

cific knowledge satisfying clearly defined requirements is placed, significantly simplifies navigation through the knowledge base and clearly localises the evolution of knowledge located on each "semantic shelf". From a meaningful point of view, the subject domain is a set of factual statements describing all elements of a given set of objects of research with the help of a given set of relations and parameters (characteristics).

A section — is a sign of a set of all possible sections included in different knowledge bases. Each section represents a conditionally didactically distinguished fragment of the knowledge base, possessing logical integrity and completeness. In the limit, the whole knowledge base of a particular ostis-system is also one large non-atomic section.

For each partition it is necessary to explicitly specify the belonging to a set of atomic or non-atomic partitions. Atomic partition — a sign of the set of all possible atomic partitions included in various documentation, i.e., partitions not decomposable into more private partitions. Non-atomic section — the sign of the set of all possible non-atomic sections that make up the various documentations, that is, sections that are decomposed into more private sections.

In the context of the *OSTIS Glossary*, it may be appropriate to identify sections that will describe some portion of that information described in the relevant sections of the *OSTIS Standard*. But nobody forbids to consider the *OSTIS Glossary* as some atomic section in which all objects are listed according to the external natural language alphabet together with their specifications. As for subject domains, it is more expedient to use them not as a means for knowledge stratification, but as a means for searching already existing information to form sections with this information.

So, let's list the **basic rules of structuring and specification of the *OSTIS Glossary* objects**:

- All information about the *OSTIS Glossary* objects should be represented in the form of semantic neighbourhoods of these objects, in which will be listed:
  - identification variants (variants of terms) of a given object in various external (natural) languages;
  - membership of the object in some subject domain with indication of the role performed within this subject domain, including membership of this object in the corresponding section of the knowledge base or the *OSTIS Standard*;
  - theoretical-multiple relations of the given object with other objects, including:
    - \* partition and decomposition of a given object into other objects;
    - \* including a given object into other objects;

- \* and other possible unions, intersections that form a given object;
- definition of the given object;
- explanation of the given object;
- description of a typical example of using the specified object;
- instances of the described object, if the given object is a concept;
- authors of the given object;
- authors of the specification, i.e. the authors who described the given object in the knowledge base;
- analogues of this object, including:
  - \* close analogues of the given object;
  - \* differences and similarities with other objects, including listed analogues;
- bibliographic sources of the object;
- possible quotations, aphorisms, metaphors and epigraphs related to the given object; This variant of specification will be called dictionary specification.

- In this case, all objects and their specifications within the *OSTIS Glossary* can be ordered:
  - as an enumeration of these objects in the lexicographic order of their terms, forming one single atomic section, respectively being the *OSTIS Glossary*;
  - and the enumeration of sections corresponding to the sections of the *OSTIS Standard*, uniting objects by common features, considered within a particular section, and representing sequences of these objects in the lexicographic order of their terms.

For the second case, it is important that the sections do not overlap with each other, i.e., do not duplicate descriptions of the same entity. From this point of view, the first option is easier to implement, because it does not require taking into account the possibility of occurrence of the found object in the already formed sections of the *OSTIS Glossary*.

- Structuring of objects of the *OSTIS Glossary* should be reduced to the formation of text from the already structured text of the *OSTIS Standard*.

#### D. Rules for identification of the *OSTIS Glossary* objects

In the previous section the rules of structuring and specifying (standardisation) of the *OSTIS Glossary* objects were fixed. However, it is necessary to standardise not only those texts that are written directly in *SC-code*, but also those texts that are written in external, e.g. natural languages. One of the features of the *SC-code* is that with it it is also possible to record some natural language files, by means of which the comprehension of the text by any human being is improved. Such a possibility is realised with the help of ostis-system files,

with the help of which external information constructions which are not text in *SC-code* [43] are denoted. With their help it is possible to specify information in an external natural language for all objects in the knowledge base. A special case of ostis-system files are files denoting identifiers of objects in the knowledge base.

Identifier is a structured sign representing an entity denoted by a string of symbols. An identifier is an information construct (most often a string of symbols) providing unambiguous identification of the corresponding object described in knowledge bases of ostis-systems, and is, most often, a name (term) corresponding to the described object, a name denoting this object in external texts of ostis-systems.

In formal texts, identifiers must be unique to uniquely match an object. Each pair of identical identifiers must denote the same object.

All objects in a knowledge base have the following **common identification rules**:

- The membership of the identified object in some object classes is explicitly specified in the external identifier of this object (in the sc-identifier) by means of appropriate conditional attributes:
  - if the last character of the sc-identifier is an “asterisk” character, then the identified object belongs to the Class of non-role relation designations;
  - if the last character of the sc-identifier is an apostrophe, then the identified object belongs to the Class of role relationship designations, each of which is a subset of Membership relation;
  - if the last character of the sc-identifier is “^”, then the identified object belongs to the Parameter designation class.
- For each object, you can construct an sc-identifier, which is a proper name that always starts with a capital letter.
- If an object is a designation of some class of objects, then this object can be matched not only with a proper name, but also with a common name, which starts with a small (lowercase) letter. The specification of each class (each concept) includes a list of equivalent (synonymous) sc-identifiers, among which there are both proper and nominative names.
- Identification of partitions in knowledge bases is performed at the expense of identification of partition objects. Instances of partition classes within the Russian language are named according to the following rules:
  - at the beginning of the identifier the word Section is written and a dot is put;
  - followed by the name of the section with a capital letter, reflecting its content.

Obviously, the same rules apply to the *OSTIS Glossary* objects. Therefore, there is no need to describe any addi-

tional rules for identifying the *OSTIS Glossary* objects.

Besides identifiers of ostis-systems knowledge base objects, it is possible to standardise the form of presentation of information in the specification of these objects: both the style of writing the text itself and the information with the help of which it is possible to explain the basic information in the specification of these objects.

#### *E. Key elements of didactic information in the OSTIS Glossary object specifications*

The most important criterion of quality of created ostis-systems of any purpose is to create conditions so that insufficiently qualified users of each ostis-system (both end-users and those responsible for its effective operation and modernisation) could acquire the required qualification quickly enough with the help of the same ostis-system. This means that each ostis-system, irrespective of its direct purpose (automation of specific types of human activities in a particular field) should also be a training system, i.e. it should be able to train its users in the direction of improving their qualification. A qualified user of any category must understand the capabilities of the ostis-system with which he interacts, must understand what the system knows and can do, as well as how its activities can be managed. Lack of understanding between ostis-systems and their users — is a violation of the interoperability requirement [44] imposed on both ostis-systems and their users.

Therefore, the key stage in the development of any knowledge base, and, in general, any information resource is the stage of development and implementation of didactic information. Didactic information should be understood as a specification of the subject domain, which provides additional information designed to enable users and developers (knowledge engineers), who use or improve the specified subject domain and its ontology, to learn their features faster. Didactic information enables [45]:

- to quickly and adequately assimilate the denotational semantics of knowledge stored in the system;
- to provide a deeper understanding and assimilation of the meaning of various kinds of entities (including various knowledge);
- to establish mutual understanding between systems and their users;
- to accelerate the process of formation of the required qualification of users in various fields.

The “didactic” effect of didactic information is provided by:

- by sufficient detail of the studied entity (completeness of the semantic neighbourhood describing the relations of this entity with other entities)
  - decomposition of the entities under consideration;
  - by specifying analogues (similar entities in different senses);

- indicating metaphors (epigraphs);
- indicating antipodes (entities that differ in different senses);
- exercises — solutions to various problems using the entities studied;
- references to knowledge stored within the same knowledge base;
- references to bibliographic sources.

Following one of the goals of the *OSTIS Glossary*, namely to provide a quality and understandable text for the end user, it is necessary to describe and record the **main elements of didactic information for the objects of the *OSTIS Glossary***. Such elements should include:

- information by means of which the basic information in the *OSTIS Glossary* object specification is defined or explained, i.e.:
  - various types of definitions, explanations, and annotations for that object;
  - examples of how to use this object;
  - instances of this object, if it is a concept;
- information describing distinctive and similar characteristics between the *OSTIS Glossary* objects, including:
  - analogies, correspondences, corollaries between objects;
  - differences and similarities between objects;
- information that supports the significance, scientific novelty, and practical applicability of the *OSTIS Glossary* object, including:
  - of the authors of these objects, as well as the authors who specified this object;
  - bibliographic sources of these objects;
  - quotations, aphorisms, metaphors and epigraphs related to the given object;
  - other.

It should be noted that the quality of the information described in the knowledge base directly depends on the quality of the presentation of this information and the quality of the information with the help of which the basic information in the knowledge base is explained. The more qualitatively the information in the knowledge base is described, structured and stratified, the lower the requirements to the readers of this information and the higher the level of understanding of the information in the knowledge base by these readers [46].

Didactic information in the knowledge base determines the level of quality of the information described in the knowledge base. The key elements of didactic information are those elements that contribute to easier and deeper mastering of the basic information in the knowledge base. The authors believe that it is the information that reveals similarities between objects in the knowledge base that is the key to improving the quality of the entire knowledge base.

## F. Conclusion

To summarize, it is important to note the following points:

- The *development of the OSTIS Glossary* is reduced to the development of the *OSTIS Standard* and a set of tools that allow to form and improve the *OSTIS Glossary* on the basis of the *OSTIS Standard*;
- When developing the *OSTIS Glossary* as part of the *OSTIS Standard*, it is important to fix and improve the principles and rules:
  - specification and formalisation of objects;
  - identification of objects and their specifications;
  - of structuring and stratification of object specifications;
  - development and consistency of objects and their specifications;
- The *quality of the OSTIS Glossary* is determined by:
  - quality of the *OSTIS Standard*, which is defined by:
    - \* the quality of the information it contains;
    - \* the quality of the means to improve this information, which is determined by:
      - the quality of the methods, principles and rules for developing this information;
      - the quality of didactic information explaining this information;
    - \* the competence and level of training of the developers of the *OSTIS Standard*.

In the final section the current state of the *OSTIS Glossary* within the *OSTIS Metasystem* will be considered, the principles of automatic formation of the *OSTIS Glossary*, information retrieval in it will be fixed, and also the current Author team of the *OSTIS Glossary* and requirements to its developers will be considered.

## IV. Implementation of the *OSTIS Glossary* within the *OSTIS Metasystem*

### A. Current specification and structure of the *OSTIS Glossary*

It is important to note that the *OSTIS Glossary* is not just a dynamically generated text from the text of the *OSTIS Standard*, which is a simplified representation of the *OSTIS Standard*. The *OSTIS Glossary* acts as a means to provide a consistent and interoperable activity for the development of new generation intelligent computer systems, and thus has documentation for its use and development. Therefore, first of all, the *OSTIS Glossary* is documentation on how to properly form a dictionary representation of the *OSTIS Standard*.

### The *OSTIS Glossary*

:= [Semantic electronic dictionary of Artificial intelligence]