The Principles of Building Personal Data Integrators Using OSTIS Technology

Pavel Titenkov, Alexey Kaeshko
Belorussian State University of Informatics and Radioelectronics
Minsk, Belarus
Email: pavel@titenkov.com

***Abstract*—This paper describes principles of building personal data integrators using OSTIS technology. The main goal of creating such kind of systems is adaptation of heterogeneous services and transforming them into one personal information space. Personal data integrator is designed to become a one-way interface to the digital world, that will make an attempt to make this interaction easier and more efficient, allowing to solve user's specific tasks based on his specific context. Goal of the paper is to make design of personal integrator system that would understand the semantics of user personal data, understand how to interact with each specific user web service and be able to extend their functionality by creating intelligent agents. It is designed using the ontological principles and OSTIS technology. Personal data integrator is designed to be connected with several web-services to harvest the available user data and transform it into a powerful fragments of knowledge-base, integrated on semantic layer, that can be easily used by intelligent agents of the system.**

Keywords*—*personal data integrator, ontology, semantic network.

I. Introduction

During the evolution of information technologies absolutely new infrastructure appeared, that we currently can't imagine our life without: internet, web-services, smartphones, apps, social networks, etc. All of that changed the style of people's life and put them into a new type of space - information space. The personification of tools for this information access founded a new class of this space - personal information space. From the use cases point of view, it's very important to adopt this personal information space for goals and purposes of each concrete person, but there are no such technical solutions to solve that problem so far.

A. Problems and goals

The high demand of quick access to information resulted in a huge amount of different information services in internet, like email services, messengers, apps for time and task management, calendars, etc. Lack of any standards for building such kind of services resulted in theirs architectural, model and technological heterogeneity [1] [7] [8]. Models that have been used in each service were different both on technology and semantic levels as well, because of no binding between used technologies and domain objects.

All these factors influence on the end user in different ways:

- the need to adapt user to the interface in each service implementation;
- the complexity of initial service setup for personal needs;
- all user personal information is spread between different services, that leads to the search problems and duplications;
- the need to use several services to make task done.

All mentioned problems cause the need of special system creation, that will provide simple and efficient interface to manage the personal information space of a user.

To archive this goal we need to solve the following problems:

- Design an efficient method of data integration on the semantic layer. It's obvious that we need some common abstraction layer to integrate services from different domain objects. Integration of any data will bring us a value only if it would be integrated on the semantic layer [5] [12], otherwise some of the problems mentioned above still would not be solved. The problem of information services integration lies in the logical and semantic combination of heterogeneous data coming from different information sources, that will provide the unified representation of it and will give one common interface for manipulations [14].
- Design the universal user model and user information space. To personalize integrator for user's needs it's required to design the user model, that of course will be different for each person, but at least it should have unified skeleton.
- Design the universal information service model.
- Design models and methods of data integration. It's important to decide to consolidate data or only to integrate. The consolidation approach will solve issue of information spreading between sources by storing everything in one system, but will raise several other technical issues, like storing huge amount of data, keeping it in sync, etc. The virtual integration means that all data will still be stored in their source services, but user will get one common interface to manipulate it. Virtual integration doesn't exclude an ability of copying some most regular used information to the system, but at least the most of the data will stay in source services.
- Design the unified user interface of personal integrator.

The personal integrator of information services should allow user to get any personal information without a need of searching for source service, without a need of manual combining of data retrieved from different sources, without a need of using several services to get one task done. So the main goal of creating such kind of system is a personal maintenance of a user information need and making a one-way interface for communication with personal information space.

B. The analysis of current personal information assistants

Today there are a number of approaches to construct intelligent assistants and integrators of information services, which can be divided into several categories depending on the form of user and system intercourse and the way of system interpretation of coming information and queries. Primarily such solutions are focused on mobile platforms and employed by big companies that usually own information services are integrated.

1) Siri: Siri is a personal assistant made by Apple. It represents question answering system for the iOs operating system. Siri is based on the natural language recognition and machine learning technologies. Siri can be adapted to user pronunciation with time and can be integrated with a number of smartphone apps, such as calendar, contacts, music library, photo gallery [7].

Potentialities:

- Understanding of voice input. Siri knows how to react to a number of definite commands, respond to the questions.
- Reminders. Siri is able to remind the user about calendar events, some out-of-date tasks and other events, the user is subscribed for.
- Reporting of reference information. Siri can interact with search systems to seek the information and give the list of found resources conformed to the request.
- Interaction with other devices (Internet of things). Siri can control a TV-set, a laptop, a garage door, lights, watching cameras, etc.
- Interaction with social media.

In spite of impressive possibilities Siri has its own shortcomings. The key challenge is a weak personification, absence of exact user model, limited API given by implementers and primitiveness of build-in commands. Actually Siri is a natural language interface to interact with some apps which doesn't integrate user services into common information space, but delegate the execution of search and other user requests to the services.

2) Google Assistant: In contrast to Siri Google Assistant appeared not long ago as upgrade version of Google Now. Assistant represent a similar collection of possibilities but thanks to the absolute integration with other Google services, particularly with search service, gives wider capabilities to the user for the purpose of answers on different types of questions.

3) Cortana: Cortana is a virtual voice assistant with the elements of artificial intelligence from Microsoft Company. Cortana has the similar possibilities integrated Micrisoft services to each other. In contrast to Siri and Google Assistant the search system Bing is used. Cortana collects all the information about user accessible on the user laptop and cell, transforms it into internal representation and sends to Microsoft. Also Cortana demands gaining and sending off statistics about all information input to the computer from the operating system. One of the key Cortana advantage is the advanced system of different personal data authorization rules which is more flexible than the competitor ones.

C. Disadvantages

All the examined services have the similar set of possibilities and differ from each other in inconspicuous details. Siri has the best quality of speech recognition and synthesis, Google Assistant has excellent user model based on deep learning algorithms, Cortana has the control system of different personal data authorization rules. In spite of these advantages all the intelligent assistants listed above don't integrate personal information on the semantic level and don't solve a problem of personal information space heterogeneity [14] [22]. Examined intelligent assistants provide universal information access natural language interface and use service-oriented approach to integrate information resources for this purpose. This approach has a number of limitations [22] [23]. If we speak about intelligent system building, the main limitation is the impossibility of expansion of initial services capability and personification for concrete person tasks and necessity [21].

II. Provided method for personal intelligent integrator implementation

Data integration base of listed solutions is the search systems. Consequently they use syntax approach based on the data similarity. There is no doubt that machine learning methods, effective enough, are used in such powerful search machines as Google. These methods allow the photographs to be found even by text description. But even if machine learning is used, information is represented as a set of bits and bytes. The semantics is not taken into account. If data have their own semantics descriptions it will be possible to get to a whole new integration level – semantic level.

When this happens data semantics can be taken into account during data transmission and be used for its integration into knowledge base.

The idea of semantic technologies application to organize data interchange between information services is obvious enough [14]. If one system gives the other one not only the data but also information about its subject entity, it helps to treat exchanging system separately from each other better than using download to intermediate format or web-services SOA. Consequently systems report facts to each other. Classical integration solution consists of information services supplied information, mediators connected to them and an integration service. The main complication of this approach is the creation of united interface between mediators and the integration service.

Our approach to mediators creation is based on building of a united integrated services ontology, described all kinds of objects supplied by services and their connections, as well as services. Therefore there are models of information resources and models of data given by them in the ontology. As soon as the state of the data, the user is interested in, is changed in the information service, the proper mediator become active. It transforms necessary information according to the model described in the form of information resource ontology.

Such approach provides a number of advantages:

- Data transmission model doesn't depend on the model which presents data of donor information service;
- Mediator can be implemented in any programming language. Information services ontology is the standard and background information for its creation.

The main advantage of ontological approach to designing is considerable rise of design system flexibility. Except designing of subject domain ontology the meaning of ontological approach to information resource integration comes to the unification of data and its context and keeping of data with its metadata that provides an opportunity to take the data nature into account. The base of the integrator is the knowledge base containing both ontology of integrating information services and ontology of supplied data subject domains. The similar concepts from different resources are the points of integration. Therefore mediators are in fact agents of integrated knowledge base. Different kinds of agents and their groups (search agents, agents of marking subject matter out, conversion agent) can attend to knowledge base except mediators.

Primarily literary sources recommend Semantic Web technologies when choosing the semantic technology for data integration [16].

Unquestionable advantage of Semantic Web project is the commitment to independent distributed development of ontology. Knowledge of subject domain can be accumulated and defined gradually with the participation of great number of people without constant agreement. There are a set of tools ready for ontology designing and semantic repository implementation.

The most talked about problems of Semantic Web tools, especially of OWL language, are:

- Absence of answer, what part should be modeled with the help of classes and what part – with the help of samples. There is the ambiguity at determining of classes and their samples in Semantic Web tools [17].
- One more shortcoming is absence of possibility to define properties of properties directly. It prevents from modeling of subject domain attributes, n-ary relations and attributes of attributes.
- Web-orientation of the project and semantic network representation close to machine representation.
- Undeveloped standards of time variables view and fuzzy subject domains view.
- Weakly studied level of the ontology verification directed to the authenticity and completeness.

It's obvious that in spite of its popularity Semantic Web tools have a number of shortcomings. Firstly because Semantic Web tools were originally directed to the machine-oriented description of information resources in the web-space without including of comprehensive approach to the problems of semantic view in the context of artificial intelligence theory.

The alternative technology of intelligent systems designing is Open Semantic Technologies for Intelligent Systems (OSTIS) [24]. OSTIS project is intended to intelligent information systems and their components design. In this case such systems will be based on the knowledge presented in the form of ontology. As part of this article systems controlled by knowledge built on the OSTIS technology will be named as ostis-systems.

If there is a set task, information systems mediated on the base of OSTIS technology can be not completely intelligent but serve for the accumulation of data formalized by complicated model. As opposite to Semantic Web technology project OSTIS tools have a strict set-theoretic interpretation and aren't attached to specific application sector. That provides more compact and technically accurate view of the information.

It's defined by a number of properties that allows talking about language means of OSTIS project as about the most preferable integration means of different resources knowledge:

- Using of ontological approach to design knowledge bases;
- Step by step evolutionary design of system knowledge base;
- Modular design based on the libraries of typical reusable components;
- Same as in Semantic Web languages in OSTIS technology binary relation is preferred but there is a way to present relations of any arity;
- Relations are represented in the form of semantic network nodes that allow to define their attributes;
- Relation samples are highlighted as separate semantic network nodes that provide the possibility to define each relation sample in a unique way;
- There are elements of key nodes and arcs alphabet to describe fuzzy, negative and temporary objects;
- Semantic repository of OSTIS technology integrates the similar entities to the united network automatically;
- Connection of external thesauruses and OWL technologies (converter from XML models, PDF) is not hard. The proper converters are evolved as a part of OSTIS technology.
- The problem of knowledge based on homogeneous semantic networks including knowledge verification and critical errors repair is solved.

The main advantage of OSTIS technology is the flexibility of designing systems. OSTIS technology has already contained models, means, methods of intelligent system designing and a pack of subject domains ontology accumulated and formalized by this time [27]. That is why designing using OSTIS technology comes to its broad knowledge base designing.

III. Implementation of proposed approach

The essence of the ontological approach during designing of such systems, meaning personal integrators, centers around the consideration of system knowledge base as the hierarchy of marked subject domains and proper ontology. The following subject domains must be marked as a part of designing system:

- subject domain of information system;
- subject domain of integrator user;
- subject domain of user information space;
- subject domain of agent-translators.

The analysis of each present subject domain is analytically complex and consists of multiple abstractions, as the result of which the most considerable and relevant to current task objects, their attributes and mutual relations is marked from the whole variety of them. Knowledge of subject domain, understanding of present processes, rules and existing limitations are the necessary condition of flexible and effective information service integrator designing [5]. From the perspective of gain knowledge we have the possibility to determine designing ontology scale which provides sufficient level of ontology detailed elaboration required to solve the tasks of information service integration and following work with them. It also helps to mark concepts and relations which are necessary to include into. Taking into account that integrated systems can be absolutely different, elaborated subject domains ontology allows keeping declarative content of knowledge stored in them in spite of their syntax and stylistic distinction of their representation [4].

Implementation of the reviewed approach to create personal intelligent integrator of information services is the OSTIS-system. At this stage agent of some heterogeneous resources data collection are implemented. They provide receipt and transmission of the information to internal representation of knowledge base, provided their semantic integration. By virtue of the foregoing approach we have the possibility to integrate not only similar services but also services from the cardinally different subject domains, because of the united and common model of the user and personal information space. Let's examine the architecture of developed system and features implemented agents a little more detailed.

The common simplified architecture of the personal intelligent integrator.

At this stage personal information integrator provides the possibility of integration with the following services:

- Facebook
- Google Tasks
- Google Calendar
- Google Mail
- Todoist
- Dropbox

The first task to solve to design the intelligent integrator was machine understanding of integrated services. In other words the system should know which resources it's integrated with, how to work with them, which protocols and accessor methods are used, how often they needs synchronization, etc. This task was solved by means of formalization of external resources model and including of this model into system knowledge base when linking the external resource of data. In such a way information about external information resources became a part of integrated information user space that allowed its flexible and simple extension.

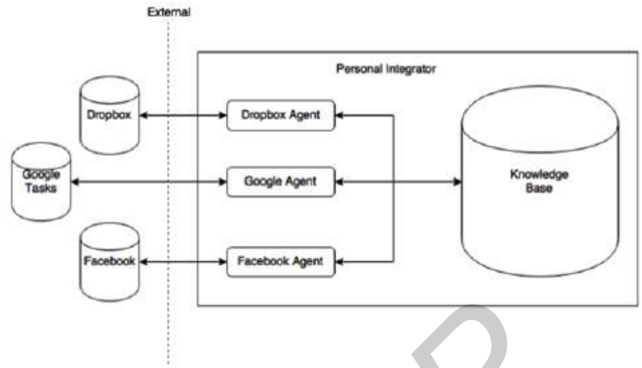

Figure 1. Simplified architecture of the personal intelligent integrator

Service formalized models formed so named user profile and are used to collect information about user. Collection of the information accomplished by specialized agents which acts as translators transforming the information received from the external resource to internal representation come from the predetermined template. The important aspect of heterogeneous resource semantic integration is the usage of knowledge of the service during collection of the information and formalization of the agent. Because of this the border between formal model and its technical implementation is removed. Also by virtue of semantic integration the user give the possibility to widen the usable services capabilities and self-adapt them. For this purpose he just has to write agent solving his specific task which won't depend on the resource of information and the service it needs to adapt because it will work with semantic integrated information. Therefore a user gives unlimited means to widen the functionality of integrated services which is inaccessible today.

Semantic integration of heterogeneous resources opens the possibility to solve the tasks which has been impossible to solve because they were on the junction of some services work. The example of such tasks can be meeting shift and its participants notification based on information about user geolocation, planning of training relying on user calendar and his medical constraints, book suggestion based on the list of read books and user current interests [7]. Therefore, information services, gained an access and had more detailed and all round user portrait, will get the possibility to improve provided services taking various aspects and each concrete case specificity into account. The user model should be formalized to uniform integration of information services and presentation of above-stated possibilities.

System user model will differ depending on concrete person interests as interests and requirements of each person are different, but the framework of the model for every person

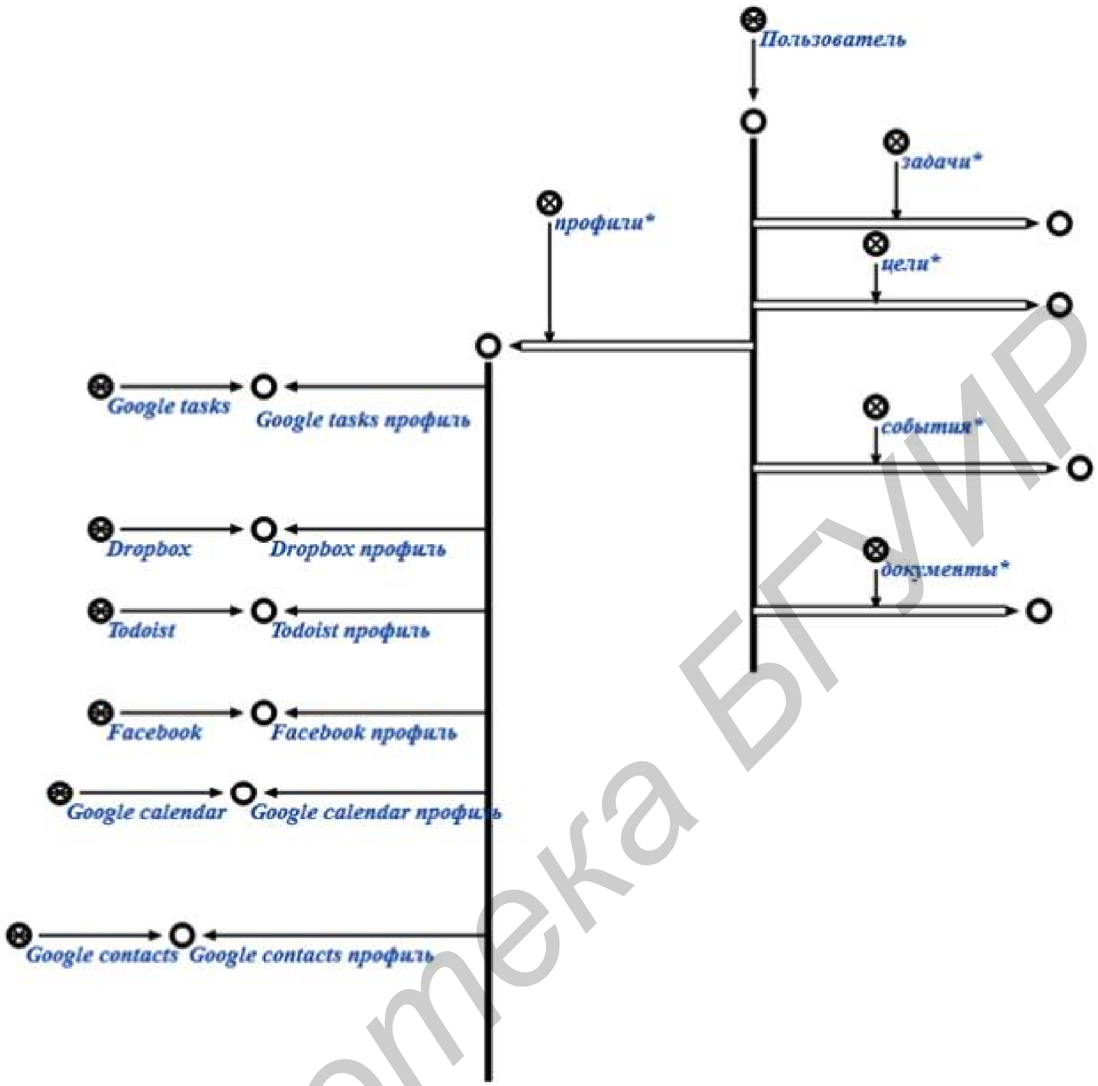

Figure 2. User profiles

will be the same. Depending on interests user will get the possibility to integrate usable service and knowledge base template which will conform to his subject domain. Therefore initial framework, common for every person, will be overgrown with more detailed models for each specific subject domain.

Current approach provides users to develop dynamically the semantic structure of their content which is made as "semantic halo" integrating information components into united semantic space. By virtue of that users will have the exact idea of the information surrounding them, possibility of semantic navigation through the system content and united universal interface to have an access to their information resources [13].

There is an example of user task model treated with the integration of some task services such as Google Tasks and Todoist on the diagram below. Each task received from these services is brought into united format by the agent-translator implemented for concrete service. Knowledge base contains the information where some information come from, so data refreshing will not break the consistence of integrated information and the user will have the possibility to work with this information not only within a designing system framework but also to continue manipulation of services integrated beforehand. Therefore the user isn't made to use designing system for solving any of his tasks but user gives an opportunity to choose and use the right service which let him solve his tasks fast and effective.

Let's examine the process of integration with service Google Tasks to understand the suggested architecture better.

The agent implementation begins with receipt of token which will thereafter be used to authenticate the application and to communicate with the service. The model of integrated service is created from the universal template and placed into the knowledge base. This model includes such information as access protocol, api version, accessible resources, api token, etc.

The example of such a formalization is below. When the service was added to the knowledge base, it became available for using by system agents.

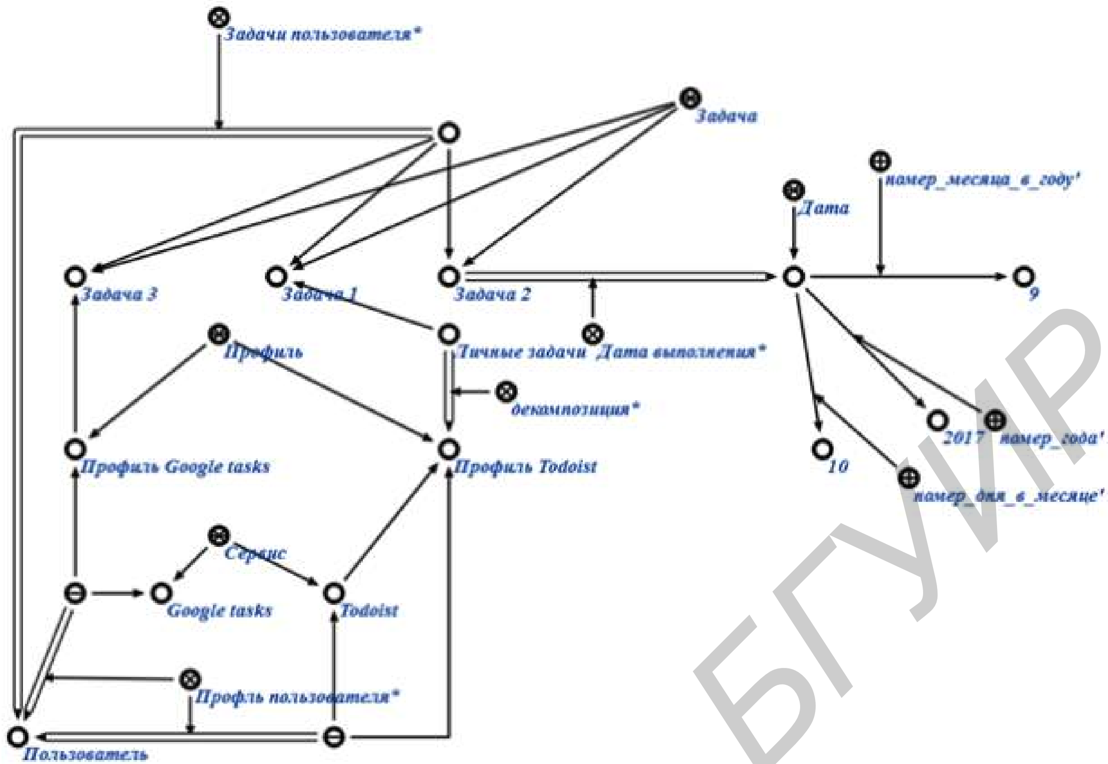

Figure 3. User tasks model

An agent is such a translator which receive the information about necessary service, find the necessary resource to maintain the operation, make requests and transform answers from JSON to the knowledge base content using the prepared templates. In case of conflicts, for example when some of the information will already exist in the knowledge base, it would be synced based on the unique identifiers. For resolving content conflicts a special interface should be designed to exclude chance of loosing important information and delegate (allow delegation) of making a decision to user.

The Google Task have the request of user tasks like this:

```
GET https://www.googleapis.com/tasks/v1
/users/@me/lists?key={API_KEY}

GET https://www.googleapis.com/tasks/v1
/lists/default/tasks?key={API_KEY}
```


```
{
 "kind": "tasks#tasks",
 "etag": "\"NEVtLf5Q_dTURZbE3G-zpPgGk\"",
 "items":
 [
      {
        "kind": "tasks#task",
        "id": "MDg0NDI4MTcwNTgwMjIyMNDk",
        "etag": "\"NEVtLf5Q_dTURZbE3E-
        zlPpPgGk\"",
        "title": "Task 1",
        "updated": "2016-07-25T17:36:41.
        000Z",
        "position": "00000000000001636797",
        "status": "completed",
        "due": "2016-06-16T00:00:00.000Z",
        "completed": "2016-06-16T21:35:26.
        000Z"
      },
      {
        "kind": "tasks#task",
        "id": "MDg0NDI4MTcwNTgwMjIyMNDk",
        "etag": "\"NEVtLf5Q_dTURZbE3E-
        zlPpPgGk\"",
        "title": "Task 2",
        "updated": "2016-06-16T21:35:27.
        000Z",
        "position": "00000000000001636798",
        "status": "completed",
        "due": "2016-06-16T00:00:00.000Z",
        "completed": "2016-06-16T21:35:27.
        000Z"
      }
    ]
}
```


This JSON format answer transform into temporary object which is brought into correlation with universal template, fill it and is preserved into the knowledge base. There is an opportunity to use libraries to simplify work with each concrete service. These libraries encapsulate process of communication with a service and provide it as a set of functions. In the case of Google API it's google-api-python-client package. Without regard for integrated service the integration process represent

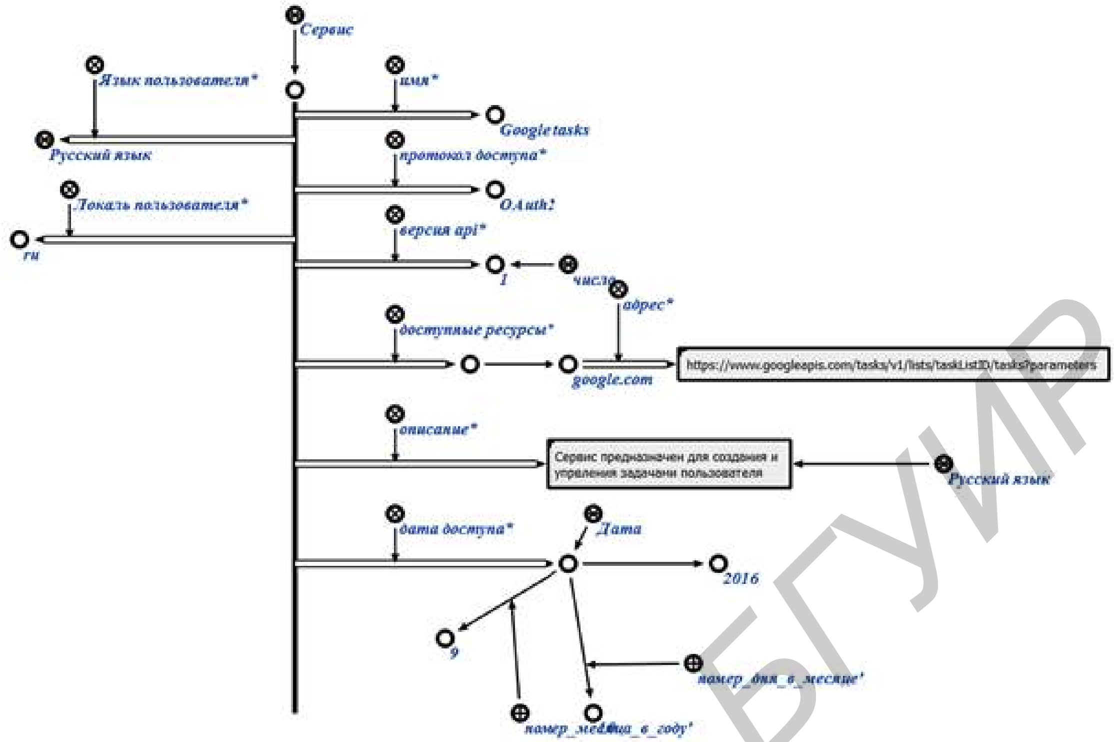

Figure 4. Web service representation

the following sequence of steps:

- registration of the application and receipt of the token to work with API service;
- addition of the service into integrator knowledge base;
- subject domain formalization;
- agent-translator elaboration.

In suggested implementation the problems of data semantic integration were solved by virtue of development of integrated services and their users' universal models. Thanks to using of semantic network as a model of representing knowledge about information service the configuration and adaptation in case of changing external resource model flexibility was reached. User universal model let us solve tasks on the junction of subject domains. Beforehand these tasks were inaccessible because of information distribution over user information space. The example of such task is the organization of meetings and events using information not only about participants' plans but also about participants' location, accessibility of places to hold events, traffic system state and weather conditions.

Also the important advantage of OSTIS technology is its modularity. All system components interact via united knowledge base that let us change system functionality without changing its common architecture. Addition of new user profile (integration with new information resource) doesn't influence on system work in general.

CONCLUSION

The complication of person information space structure leads to the necessity of development of new approaches to satisfy his information requirements. Information resource diversity leads to the problem of information perception fragmentariness. This problem brings to the appearance of new information system class – personal information integrators. In spite of existing set of personal assistants such as Siri, Google Assistant and Cortana the problem of semantic heterogeneity of personal information space wasn't solved.

On the base of OSTIS technology the system of personal information integrator was developed in which the problem of heterogeneous resources integration was solved. The design of information resources subject domain and user model ontology allowed to solve problem of data transfer between heterogeneous information services and to integrate received data on the semantic layer. Using the semantic networks approach the problem of unified information integration has been solved. According to this approach all collected information is stored inside a knowledge base of integrator system. OSTIS technology provides the united interface to work with knowledge base centered around SCg language. You can see this interface on 2-4 figures. Current implementation proposal has some problems, that just going to be solved and not deeply covered in scope of this paper. Designing the ontologies of each domain subject described in the paper is going to be improved to become a well-covered skeleton for future developments in this field. The user interface for personal integrator that would become efficient and domain subject independent is also a big task to do. To sum it all up the further development of personal assistant will be run into three directions: the increase of quantity of integrated information services, improving the ontology of domain subject and designing an approach for universal user interface.

REFERENCES

[1] V. Kashyap and A. Sheth. Semantic heterogeneity in global information systems: the role of metadata context and ontology.

[2] W. Kim: On resolving schematic heterogeneity in multidatabase systems. Intl. Journal on Distributed and Parallel Databases, Vol. 1:251-279, (1993)

[3] Wache H., and other. Ontology-Based Integration of Information. A Survey of Existing Approaches. Proceedings of the IJCAI-01 Workshop: Ontologies and Information Sharing, 2001. [Online]. Available: http://www.cs.vu.nl/ heiner/public/ois-2001.pdf

[4] Bianchini D., De V. Antonellis «Ontology-based Integration for Sharing Knowledge over the Web». [Online]. Available:http://www.doc.ic.ac.uk/pjm/diweb2004/DIWeb2004Part8.pdf

[5] Menzel, Christopher, Ontology Theory, in J. Euzenat, A. Gomez-Perez, N. Guarino, and H. Stuckenschmidt (eds.), Ontologies and Semantic Interoperability, CEUR Workshop Proceedings, 64 (2002). [Online]. Available: http://CEUR-WS.org/Vol-64/menzel.pdf.

[6] BotSpot. A Collection of Bots in the Web.[Online]. Available: http://www.botspot.com/

[7] Payne T. R., Singh R. and Sycara K. 2002. Calendar agents on the semantic web.

[8] G. Czibula, A.M. Guran, I.G. Czibula and G.S. Cojocar, IPA - An intelligent personal assistant agent for task performance support, in IEEE 5th International Conference on Intelligent Computer Communication and Processing. ICCP 2009, aug. 2009, pp. 31–34.

[9] Chen L. and Sycara K. 1998. WebMate: A personal agent for browsing and searching. In Sycara, K. P., and Wooldridge, M., eds., Proceedings of the 2nd International Conference on Autonomous Agents, 132–139. New York: ACM Press.

[10] M.H. Wang, C.S. Lee, K.L. Hsieh, C.Y. Hsu and C.C. Chang, Intelligent ontological multi-agent for healthy diet planning, in IEEE International Conference on Fuzzy Systems. FUZZ-IEEE 2009, aug. 2009, pp. 751–756.

[11] C. Ding, J.C. Patra, and F. C. Peng, Personalized web search with self-organizing map, in IEEE International Conference on e-Technology, e-Commerce and e- Service. EEE'05, mar. 2005, pp. 144–147

[12] A. Kiryakov, B. Popov, I. Terziev, D. Manov, and D. Ognyanoff, Semantic annotation, indexing, and retrieval, Web Semantics: Science, Services and Agents on the World Wide Web, vol. 2, no. 1, pp. 49–79, 2004.

[13] B. Brewington. Mobile agents for distributed information retrieval. Chapter 15 in [68].

[14] Pannu A. and Sycara, K. 1996. A learning personal agent for text filtering and notification.

[15] S. J. Russell and P. Norvig, Artificial Intelligence: A Modern Approach. Pearson Education, 2003.

[16] Berners-Lee, T. 2000. The semantic web vision.

[17] T. Berners-Lee, J. Hendler, and O. Lassila, The semantic web, Scientific American, vol. 284, no. 5, pp. 34–43, May 2001.

[18] Cheyer A. and Martin D. 2001. The Open Agent Architecture. Journal of Autonomous Agents and Multi-Agent Systems 4(1/2): 143–148.

[19] M. Klusch (Ed.). Intelligent Information Agents. (Springer, 1999)

[20] T. Mitchell. Machine Learning, (McGraw-Hill, 1997)

[21] G. Weiß (Ed.). Multiagent Systems. (MIT Press, 1999)

[22] G. Weiß (Ed.). Distributed Artificial Intelligence meets Machine Learning. Selected papers from ECAI-96 Workshop LDAIS and ICMAS-96 Workshop LIOME, LNCS 1221, (Springer, 1997)

[23] G. Weiß and S. Sen (Eds.). Adaptation and Learning in Multi-Agent Systems. Proc. of IJCAI-95 Workshop, CA, LNCS 1042, (Springer, 1995)

[24] Открытая семантическая технология проектирования интеллектуальных систем [Электронный ресурс]. – 2015. - Режим доступа: http://ostis.net. – Дата доступа: 15.12.2015.

[25] Базы знаний интеллектуальных систем / Т.А. Гаврилова, В.Ф. Хорошевский – СПб: Питер, 2000.

[26] Представление и обработка знаний в графодинамических ассоциативных машинах / В.В. Голенков [и др.]; под ред. В.В. Голенкова. – Минск : БГУИР, 2001.

[27] Интеграция знаний в информационных системах. / Н.А. Гулякина, В.П. Ивашенко // Доклады БГУИР. – 2004. – N6. – С. 113-119.

[28] Смирнов А.В., Пашкин М.П., Шилов Н.Г., Левашова Т.В. Онтологии в системах искусственного интеллекта: способы построения и организации // Новости искусственного интеллекта.

ПРИНЦИПЫ ПОСТРОЕНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ ИНТЕГРАТОРОВ ИНФОРМАЦИОННЫХ СЕРВИСОВ НА ОСНОВЕ ТЕХНОЛОГИИ OSTIS

Титенков П.В., Каешко А.И.

В статье рассматривается принципиально новый класс систем персональных информационных интеграторов, которые призваны решить проблему гетерогенности информационного пространства пользователя. Проблема гетерогенности информационного пространства выражается в необходимости использования нескольких сервисов для решения одной задачи, разрозненности личной информации и ее распределенности по нескольким серсвисам, затрудненных операциях поиска и во многом другом. В статье приводится анализ имеющихся интеллектуальных ассистентов, которые несмотря на впечатляющий набор возможностей, не интегрируют информационной сервисы на семантическом уровне, а лишь занимаются приведением естественно-языковых запросов к некоторому шаблонизируемому формату и делегированием их выполнения исходным сервисам. Рассматриваются технологии интеграции данных, приводятся достоинства и недостатки наиболее популярных подходов. Приводится анализ затрагиваемых предметных областей и соответствующих им онтологий. Рассмотрен пример реализации персонального интегратора нескольких информационных сервисов. Решение проблемы гетерогенности информационного пространства открывает широкие возможности для написания интеллектуальных агентов, которые обладая семантически целостной базой знаний, получат уникальный контекст для своей работы и смогут решать задачи находящиеся на стыке работы нескольких сервисов и не представляющие возможным их решение на сегодняшний день.