ABSTRACT

The Semantic Web and Linked Data gained traction in the last years. However, the majority of information still is contained in unstructured documents. This can also not be expected to change, since text, images and videos are the natural way how humans interact with information. Semantic structuring on the other hand enables the (semi-)automatic integration, repurposing, rearrangement of information. NLP technologies and formalisms for the integrated representation of unstructured and semantic content (such as RDFa and Microdata) aim at bridging this semantic gap. However, in order for humans to truly benefit from this integration, we need ways to author, visualize and explore unstructured and semantic information in a holistic manner. In this paper, we present the WYSIWYM (What You See is What You Mean) concept, which addresses this issue and formalizes the binding between semantic representation models and UI elements for authoring, visualizing and exploration. With RDFaCE and Pharmer we present and evaluate two complementary showcases implementing the WYSIWYM concept for different application domains.

Categories and Subject Descriptors
H.5.2 [User interfaces]: GUIs, Interaction styles

General Terms
Design, Human factors

1. INTRODUCTION

The Semantic Web and Linked Data gained traction in the last years. However, the majority of information still is contained in and exchanged using unstructured documents, such as Web pages, text documents, images and videos. This can also not be expected to change, since text, images and videos are the natural way how humans interact with information. Semantic structuring on the other hand provides a wide range of advantages compared to unstructured information. It facilitates a number of important aspects of information management:

- For search and retrieval enriching documents with semantic representations helps to create more efficient and effective search interfaces, such as faceted search or question answering [27].
- In information presentation semantically enriched documents can be used to create more sophisticated ways of flexibly visualizing information, such as by means of semantic overlays as described in [3].
- For information integration semantically enriched documents can be used to provide unified views on heterogeneous data stored in different applications by creating composite applications such as semantic mashups [2].
- To realize personalization, semantic documents provide customized and context-specific information which better fits user needs and will result in delivering customized applications such as personalized semantic portals [23].
- For reusability and interoperability enriching documents with semantic representations facilitates exchanging content between disparate systems and enables building applications such as executable papers [18].

Natural Language Processing (NLP) technologies (e.g., named entity recognition and relationship extraction) as well as formalisms for the integrated representation of unstructured and semantic content (such as RDFa and Microdata) aim at bridging the semantic gap between unstructured and semantic representation formalisms. However, in order for humans to truly benefit from this integration, we need ways to author, visualize and explore unstructured and semantic information in a holistic manner.

In this paper, we present the WYSIWYM (What You See Is What You Mean) concept, which addresses the issue of an integrated visualization, exploration and authoring of un-structured and semantic content. Our WYSIWYM concept formalizes the binding between semantic representation models and UI elements for authoring, visualizing and exploration. We analyse popular tree, graph and hyper-graph based semantic representation models and elicit a list of semantic representation elements, such as entities, various relationships and attributes. We provide a comprehensive survey of common UI elements for authoring, visualizing and exploration, which can be configured and bound to individual semantic representation elements. Our WYSIWYM concept also comprises cross-cutting helper components, which can be employed within a concrete WYSIWYM interface.
The contributions of this work are in particular:

1. A formalization of the WYSIWYM concept based on definitions for the WYSIWYM model, binding and concrete interfaces.
2. A comprehensive survey of semantic representation elements of tree, graph and hyper-graph knowledge representation formalisms as well as UI elements for authoring, visualization and exploration of such elements.
3. Two complementary use cases, which evaluate different, concrete WYSIWYM interfaces in a generic as well as domain specific context.

The WYSIWYM formalization can be used as a basis for implementations; allows to evaluate and classify existing user interfaces in a defined way; provides a terminology for software engineers, user interface and domain experts to communicate efficiently and effectively. We aim to contribute with this work to making Semantic Web applications more user friendly and ultimately to create an ecosystem of flexible UI components, which can be reused, repurposed and choreographed to accommodate the UI needs of dynamically evolving information structures.

The remainder of this article is structured as follows: In Section 2 we describe the background of our work and discuss the related work. Section 3 describes the fundamental WYSIWYM concept proposed in the paper. Subsections of Section 3 present the different components of the WYSIWYM model. In Section 4 we introduce two implemented WYSIWYM interfaces together with their evaluation results. Finally, Section 5 concludes with an outlook on future work.

2. RELATED WORK

WYSIWYG. The term WYSIWYG as an acronym for What-You-See-Is-What-You-Get is used in computing to describe a system in which content (text and graphics) displayed on-screen during editing appears in a form closely corresponding to its appearance when printed or displayed as a finished product. The first usage of the term goes back to 1971 in the print industry to express the idea that what the user sees on the screen is what the user gets on the printer. Xerox PARC’s Bravo was the first WYSIWYG editor-formatter in 1979. Later on by the emergence of Web and HTML technology, the WYSIWYG concept was also utilized in Web-based text editors. The aim was to reduce the effort required by users to express the formatting directly as valid HTML markup. In a WYSIYG editor users can edit content in a view which matches the final appearance of published content with respect to fonts, headings, layout, lists, tables, images and structure. The first programs for building Web pages with a WYSIYG interface were Netscape Gold, Claris HomePage, and Adobe PageMill.

WYSIWYG text authoring is meanwhile ubiquitous on the Web and part of most content creation and management workflows. It is part of content management systems (CMS), weblogs, wikis, fora, product data management systems and online shops, just to mention a few. However, the WYSIWYG model has been criticized, primarily for the verbosity, poor support of semantics and low quality of the generated code and there have been voices advocating a change towards a WYSIWYM (What-You-See-Is-What-You-Mean) model.

WYSIWYM. The first use of the WYSIWYM term occurred in 1995 aiming to capture the separation of presentation and content when writing a document. The LyX editor[19] was the first WYSIWYM word processor for structure-based content authoring. Instead of focusing on the format or presentation of the document, a WYSIWYM editor preserves the intended meaning of each element. For example, page headers, sections, paragraphs, etc. are labeled as such in the editing program, and displayed appropriately in the browser. Another usage of the WYSIWYM term was by Power et al. [22] in 1998 as a solution for Symbolic Authoring. In symbolic authoring the author generates language-neutral “symbolic” representations of the content of a document, from which documents in each target language are generated automatically, using Natural Language Generation technology.

The WYSIWYM term as defined and used in this paper targets the novel aspect of integrated visualization, exploration and authoring of unstructured and semantic content. The rationale of our WYSIWYM concept is to enrich the existing WYSIWYG presentational view of the content with UI components revealing the semantics embedded in the content and enable the exploration and authoring of semantic content. Instead of separating presentation, content and meaning, our WYSIWYM approach aims to integrate these aspects to facilitate the process of Semantic Content Authoring. There are already some approaches (i.e. visual mapping techniques), which go into the direction of integrated visualization and authoring of structured content.

Visual Mapping Techniques. Visual mapping techniques (a.k.a. knowledge representation techniques) are methods to graphically represent knowledge structures. Most of them have been developed as paper-based techniques for brainstorming, learning facilitation, outlining or to elicit knowledge structures. According to their basic topology, most of them can be related to the following fundamentally different primary approaches [6][25]:

- Mind-Maps. Mind-maps are created by drawing one central topic in the middle together with labeled branches and sub-branches emerging from it. Instead of distinct nodes and links, mind-maps only have labeled branches. A mind-map is a connected directed acyclic graph.
graph with hierarchy as its only type of relation.

- **Concept Maps.** Concept maps consist of labeled nodes and labeled edges linking all nodes to a connected directed graph. The basic node and link structure of a connected directed labeled graph also forms the basis of many other modeling approaches like Entity-Relationship (ER) diagrams and Semantic Networks. These forms have the same basic structure as concept maps but with more formal types of nodes and links.

- **Spatial Hypertext.** A spatial hypertext is a set of text nodes that are not explicitly connected but implicitly related through their spatial layout, e.g., through closeness and adjacency — similar to a pin-board. Spatial hypertext can show fuzzily related items. To fuzzily relate two items in a spatial hypertext schema, they are simply placed near to each other, but possibly not quite as near as to a third object. This allows for so-called “constructive ambiguity” and is an intuitive way to deal with vague relations and orders.

**Binding data to UI elements.** There are already many approaches and tools which address the binding between data and UI elements for visualizing and exploring semantically structured data. Dadzie and Rowe [3] present the most exhaustive and comprehensive survey to date of these approaches. For example, Fresnel [21] is a display vocabulary for core RDF concepts. Fresnel’s two foundational concepts are lenses and formats. Lenses define which properties of an RDF resource, or group of related resources, are displayed and how these properties are ordered. Formats determine how resources and properties are rendered and provide hooks to existing styling languages such as CSS.

Parallax, Tabulator, Explorer, Rhizomer, Spitzler, Fenfire, RDF-Gravity, IsaViz and i-Disc for Topic Maps are examples of tools available for visualizing and exploring semantically structured data. In these tools the binding between semantics and UI elements is mostly performed implicitly, which limits their versatility. However, an explicit binding as advocated by our WYSIWYM model can be potentially added to some of these tools.

In contrast to the structured content, there are many approaches and tools which allow binding semantic data to UI elements within unstructured content (cf. our comprehensive literature study [10]). As an example, Dido [9] is a data-interactive document which lets end users author semantic content mixed with unstructured content in a web-page. Dido inherits exploration capabilities from the underlying Exhibit framework. Loomp as a prove-of-concept for the One Click Annotation strategy is another example in this context. Loomp is a WYSIWYG web editor for enriching content with RDFa annotations. It employs a partial mapping between UI elements and data to hide the complexity of creating semantic data.

### 3. WYSIWYM CONCEPT

In this section we introduce the fundamental WYSIWYM concept and formalize key elements of the concept. Formalizing the WYSIWYM concept has a number of advantages: First, the formalization can be used as a basis for implementations (cf. Section 4). Secondly it allows to evaluate and classify existing user interfaces according to the conceptual model in a defined way. Also, the formalization serves the purpose of providing a terminology for software engineers, user interface and domain experts to communicate efficiently and effectively.

Figure 1 provides a schematic overview of the WYSIWYM concept. The rationale is that elements of a knowledge representation formalism (or data model) are connected to suitable UI elements for visualization, exploration and authoring. Formalizing this conceptual model results in three core definitions (1) for the abstract WYSIWYM model, (2) bindings between UI and representation elements as well as (3) a concrete instantiation of the abstract WYSIWYM model, which we call a WYSIWYM interface.

**Definition 1 (WYSIWYM model).** The WYSIWYM model can be formally defined as a quintuple $(D, V, X, T, H)$ where:

- $D$ is a set of semantic representation data models, where each $D_i \in D$ has an associated set of data model elements $E_{D_i}$;
- $V$ is a set of tuples $(v, C_v)$, where $v$ is a visualization technique and $C_v$ a set of possible configurations for the visualization technique $v$;
- $X$ is a set of tuples $(x, C_x)$, where $x$ is an exploration technique and $C_x$ a set of possible configurations for the exploration technique $x$;
- $T$ is a set of tuples $(t, C_t)$, where $t$ is an authoring technique and $C_t$ a set of possible configurations for the authoring technique $t$;
- $H$ is a set of helper components.

The WYSIWYM model represents an abstract concept from which concrete interfaces can be derived by means of bindings between semantic representation model elements and configurations of particular UI elements.

**Definition 2 (Binding).** A binding $b$ is a function which maps each element of a semantic representation model $e (e \in E_{D_i})$ to a set of tuples $(u_i, c)$, where $u_i$ is a user interface technique $u_i$ ($u_i \in V \cup X \cup T$) and $c$ is a configuration $c \in C_{ui}$.

![WYSIWYM Interface](http://simile-widgets.org/exhibit/)

**Figure 1:** Schematic view of the WYSIWYM model.
Complexity of visual mapping
low

ivation of the WYSIWYM model
I

otion in a binding, we attain a concrete instantiation of our
both are bound to each other encoding a certain configura-
ing between a particular UI and data model element. Once
using a certain configuration (cells). The shades of gray
UI elements (rows) and how they can be bound together
Figure 4: Comparison of existing visual mapping
techniques in terms of semantic expressiveness and
complexity of visual mapping.

Figure 2: Comparison of existing visual mapping
techniques in terms of semantic expressiveness and
complexity of visual mapping.

Figure 4 gives an overview on all data model (columns) and
UI elements (rows) and how they can be bound together
using a certain configuration (cells). The shades of gray
in a certain cell indicate the suitability of a certain binding
between a particular UI and data model element. Once a
selection of data models and UI elements was made and
both are bound to each other encoding a certain configura-
in a binding, we attain a concrete instantiation of our
WYSIWYM model called WYSIWYM interface.

Definition 3 (WYSIWYM Interface). An instantiation
of the WYSIWYM model I called WYSIWYM inter-
face now is a hextuple \((D_I, V_I, X_I, T_I, H_I, b_I)\), where:

- \(D_I\) is a selection of semantic representation data models \((D_I \subset D)\);
- \(V_I\) is a selection of visualization techniques \((V_I \subset V)\);
- \(X_I\) is a selection of exploration techniques \((X_I \subset X)\);
- \(T_I\) is a selection of authoring techniques \((T_I \subset T)\);
- \(H_I\) is a selection of helper components \((H_I \subset H)\).
- \(b_I\) is a binding which binds a particular occurrence of a
data model element to a visualization, exploration and/or authoring technique;

Note, that we limit the definition to one binding, which
means that only one semantic representation model is sup-
ported in a particular WYSIWYM interface at a time. It
could be also possible to support several semantic representa-
tion models (e.g. RDFa and Microdata) at the same time.
However, this can be confusing to the user, which is why
we deliberately excluded this case in our definition. In the
remainder of this sections we discuss the different parts of
the WYSIWYM concept in more detail.

3.1 Semantic Representation Models
Semantic representation models are conceptual data models
to express the meaning of information thereby enabling rep-
resentation and interchange of knowledge. Based on their
expressiveness, we can roughly divide popular semantic rep-

presentation models into the three categories tree-based, graph-
based and hypergraph-based (cf. Figure 2). Each semantic
representation model comprises a number of representation elements, such as various types of entities and rela-
tionships. For visualization, exploration and authoring it is
of paramount importance to bind the most suitable UI ele-
ments to respective representation elements. In the sequel
we briefly discuss the three different types of representation
models.

Tree-based. This is the simplest semantic representation
model, where semantics is encoded in a tree-like structure.
It is suited for representing taxonomic knowledge, such as
thesauri, classification schemes, subject heading lists, con-
cept hierarchies or mind-maps. It is used extensively in bi-
ology and life sciences, for example, in the APG III system
(Angiosperm Phylogeny Group III system) of flowering plant
classification, as part of the Dimensions of the XBRL (eX-
tensible Business Reporting Language) or generically in the
SKOS (Simple Knowledge Organization System). Elements
of tree-based semantic representation usually include:

- \(E_1\): Item – e.g. Magnoliidae, the item representing
  all flowering plants.
- \(E_2\): Item type – e.g. biological term for Magnoli-
  idae.
- \(E_3\): Item-subitem relationships – e.g. Magnoliidae
  referring to subitem magnolias.
- \(E_4\): Item property value – e.g. the synonym flowering
  plant for the item Magnoliidae.
- \(E_5\): Related items – e.g. the sibling item Eudicots to
  Magnoliidae.

Tree-based data can be serialized as Microdata or Microfor-
mats.

Graph-based. This semantic representation model adds more
expressiveness compared to simple tree-based formalisms.
The most prominent representative is the RDF data model,
which can be seen as a set of triples consisting of subject,
predicate, object, where each component can be a URI, the
object can be a literal and subject as well as object can be a
blank node. The most distinguishing features of RDF from a
simple tree-based model are: the distinction of entities in
classes and instances as well as the possibility to express
arbitrary relationships between entities. The graph-based
model is suited for representing combinatorial schemes such
as concept maps. Graph-based models are used in a very
broad range of domains, for example, in the FOAF (Friend
of a Friend) for describing people, their interests and inter-
connections in a social network, in MusicBrainz to publish
information about music albums, in the medical domain (e.g.
DrugBank, Diseases, ChEMBL, SIDER) to describe the
relations between diseases, drugs and genes, or generically
in the SIOC (Semantically-Interlinked Online Communities)
vocabulary. Elements of RDF as a typical graph-based data
model are:

- \(E_1\): Instances – e.g. Warfarin as a drug.
- \(E_2\): Classes – e.g. anticoagulants drug for Warfarin.
- \(E_3\): Relationships between entities (instances or classes)
  – e.g. the interaction between Aspirin as an an-
tiplatelet drug and Warfarin which will increase the
  risk of bleeding.
- \(E_4\): Literal property values – e.g. the half-life for the

<table>
<thead>
<tr>
<th>Semantic Expressiveness</th>
<th>Complexity of Visual Mapping</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>high</td>
</tr>
<tr>
<td>high</td>
<td>low</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Representations</th>
<th>RDF/XML</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree-based</td>
<td>Magnoliidae</td>
</tr>
</tbody>
</table>
| RDF                      | biological term for Magnoli-
                               idae. |
| RDFa                     | Magnoliidae referring to
                               subitem magnolias. |
| Turtle/N3/N-Triples      | Item property value – e.g.  
                               the synonym flowering
                               plant for the item
                               Magnoliidae. |
| JSON-LD                  | Related items – e.g. the
                               sibling item Eudicots to
                               Magnoliidae. |
Amoxicillin.
- $E_4$: Value – e.g. 61.3 minutes.
- $E_5$: Language tag – e.g. en.
- $E_6$: Datatype – e.g. xsd:float.

RDF-based data can be serialized in various formats, such as RDFa, RDF/XML, JSON-LD or Turtle/N3/N-Triples.

**Hypergraph-based.** A hypergraph is a generalization of a graph in which an edge can connect any number of vertices. Since hypergraph-based models allow n-ary relationships between arbitrary number of nodes, they provide a higher level of expressiveness compared to tree-based and graph-based models. The most prominent representative is the **Topic Maps** data model developed as an ISO/IEC standard which consists of topics, associations and occurrences. The semantic expressivity of Topic Maps is, in many ways, equivalent to that of RDF, but the major differences are that Topic Maps (i) provide a higher level of semantic abstraction (providing a template of topics, associations and occurrences, while RDF only provides a template of two arguments linked by one relationship) and (hence) (ii) allow n-ary relationships (hypergraphs) between any number of nodes, while RDF is limited to triplets. The hypergraph-based model is suited for representing complex schemes such as spatial hypertext. Hypergraph-based models are used for a variety of applications. Amongst them are musicDNA, as an index of musicians, composers, performers, bands, artists, producers, their music, and the events that link them together, TM4L (Topic Maps for e-Learning), clinical decision support systems and enterprise information integration. Elements of Topic Maps as a typical hypergraph-based data model are:

- $E_1$: Topic name – e.g. University of Leipzig.
- $E_2$: Topic type – e.g. organization for University of Leipzig.
- $E_3$: Topic associations – e.g. member of a project which has other organization partners.
- $E_4$: Topic role in association e.g. coordinator.
- $E_5$: Topic occurrences – e.g. address.
  - $E_6$: value – e.g. Augustusplatz 10, 04109 Leipzig.
  - $E_7$: datatype – e.g. text.

Topic Maps-based data can be serialized as an XML-based syntax called XTM (XML Topic Map), LTM (Linear Topic Map Notation), CTM (Compact Topic Maps Notation) and AsTMa (Asymptotic Topic Map Notation).

### 3.2 Visualization

The primary objectives of visualization are to present, transform, and convert semantic data into a visual representation, so that, humans can read, query and edit them efficiently. We divide existing techniques for visualization of knowledge encoded in text, images and videos into the three categories **Highlighting**, **Associating** and **Detail view**. Highlighting includes UI techniques which are used to distinguish or highlight a part of an object (i.e. text, image or video) from the whole object. Associating deals with techniques that visualize the relation between some parts of an object. Detail view includes techniques which reveal detailed information about a part of an object. For each of the above categories, the related UI techniques are as follows:

**Highlighting**
- $V_1$: **Framing and Segmentation** (borders, overlays and backgrounds). This technique can be applied to text, images and videos, we enclose a semantic entity in a coloured border, background or overlay. Different border styles (colours, width, types), background styles (colours, patterns) or overlay styles (when applied to images and videos) can be used to distinguish different types of semantic entities (cf. Figure 3 no. 1, 2).
- $V_2$: **Text formatting** (color, font, size, etc.). In this technique different text styles such as font family, style, weight, size, colour, shadows and other text decoration techniques are used to distinguish semantic entities within a text (cf. Figure 3 no. 6). The problem with this technique is that in an HTML document, the applied semantic styles might overlap with existing styles in the document and thereby add ambiguity to recognizing semantic entities.
- $V_3$: **Image color effects.** This technique is similar to text formatting but applied to images and videos. Different image color effects such as brightness/contrast, shadows, glows, bevel/emboss are used to highlight semantic entities within an image (cf. Figure 3 no. 7). This technique suffers from the problem that the applied effects might overlap with the existing effects in the image thereby making it hard to distinguish the semantic entities.
- $V_4$: **Marking** (icons appended to text or image). In this technique, which can be applied to text, images and videos, we append an icon as a marker to the part of object which includes the semantic entity (cf. Figure 3 no. 9). The most popular use of this technique is currently within maps to indicate specific points of interest. Different types of icons can be used to distinguish different types of semantic or correlated entities.
- $V_5$: **Bleeping**. A bleep is a single short high-pitched signal in videos. Bleeping can be used to highlight semantic entities within a video. Different type of bleep signals can be defined to distinguish different types of semantic entities.
- $V_6$: **Speech** (in videos). In this technique a video is augmented by some speech indicating the semantic entities and their types within the video.

**Associating**
- $V_7$: **Line connectors**. Using line connectors is the simplest way to visualize the relation between semantic entities in text, images and videos (cf. Figure 3 no. 4). If the value of a property is available in the text, line connectors can also reflect the item property values. Problematic is that normal line connectors can not express the direction of a relation.
- $V_8$: **Arrow connectors**. Arrow connectors are extended line connectors with arrows to express the direction of a relation in a directed graph.

**Detail view**
- $V_9$: **Callouts**. A callout is a string of text connected by a line, arrow, or similar graphic to a part of text, image or video giving information about that part. It is used in conjunction with a cursor, usually a pointer. The
Figure 3: Screenshots of user interface techniques for visualization and exploration: 1-framing using borders, 2-framing using backgrounds, 3-video subtitle, 4-line connectors and arrow connectors, 5-bar layouts, 6-text formatting, 7-image color effects, framing and line connectors, 8-expandable callout, 9-marking with icons, 10-tooltip callout, 11-faceting

user hovers the pointer over an item, without clicking it, and a callout appears (cf. Figure 3 no. 10). Callouts come in different styles and templates such as infotips, tooltips, hint and popups. Different sort of semantic information can be embedded in a callout to indicate the type of semantic entities, property values and relationships. Another variant of callouts is the status bar which displays semantic information in a bar appended to the text, image or video container. A problem with dynamic callouts is that they do not appear on mobile devices (by hover), since there is no cursor.

- **Video subtitles**: Subtitles are textual versions of the dialog or commentary in videos. They are usually displayed at the bottom of the screen and are employed for written translation of a dialog in a foreign language. Video subtitles can be used to reflect detailed semantics embedded in a video scene when watching the video. A problem with subtitles is efficiently scaling the text size and relating text to semantic entities when several semantic entities exist in a scene.

### 3.3 Exploration

To increase the effectiveness of visualizations, users need to be capable to dynamically explore the visual representation of the semantic data. The dynamic exploration of semantic data will result in faster and easier comprehension of the targeted content. Techniques for exploration of semantics encoded in text, images and videos include:

- **X1: Zooming**: In a zoomable UI, users can change the scale of the viewed area in order to see more detail or less.Zooming in a semantic entity can reveal further details such as property value or entity type. Zooming out can be employed to reveal the relations between semantic entities in a text, image or video. Supporting rich dynamics by configuring different visual representations for semantic objects at different sizes is a requirement for a zoomable UI. The iMapping approach[6] which is implemented in the semantic desktop is an example of the zooming technique.

- **X2: Faceting**: Faceted browsing is a technique for accessing information organized according to a faceted classification system, allowing users to explore a collection of information by applying multiple filters (cf. Figure 3 no. 11). Defining facets for each component of the predefined semantic models enable users to browse the underlying knowledge space by iteratively narrowing the scope of their quest in a predetermined order. One of the main problems with faceted browsers is the increased number of choices presented to the user at each step of the exploration[5].

- **X3: Bar layouts**: In the bar layout, each semantic en-
tity within the text is indicated by a vertical bar in the left or right margin (cf. Figure 3 no. 5). The colour of the bar reflects the type of the entity. The bars are ordered by length and order in the text. Nested bars can be used to show the hierarchies of entities. Semantic entities in the text are highlighted by a mouse-over the corresponding bar. This approach is employed in Loomp [17].

- $X_4$: Expandable callouts. Expandable callouts are interactive and dynamic callouts which enable users to explore the semantic data associated to a predefined semantic entity (cf. Figure 3 no. 8). This technique is employed in OntosFeeder [11].

### 3.4 Authoring

Semantic authoring aims to add more meaning to digitally published documents. If users do not only publish the content, but at the same time describe what it is they are publishing, then they have to adopt a structured approach to authoring. A semantic authoring UI is a human accessible interface with capabilities for writing and modifying semantic documents. The following techniques can be used for authoring of semantics encoded in text, images and videos:

- $T_1$: Form editing. In form editing, a user employs existing form elements such as input/check/radio boxes, drop-down menu, slider, spinner, buttons, date/color picker etc. for content authoring.
- $T_2$: Inline edit. Inline editing is the process of editing items directly in the view by performing simple clicks, rather than selecting items and then navigating to an edit form and submitting changes from there.
- $T_3$: Drawing. Drawing as part of informal user interfaces [14], provides a natural human input to annotate an object by augmenting the object with human-understandable sketches. For instance, users can draw a frame around semantic entities, draw a line between related entities etc. Special shapes can be drawn to indicate different entity types or entity roles in a relation.
- $T_4$: Drag and drop. Drag and drop is a pointing device gesture in which the user selects a virtual object by grabbing it and dragging it to a different location or onto another virtual object. In general, it can be used to invoke many kinds of actions, or create various types of associations between two abstract objects.
- $T_5$: Context menu. A context menu (also called contextual, shortcut, or pop-up menu) is a menu that appears upon user interaction, such as a right button mouse click. A context menu offers a limited set of choices that are available in the current state, or context.
- $T_6$: (Floating) Ribbon editing. A ribbon is a command bar that organizes functions into a series of tabs or toolbars at the top of the editable content. Ribbon tabs/toolbars are composed of groups, which are a labeled set of closely related commands. A floating ribbon is a ribbon that appears when user rolls the mouse over a target area. A floating ribbon increases usability by bringing edit functions as close as possible to the user’s point of focus. The Aloha WYSIWYG editor [1] is an example of floating ribbon based content authoring.
- $T_7$: Voice commands. Voice commands permit the user’s hands and eyes to be busy with another task, which is particularly valuable when users are in motion or outside. Users tend to prefer speech for functions like describing objects, sets and subsets of objects [20].
- $T_8$: (Multi-touch) gestures. A gesture (a.k.a. sign language) is a form of non-verbal communication in which visible bodily actions communicate particular messages. Technically, different methods can be used for detecting and identifying gestures. Movement-sensor-based and camera-based approaches are two commonly used methods for the recognition of in-air gestures [19]. Multi-touch gestures are another type of gestures which are defined to interact with multi-touch devices such as modern smartphones and tablets. Users can use gestures to determine semantic entities, their types and relationship among them. The main problem with gestures is their high level of abstraction which makes it hard to assert concrete property values.

### 3.5 Bindings

Figure 4 gives an overview on possible bindings between the user interface and semantic representation elements defined in previous sections. These possible bindings were obtained from a comprehensive survey of existing tools and an analysis of possible connections between a specific UI element and a semantic model element. The following binding configurations are available and refereed to from the cells of Figure 4.

- Defining a special border or background style ($C_1$), text style ($C_2$), image color effect ($C_3$), beep sound ($C_4$), bar style ($C_5$), sketch ($C_7$), draggable or drop-pable shape ($C_8$), voice command ($C_9$), gesture ($C_{10}$) or a related icon ($C_{11}$) for each type.
- Progressive shading ($C_{11}$) by defining continuous shades within a specific color scheme to distinguish items in different levels of the hierarchy.
- Hierarchical bars ($C_{12}$) by defining special styles for nested bars.
- Grouping by similar border or background style ($C_{13}$), text style ($C_{14}$), icons ($C_{15}$) or image color effects ($C_{16}$).

### 3.6 Helper Components

In order to facilitate, enhance and customize the WYSIWYM model, we utilize a set of helper components, which implement cross-cutting aspects. A helper component acts as an extension on top of the core functionality of the WYSIWYM model. The following components can be used to improve the quality of a WYSIWYM UI depending on the requirements defined for a specific application domain:

- $H_1$: Automation means the provision of facilities for automatic annotation of text, images and videos to reduce the need for human work and thereby facilitating the efficient annotation of large item collections. For example, users can employ existing NLP services (e.g. named entity recognition, relationship extraction) for automatic text annotation.
- $H_2$: Real-time tagging is an extension of automation, which allows to create annotations proactively while the user is authoring a text, image or video. This will significantly increase the annotation speed and users.
### Visual Representation-Based Approaches

<table>
<thead>
<tr>
<th>Structure encoded in:</th>
<th>UI categories</th>
<th>UI techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Highlighting</strong></td>
<td>text</td>
<td>Framing and segmentation (e.g., Taxonomies)</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>Text formatting (e.g., color, font, size etc.)</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>Marking (appended icons)</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>Line connectors</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>Arrow connectors</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>Callouts (text)</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>Callouts (images)</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>Callouts (videos)</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>Callouts (subtitle)</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>Zooming</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>Faceting</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>Bar layout</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>Expandable callouts</td>
</tr>
<tr>
<td></td>
<td>images</td>
<td>Zooming</td>
</tr>
<tr>
<td></td>
<td>images</td>
<td>Faceting</td>
</tr>
<tr>
<td></td>
<td>images</td>
<td>Faceting (excerpt)</td>
</tr>
<tr>
<td></td>
<td>videos</td>
<td>Faceting</td>
</tr>
<tr>
<td></td>
<td>videos</td>
<td>Form editing</td>
</tr>
<tr>
<td></td>
<td>videos</td>
<td>Inline edit</td>
</tr>
<tr>
<td></td>
<td>videos</td>
<td>Drawing</td>
</tr>
<tr>
<td></td>
<td>videos</td>
<td>Drag and drop</td>
</tr>
<tr>
<td></td>
<td>videos</td>
<td>Context menu</td>
</tr>
<tr>
<td></td>
<td>videos</td>
<td>(Floating) Ribbon editing</td>
</tr>
<tr>
<td></td>
<td>videos</td>
<td>Voice commands</td>
</tr>
<tr>
<td></td>
<td>videos</td>
<td>(Multi-Touch) Gestures</td>
</tr>
</tbody>
</table>

**Table Notes:**
- **No binding**
- **Partial binding**
- **Full binding**

**Figure 4:** Possible bindings between user interface and semantic representation model elements.
are not distracted since they do not have to interrupt their current authoring task.

- **H3:** Recommendation means providing users with pre-filled form fields, suggestions (e.g., for URIs, namespaces, properties), default values etc. These facilities simplify the authoring process, as they reduce the number of required user interactions. Moreover, they help preventing incomplete or empty metadata. In order to leverage other user’s annotations as recommendations, approaches like Paragraph Fingerprinting can be implemented.

- **H1:** Personalization and context-awareness describes the ability of the UI to be configured according to users’ contexts, background knowledge and preferences. Instead of being static, a personalized UI dynamically tailors its visualization, exploration and authoring functionalities based on the user profile and context.

- **H2:** Collaboration and crowdsourcing enables collaborative semantic authoring, where the authoring process can be shared among different authors at different locations. There are a vast amounts of amateur and expert users which are collaborating and contributing on the Social Web. Crowdsourcing harnesses the power of such crowds to significantly enhance and widen the results of semantic content authoring and annotation. Generic approaches for exploiting single-user Web applications for shared editing can be employed in this context.

- **H5:** Accessibility means providing people with disabilities and special needs with appropriate UIs. The underlying semantic model in a WYSIWYM UI can allow alternatives or conditional content in different modalities to be selected based on the type of the user disability and information need.

- **H7:** Multilinguality means supporting multiple languages in a WYSIWYM UI when visualizing, exploring or authoring the content.

4. IMPLEMENTATION AND EVALUATION

In order to evaluate the WYSIWYM model, we implemented the two applications RDFaCE and Pharmer, which we present in the sequel.

RDFaCE. (RDFa Content Editor) is a WYSIWYM interface for semantic content authoring. It is implemented on top of the TinyMCE rich text editor. RDFaCE extends the existing WYSIWYG user interfaces to facilitate semantic authoring within popular CMSs, such as blogs, wikis and discussion forums. The RDFaCE implementation (cf. Figure 5, left) is open-source and available for download together with an explanatory video and online demo at [http://aksw.org/Projects/RDFaCE](http://aksw.org/Projects/RDFaCE).

RDFaCE, as a WYSIWYM instantiation can be described using the following hextuple:

- **D:** RDFa, Microdata;
- **V:** Framing using borders (C: special border color defined for each type). Callouts using dynamic tooltips;
- **E:** Faceting based on the type of entities;
- **T:** Form editing, Context Menu, Ribbon editing;
- **H:** Automation, Recommendation;
- **b:** bindings defined in [Figure 4].

Microdata support is implemented in RDFaCE-Lite available at [http://rdface.aksw.org/lite](http://rdface.aksw.org/lite).

Table 1: Usability evaluation results for RDFaCE.

<table>
<thead>
<tr>
<th>Usability Factor/Grade</th>
<th>Poor</th>
<th>Fair</th>
<th>Neutral</th>
<th>Good</th>
<th>Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fit for use</td>
<td>0%</td>
<td>12.50%</td>
<td>31.25%</td>
<td>43.75%</td>
<td>12.50%</td>
</tr>
<tr>
<td>Ease of learning</td>
<td>0%</td>
<td>12.50%</td>
<td>50%</td>
<td>31.25%</td>
<td>6.25%</td>
</tr>
<tr>
<td>Task efficiency</td>
<td>0%</td>
<td>0%</td>
<td>56.25%</td>
<td>37.50%</td>
<td>6.25%</td>
</tr>
<tr>
<td>Ease of remembering</td>
<td>0%</td>
<td>0%</td>
<td>37.50%</td>
<td>50%</td>
<td>12.50%</td>
</tr>
<tr>
<td>Subjective satisfaction</td>
<td>0%</td>
<td>18.75%</td>
<td>50%</td>
<td>25%</td>
<td>6.25%</td>
</tr>
<tr>
<td>Understandability</td>
<td>6.25%</td>
<td>18.75%</td>
<td>31.25%</td>
<td>37.50%</td>
<td>6.25%</td>
</tr>
</tbody>
</table>

Figure 6: Usability evaluation results for Pharmer

In order to evaluate RDFaCE usability, we conducted an experiment with 16 participants of the ISSLOD 2011 summer school. The user evaluation comprised the following steps: First, some basic information about semantic content authoring along with a demo showcasing different RDFaCE features was presented to the participants as a 3 minutes video. Then, participants were asked to use RDFaCE to annotate three text snippets – a wiki article, a blog post and a news article. For each text snippet, a timeslot of five minutes was available to use different features of RDFaCE for annotating occurrences of persons, locations and organizations with suitable entity references. Subsequently, a survey was presented to the participants where they were asked questions about their experience while working with RDFaCE. Questions were targeting six factors of usability namely Fit for use, Ease of learning, Task efficiency, Ease of remembering, Subjective satisfaction and Understandability. Results of the survey are shown in Table 1. They indicate on average good to excellent usability for RDFaCE. A majority of the users deem RDFaCE being fit for use and its functionality easy to remember. Also, easy of learning and subjective satisfaction was well rated by the participants. There was a slightly lower (but still above average) assessment of task efficiency and understandability, which we attribute to the short time participants had for familiarizing themselves with RDFaCE and the quite comprehensive functionality, which includes automatic annotations, recommendations and various WYSIWYM UI elements.

Pharmer. Pharmer is a WYSIWYM interface for the authoring of semantically enriched electronic prescriptions. It...
enables physicians to embed drug-related metadata into e-prescriptions thereby reducing the medical errors occurring in the prescriptions and increasing the awareness of the patients about the prescribed drugs and drug consumption in general. In contrast to database-oriented e-prescriptions, semantic prescriptions are easily exchangeable among other e-health systems without need to changing their related infrastructure. The Pharmer implementation (cf. Figure 5, right) is open-source and available for download together with an explanatory video and online demo at http://code.google.com/p/pharmer/. It is based on HTML5 contenteditable element. Pharmer as a WYSIWYM instantiation is defined using the following hextuple:

- D: RDFa.
- V: Framing using borders and background (C: special background color defined for each type), Callouts using dynamic popups.
- E: Faceting based on the type of entities.
- T: Form editing, Inline edit.
- b: bindings defined in Figure 4.

In order to evaluate the usability of Pharmer, we performed a user study with 13 subjects. Subjects were 3 physicians, 4 pharmacist, 3 pharmaceutical researchers and 3 students. We first showed them a 3-minute tutorial video of using different features of Pharmer then asked each one to create a semantic prescription with Pharmer. After finishing the task, we asked the participants to fill out a questionnaire. We used the System Usability Scale (SUS) as a standardized, simple, ten item Likert scale-based questionnaire to grade the usability of the system. SUS yields a single number in the range of 0 to 100 which represents a composite measure of the overall usability of the system. The results of our survey (cf. Figure 6) showed a mean usability score of 75 for Pharmer WYSIWYM interface which indicates a good level of usability. Participants particularly liked the integration of functionality and the ease of learning and use. The confidence in using the system was slightly lower, which we again attribute to the short learning phase and diverse functionality.

5. CONCLUSIONS

Bridging the gap between unstructured and semantic content is a crucial aspect for the ultimate success of semantic technologies. With the WYSIWYM concept we presented in this article an approach for integrated visualization, exploration and authoring of unstructured and semantic content. The WYSIWYM model binds elements of a knowledge representation formalism (or data model) to a set of suitable UI elements for visualization, exploration and authoring. Based on such a declarative binding mechanism, we aim to increase the flexibility, reusability and development efficiency of semantics-rich user interfaces.

We deem this work as a first step in a larger research agenda aiming at improving the usability of semantic user interfaces, while retaining semantic richness and expressivity. In future work we envision to adopt a model-driven approach to enable automatic implementation of WYSIWYM interfaces by user-defined preferences. This will help to reuse, re-purpose and choreograph WYSIWYM UI elements to accommodate the needs of dynamically evolving information structures and ubiquitous interfaces. We also aim to bootstrap an ecosystem of WYSIWYM instances and UI elements to support structure encoded in different modalities, such as images and videos. Creating live and context-sensitive WYSIWYM interfaces which can be generated on-the-fly based on the ranking of available UI elements is another promising research venue.

Acknowledgments

We would like to thank our colleagues from the AKSW research group for their helpful comments and inspiring discussions during the development of the WYSIWYM model. This work was partially supported by a grant from the European Union’s 7th Framework Programme provided for the project LOD2 (GA no. 257943).
6. REFERENCES


