can be found in large amounts in Internet multimedia and social media channels such as youtube, vimeo, facebook, etc. In parallel, several efforts have been made to organize dance videos as rich multimedia content, such as eClap [4], where the dance videos can be browsed and searched using metadata concerning the performer, the dance company, the title etc. BalOnSe is a web application created with the goal of helping individuals or teams keep better track of the content of a set of videos. Ballet, like any sport, martial art and dance has a very specific vocabulary for its movements and techniques. For the purposes of our project, we used as an example set a number of selected videos, of well known ballet variations, i.e., solo pieces. While a variety of sophisticated technologies exist to analyse and capture whole body movement, the video of dance extracts still remains the most direct medium to communicate, disseminate, and reflect on a dance piece for educational, analytical and research purposes, and also serve as the basis further automated segmentation and processing. Different groups have explored video annotation interfaces, which can facilitate the communication between different stakeholders, in the past decade, with different degrees of automation and semantic analysis [14]. In our experiment, we focused on what linguistic terms might be useful for the user considering terminology within the context of the ballet genre, but also through a more generic perspective. We are aiming to facilitate a usable interface that eventually can serve a variety of users and dance amateurs who may be less familiar with the ballet genre.

MOVEMENT ANNOTATION TOOLS
Describing movement and segmenting recognized movement entities that are meaningful in various contexts such as gestural, non-verbal communication, sign language, sports activities, dance, etc, currently remains an important and challenging research area. Two main open issues remain: the first is the automated segmentation of video of movement sequences, based on pattern recognition or other state-of-the art methodologies, and the second problem is the development of the semantic models which represent the domain specific kinetic vocabularies. For this work, which we present in this paper, we considered previous advancement in the fields. One of the significant works in the field is the Anvil interface and the corresponding schema of manual annotation for conversational gestures, which eve
ntually supports the recreation of 2D animation based on time and special descriptions of the gestures on videos [9]. As Bertini et. al explain [1], a typical way to perform video annotation requires to classify video elements (e.g. events and objects) according to some pre-defined ontology of the video content domain, while in the same paper the present pictorially enriched ontologies based both on linguistic and visual concepts and the implementation of solutions for video annotation and retrieval based on these extended ontologies. Ramadoss and Rakummar [12] have presented the system architecture of a manual annotation tool, a semi-automatic authoring tool and a search engine for the choreographers, dancers and students of pop Indian dance to demonstrate how the dance media can be semantically annotated and how this information can be used for the retrieval of dance media objects [12]. In their paper, which they present the semantic models used, it is clarified that the video clips of Indian dance, have been an example of dance videos that allowed the authors to consider the content as narratives. In this case, movements could be easily mapped with particular segments of the song, and the mood described by movement and lyrics [11]. Singh et. al [14], presented the Choreographic Notebook, which is a multimodal annotation tool, supporting the use of text, digital ink, and to be used during the production process of contemporary dance. A similar multimodal annotation tool and approach have been presented by Cabral et. al [2], in the Creation-Tool. The two last examples are applicable in cases where the movement has no narrative or symbolic meaning, and hence the elements of the annotations can be abstract shapes on the video screenshots. Also the purpose of use is different, since it aims mainly at the collaboration and sketching during the choreography process.

DANCE REPRESENTATION MODELS
Though the aforementioned examples of related works share some commonalities, in terms that they all present tools to annotate movement, it is clear that the underlying schemata of the semantic descriptions completely differ due to the different dance contexts and the dance genre they support. Semantic representation models for dance practices is another open research issue, and most of the works are based on the following: a) Universal systems of analyzing and notating dance such as Laban Movement Analysis (LMA), Laban Efforts, Labanotation, Benesh Notation, Eshkol Wachman, etc. In their work ElRaheb [5] and Ioannidis, present a Labanotation based ontology for describing movement, while Saad, Shatina et. al [13] propose a Benesh based ontology for representing movement. b) Ad-hoc schemata, which serve the particular content of the application, the dance genre, the purpose of the developed tool and the users, group to whom the interface is targeted. Ontologies as conceptual models can either be a) upper ontologies, 2) domain ontologies, and 3) application ontologies [2]. In our approach we developed a domain ontology to represent the terminology of ballet syllabus, combined with a Generic Movement Concepts [6]. More details on the ontology is provided in the next section.

CLASSICAL BALLET SYLLABUS
Classical ballet is one of the most widespread genres of dance which originates back in the 16th Century in Italy and developed sequentially in France, and Russia in the later centuries to become one of the standard techniques in curricula of most dance schools and academies worldwide. Though in many ballet plays, pantomime gestures exist, ballet in general is a genre of dance with no symbolic meaning of movement and gesture. On the other hand, the segmentation of a dance piece into meaningful entities can be largely based on the clear and well-defined syllabus. It is this aspect of ballet dance we propose to use as an annotation schema for these types of videos. Since, also the education of any ballet dancer relies strongly on the knowledge of the syllabus and the corresponding terminology, we believe that such a tool as BalOnSe could support the learning process. The hierarchy of the movements is browsable through the application, and provides an easily accessible vocabulary of the syllabus, which associates the terms with examples through the videos, by using the “search by movement functionality”. Although several schools of technique occur, such as Cecchetti\textsuperscript{I}, Vaganova\textsuperscript{I}, the Royal Academy of Dance\textsuperscript{II}, etc., a basic standardized syllabus of movements is common (with some alterations) along with the related terminology. Many terms of the syllabus have become a standard for the communication of dancer, even in other dance genres such as modern dance techniques (Cunningham, Limon), contemporary dance. For example, the five positions of the feet used in ballet, can be considered common knowledge across dancers of almost any dance genre, which is practiced in an institutional arrangement, or independent dance schools. In contrary to other systems of describing, analyzing and notating movement, such as Labanotation [6], the ballet syllabus and terminology consists of a common language among ballet dancers, students and educators, worldwide.

THE BALLET.OWL ONTOLOGY
The BalOnSe application integrates ballet.owl, ontology in OWL-2, which was developed for this purpose. The ontology consists of 151 classes, (512 axioms) which represent a hierarchical taxonomy of ballet syllabus vocabulary. Following the distinction, which was proposed by ElRaheb and Ioannidis [6], the top of the taxonomy of the movement terms is distinguished between two main classes:

a) Generic Movement Concepts: which refers to the common, everyday language of non-experts about movement including terms such as run, walk, turn, etc. (hasSubClass Generic Actions)

b) Specific Movement Vocabularies (hasSubClass Ballet Vocabulary): This refers to any domain specific terminologies coming from particular dance genre practices or techniques. Ballet Vocabulary, which is the terminology for the ballet technique syllabi, is one of the
In our investigation, we included twelve Generic Actions, which can summarize dance or stage movement activities. The Generic Actions included in the applications are the following in alphabetical order: Arm Gesture, Balance, Bend, Extend, Fall, Jump, Leg Gesture, Position, Run, Stillness, Turn, Walk. These Generic Actions, are used to categorize the different types of movements that exist in ballet syllabi, as shown in Figure 1. For example: Assemble (and all of its subclasses), Brise, Jete, Tour en l’air etc. are subclasses of the Ballet Vocabulary since they are part of the syllabus, but also subclasses of the Generic Action Jump. The axioms which are expressed about TourEnL’Air which is a type of turn done in the air while jumping, is the following:

- TourEnL’Air SubClassOf Jump
- TourEnL’Air SubClassOf Turn
- Jump SubClassOf GenericAction
- Turn SubClassOfGenericAction

To this point note that the application is developed for the annotation of ballet performances, however, the Generic Movement Concepts should cover terms that can describe actions, beyond any techniques, or dance specific knowledge. In the interface, the Generic Actions aim at helping the user choose from a very specific list of movements that are easily understood also by any non-dance expert. For this reason, we limited the Generic Actions only to a very short list, while the Specific Vocabulary Movements includes more than 100 classes. Short definitions are given for the terms, in the form of help comments, and the ontology can be also be browsed. The Generic Actions list, which is used in this version, is a result of a thorough investigation of possible basic movement categories. One of these candidate categories are the “seven movements” of ballet, as historically have been introduced in “Lettres sur la danse et les ballet” in 1760 and are theories which are valid in ballet practice until now: 1) plier- to bend, 2) sauter- to jump/leap3) tourner- to turn 4) etendre- to stretch, 5) relever- to rise up, 6) elancer- to dart, 7) glisser-to glide. As Guest clarifies: “Viewing the seven movements of dancing through Laban Movement Analysis, five of the categories identify forms or structures, and two identify effort qualities. The five movements that address forms are basic to human movement and appear in most dance styles. These are plier, etendre, relever, sauter, and tourner. This sliding technique is the impetus for the arc-like leg gestures in terre a terre, adagio, allegro, and grand allegro movements of ballet. The gliding movement quality manifests in adagio and in the soaring leaps of grand allegro. The darting quality is explicit in allegro and grand allegro movement. Gliding and darting qualities symbolize the dynamic image of ballet. In summary, plier, etendre, relever, sauter, and tourner identify forms and are common to basic human movement and most dance styles; whereas, glisser and elancer are salient effort qualities specific to ballet style”.[8] This explains why the list of the ballet seven categories of movement, did not seem appropriate to be used as is, since some terms (glide and dart) seem to be genre specific qualities of movement, rather than common activities easily understood by a variety of users. Besides the very high level concepts of Laban Movement Analysis (LMA) for actions, we considered the basic alphabet of the Language of Dance [10] by A.H. Guest. The alphabet includes basic actions, which are derived from LMA and are used in movement practice for both adults and children. The LOD consists of the following actions which are organized also in categories: (initial statements) 1) Action (any action), 2) Stillness, (anatomical possibilities) 3) Flexion, 4) Extension, 5) Rotation, (spatial aspects) 6) Travelling, 7) Direction, (supporting) 8) Support, 9) Spring, (center of gravity) 10) Balance 11) Falling. Although many similar lists of actions can occur in other systems, it is obvious that there is a basic core of actions, which exist in the different high level lists of actions, in both theories and practices of dance and is this core that we adopted in our ontology. In the following section, we discuss the actions we propose, commenting commonalities and differences with some of the aforementioned lists of actions.

1. Arm Gesture: The action of moving the arm(s) in any way. Actually both Arm Gesture and Leg Gesture are under the class Gesture.
2. Leg Gesture: The action of doing any movement with the leg, while the leg is free of weight and not supporting the body. Using the Laban definition for Gesture, we define as such, any movement that occurs without bearing or supporting the weight [7]. Both Leg Gesture and Arm Gesture are subcategories of Gesture.
3. Turn: The action of (continuously) changing the direction of the body. The action of turning (turner-to turn) is also in the core of different lists
4. Bend: The action of bending any part of the body, such as bending arms, knees, curving the torso, backbending etc.
5. Extend: The action of extending any part of the body, Bend/Extend, or Flexion/Extension or Etendre/Plier, are two actions in the core of any movement actions.
6. Jump: The action of elevating the whole body from the ground. Also seen as “spring”, is one of the movement actions in the core of all basic lists (sauter-to jump). It is usually analysed in three stages: the preparation (with bending the knees), the elevation, and the landing (again the knees bend). There are five main categories of jumps: 1) From one to the same foot, 2) From one to the other foot, 3) From both to one foot 4) From one foot to both 5) From one foot to the other. This definition is based on the practice an analysis of jumps in dance, as also expressed in Labanotation the main notation system for analyzing and notating movement [7], and is also used in other movement ontologies [4].
7. **Balance**: The action of balancing in one or more supporting body parts. E.g., Stand one foot, Handstand. These terms could be substituted by the term “Support”, to be more consistent with the Laban/Labanotation terminology; however, we have realized that this word seems to be clearer for the non-expert user.

8. **Fall**: The action of dropping-giving one body part or the whole body into gravity. Though this action rarely occurs in classical ballet, we included the term, as it is one of the fundamentals actions, as forming a unit with Balance in LOD [10] and many dance theories and practices of modern and contemporary dance.

9. **Walk**: The action of changing the support from one foot to the other while progressing in space. Walking can be in any direction.

10. **Run**: The action of running. Continuous changes of support from one foot to the other, with complete transfer of weight, while between these changes there is a moment where the whole body is off the ground.

11. **Position**: The action of giving into the whole body or a particular body part a particular shape. The ballet syllabus is rich in positions: the positions of the feet, of the arms of the body in relation to space (croise, efface, en face), etc.

12. **Stillness**: The action of pausing, remaining still, or holding.
ANNOTATION AND INTERFACE FUNCTIONALITY

In this section, we briefly present the functionality of the application and the interface, which we developed taking into account the fundamental principles of usability, and the recent trends in web design. The main characteristic of our system include the following features:

- Simple web-based interface for users with varying degrees of technological expertise.
- An archival system of videos for both the metadata and the annotations
- Context-related video navigational tools with semantic reasoning and search functionality
- Domain specific organized vocabulary

When presented with the application the first thing a user is going to observe is the navigation bar. The navigation bar consists of five options: Home, Show latest video, Vocabulary, Advanced Search, Search by Movement:

- The Home page is the first page a user is going to see. Its purpose is to give the user a general feel of the application and to show him/her the videos.
- The Show latest video page is just a quick way for the user to redirect to the latest video. This option leads to the main annotation page of the videos (Figure 4). In this screen the vocabulary is actually used as a menu, this is why in the design we tried to achieve a good balance between the rich semantic hierarchy of movements vs. what is usable (many choices at one level vs. less choices in more depth levels).
- The Vocabulary page displays the Ontology Tree, also by clicking on the tree nodes after the 0th depth information about the selected class will appear. In this screen the user can browse through the different classes and see definitions and comments about each one of the terms.
- The advanced search page contains all the ways a user can search for a video by using information regarding the video metadata, such as title, featured dancer, etc. More specific about the different attributes a video can contain are given in the Database Structure segment.
- Finally the Search by movement.

SYSTEM ARCHITECTURE

The application provides an archival system for the videos while both the metadata of the videos and the annotations of the users are archived in a relational database schema.

In this section we briefly present the database schema, which implements a basic entity relationship model with three main entities: (User) Account, Video, and Annotation.

In more detail each table has the following attributes:
Account holds the data that have to do with a specific user account. The data that the table holds are: The name, which is the username, the account password and finally the email, though this attribute is not currently being used.

Video holds all the data regarding the application’s videos, plus all corresponding metadata, which are following:

- **Title** of the video
- **Genre**, (in the current version the only possible value is Ballet, but we consider adding more dance genre videos and vocabularies in future versions of the application)
- **Dancer** contains all the dancers’ names featured in the video.
- **Dance-work** field might contain info like the Name of the play, the scene that is featured and the act of the play. For example “Don-Quixote, Kitri variation, Act 1”.
- **Dance_company** is the related dance company e.g., Bolshoi Ballet.

Annotation is the table which stores the users annotations. It has the following attributes:

- The **name of the annotation**. A record is saved for each tag that is chosen by the user for a specific segment in a video. If the annotation is chosen from the vocabulary then it’s a reference to an ontology class name, otherwise it’s the custom tag that the user inserted.
- **Comment** is a user defined characterization of the annotation the user just created.
- **PostTime** is the time the user made the annotation
- **OntologyInfo** is the ontology URI to the referenced ontology class name or null otherwise. This attribute provides the opportunity to have log history of tags added by the same users in different times.
- **StartTime** is the time the annotation starts being observed in the video.
- **StopTime** is the time the annotation stops being observed in the video. StartTime and StopTime are expressed in second of the video time.

The corresponding tables of the database schema are connected with one another in the following ways:

- The Account table has a 1-N relationship with the Video table and with the Annotation table.
- The Video table has a N-1 relationship with the Account table and a 1-N with the Annotation table.
- The Annotation table has a N-1 relationship with the Account table and with the Video table.

Therefore for each video all the annotations provided by different users are stored, as well as different annotations of the same users provided in different times. The database stores the multiple tags of users, and stores the information of which user have provided these annotations. In the current version of the application, where videos and users are limited, this information is not exploited. Since, however, the objective is to provide a tool for groups of people to keep track of video archives, for a variety of reasons e.g., exchange, research, education, the systems foresees the need of indicating what annotations are provided by whom. In the current version, the annotations of users are stored and each time the video appears, a tag cloud shows all previous annotations.

As shown in Figure 6, the tag cloud under each video, includes both tags that have been selected from the defined ontology, but also free text “this is a very good example for Cabriole”. The tag cloud shows all the terms that have been used to annotate the different segments of the video, while the size of the fonts represents the frequency each terms occurs in this video. Note that one term can be used several times by the same or different users to annotate different segments of the video. This means that if a user annotates three different segments by picking the term “cabriole” (which is a specific jump with beating the extended legs in the air, in front of the body), then all of the three recordings count in the tag-cloud calculations. Eventually, the tag cloud shows the dominance of some terms over others. This might mean that the specific performance contains many [“cabrioles”](if this is the dominant term) or that most people have observed or noticed or decided to annotate this for their
own reason. In our work, so far, we are not in the position to answer which of the above should be the case, as this requires extended users experiments. The application, however, is appropriate for supporting this types of experiments, which are eventually interesting questions related to how people use specific terminologies, observe dance, and recognize specific parts of standard syllabi according to their backgrounds.

Finally, the application has been developed using mainly open-source technologies aiming at a sustainable, reusable and extensible system. Some of the technologies used for the are for the interface is JavaEE, Spring MVC framework, Apache Maven, Hibernate ORM, H2 database, AngularJS, Twitter Bootstrap CSS. For the integration of the ontology open-source Jena API and Pellet Java based OWL 2 have been used.

![Screenshot of tag cloud based on previous annotation](image)

**Figure 6**-Screenshot of tag cloud based on previous annotation

**FUTURE WORK**

At the moment we have done an initial round of expert evaluation of the tool and the results are taken into account for the next version, which will be ready to be evaluated by users. An issue is the extensive use of expert ballet syllabus terminology in the vocabulary and the effect that this may have on users not familiar with it. On the other hand, it is important to identify the main user groups for this tool, as ballet dance enthusiasts, even non-expert ones, are bound to be familiar with ballet terminology. In our future work we plan to further evaluate our application in terms of usability, and user experience, and do further experiments to get more feedback from a variety of user groups about the way the ontology and terms are used in action. There are some indications from the internal evaluation that the fact that the ontology includes if –not all– a big percentage of the ballet syllabus terminology, this might cause frustration to less experienced users with this terminology. The application, as well as the ontology, which can be used independently can be used in a variety of contexts, including exchange between research or studying groups, educational purposes, and also as a way to gather users observations on videos online. Nevertheless, although the current version is built upon the usability principles of error prevention, help, and error recovery, in terms of annotations, we did not implement any algorithms to validate and check the semantic correctness of the annotation.

**CONCLUSIONS**

In this paper we presented a web-based application, which allows the user to annotate dance videos, using both free text tags, and terms from a predefined ontology of ballet syllabus terminology, while providing video archiving functionalities. The interface is designed and developed using recent principles of usability (help, error prevention, feedback, etc) while the generic movement vocabulary supports the users who are less familiar with the vocabulary, or just desiring a more abstract description of movement. In this work, we have also developed a domain specific vocabulary for representing hierarchies of ballet syllabi, and shown some examples of uses of the application. Our main contribution is 1) the development of a web-based user interface and application which is completely dedicated to the creation of small content oriented archives of dance videos, 2) the introduction of the functionality of searching videos by dance terminology keywords which are provided by the users, 3) the implementation and presentation of a ballet syllabus ontology, which can be used in other applications, be extended or integrated with other similar ontologies. The application can potentially work with other similar ontologies and sets of video, e.g., for other dance genres.

**ACKNOWLEDGMENTS**

This paper includes part of the research which is conducted in the framework of the WhoLoDancE H2020 EU funded project. In addition, we thank all the volunteers who have helped in the internal evaluation of the application.

**REFERENCES**

1. Bertini, Marco, Alberto Del Bimbo, and Carlo Torniai. "Enhanced ontologies for video annotation and retrieval." Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval. ACM, 2005.

2. Cabral, Diogo, et al. "Evaluation of a multimodal video annotator for contemporary dance." Proceedings of the International Working Conference on Advanced Visual Interfaces. ACM, 2012.

3. Chandrasekaran, Balakrishnan, John R. Josephson, and V. Richard Benjamins. "What are ontologies, and why do we need them?." IEEE Intelligent systems 1 (1999): 20-26.

4. elibrary for performing arts-http://www.eclap.eu

5. El Raheb, Katerina, and Yannis Ioannidis. "A labanotation based ontology for representing dance movement." Gesture and Sign Language in Human-Computer Interaction and Embodied Communication. Springer Berlin Heidelberg, 2011. 106-117.

6. El Raheb, Katerina, and Yannis Ioannidis. "From Dance Notation to Conceptual Models: A Multilayer
7. Guest, Ann Hutchinson. Labanotation: the system of analyzing and recording movement. Routledge, 2014.

8. Guest, Ann Hutchinson. "Seven Basic Movements in Dancing." Dance Research: The Journal of the Society for Dance Research 13.1 (1995): 7-20.

9. Kipp, Michael, Michael Neff, and Irene Albrecht. "An annotation scheme for conversational gestures: how to economically capture timing and form." Language Resources and Evaluation 41.3-4 (2007): 325-339.

10. Language of Dance Center -http://www.lodc.org

11. Ramadoss, Balakrishnan, and Kannan Rajkumar. "Semantic Modeling and Retrieval of Dance Video Annotations." INFOCOMP Journal of Computer Science 6.1 (2007): 9-17.

12. Ramadoss, Balakrishnan, and Kannan Rajkumar. "Semi-automated annotation and retrieval of dance media objects." Cybernetics and Systems: An International Journal 38.4 (2007): 349-379.

13. Saad, Shatina, et al. "An ontology for video human movement representation based on benesh notation." Multimedia Computing and Systems (ICMCS), 2012 International Conference on. IEEE, 2012.

14. Singh, Vikash, et al. "The choreographer's notebook: a video annotation system for dancers and choreographers." Proceedings of the 8th ACM conference on Creativity and cognition. ACM, 2011.

\[\text{i} \quad \text{http://www.cecchetti.org/}\]

\[\text{ii} \quad \text{http://www.vaganova.ru/}\]

\[\text{iii} \quad \text{http://www.rad.org.uk/}\]