Chemical formula depiction using web service

***Pavel Horovčák*[1], *Dušan Dugáček*[2] *and* *Peter Cirbes*[1]**

Accurate minerals' chemical formulas depiction in websites requires an adequate working out of their representation. Within the latest approaches and technologies lies a SOA (Service Oriented Architecture), realized predominantly by WS (Web Service). The paper is devoted to SOA utilization in problem solution of chemical formulas depiction in internet environment by creation for this purpose intended web service. The application functionality is documented by its calling in testing website. In this way the SOA carries possibility of executive, effective, freely utilizable and interactive web applications creation.

***Key words**: SOA, Web Services, chemical formula, mineral, integration, PHP*

Introduction

Correct depiction of minerals' chemical forms on web sites requires solution for their portrayal. Indexing of quantifiers in chemical formulas is the most notable problem. In the information systems are chemical formulas usually placed in database table as text constants written in standard ASCII form. Similar problem occurs in other tasks and situations, which leads to depiction of chemical formula on web site. Structure of inorganic chemical forms is from the aspect of depiction relatively simple, because is enough to display quantifiers of amount of atoms or valence of ions as subscript or superscript text (upper indexes or lower indexes). Depiction of chemical formula in the form of picture (by means of <img> element) is one of the most used methods till now. This manner has several negatives (problems with font size changing; need to spend more effort to create graphical form and need to maintain relevant links to files that represents them). Another way to represent forms is markup language MathML (Mathematical Markup Language). This language was developed in 1998, is XML based and primary meant for inscription of formulas on web sites (Bos, 2008). Because web browsers developers don't pay attention to implementation of this standard, its application is problematic now (Design, 2008). MathML is supported by Amaya browser (court browser of W3C) and new versions of Mozilla (Firefox) in XHTML documents. In the most widespread browser MS Internet Explorer is support for MathML still missing (need to use MathPlayer plugin). Similar situation is with another markup language CML (Chemical Markup Language), which allows representation of chemical molecules structure. It is also XML based standard and for its interpretation is needed specialized browser (Murray-Rust, 1997, 2002). CML is integrated part of several development environments (OpenBabel, JMOL, JChemPaint and XDrawChem, available at http://www.sourceforge.net). For that reason as best suited solution seems usage of SOA (Service Oriented Architecture) that allows integration for that reason developed and created web service to client application (user's web site).

Web service for chemical formula depiction

Web service (WS) represents one of the main possibilities for service realization in SOA. SOA is "a loosely-coupled architecture designed to meet the business needs of the organization" (Evdemon, 2007). According to (Exforsys, 2008) there is no general agreed definition of SOA besides it is architecture that is based on orientation on services as main construct principle. What says definition of web service? For a number of definitions we mention two: a) Web service as a sum of new approaching standards that describes application architecture oriented on services that is based on components and build upon principle of architecture oriented on services SOA. (Champion et all, 2002); b) WS is software system identified by URI whose public interfaces and bindings are defined by XML and definition of this system can

[1] *doc. Ing. Pavel Horovčák, CSc., Ing. Peter Cirbes*, Technická univerzita, Boženy Němcovej 3, 040 00 Košice, Fakulta Baníctva, ekológie, riadenia a geotechnológií, Ústav riadenia a informatizácie výrobných procesov, Tel. (421 55) 602 5176, Pavel.Horovcak@tuke.sk

[2] *Ing. Dušan Dugaček*, Technická univerzita, Boženy Němcovej 3, 040 00 Košice, Fakulta Baníctva, ekológie, riadenia a geotechnológií, Ústav geovied

(Review a revised version 28. 1. 2010)

be discovered by other systems. These systems can collaborate with WS by manner that is described in its definition by means of XML messages transferred via internet protocols (Samtani, 2002).

WS provides standard description of its interface in form of WSDL (Web Services Description Language) that allows client to use the service. This description is generally deployed at different server than web service itself and presents some kind of connection between provider of service (server) and consumer (client) (fig. 1). SOA is however not (only) equal to WS, because each WS is indeed in compliance with SOA, but not each service within the frame of SOA has to be WS. At the present time, to the existing WS standards (XML, SOAP, WSDL and HTTP) new extending specifications were added (e.g. WS-ReliableMessaging, WS-Addressing, WS-Notification, WS-Security, WS-Policy and WS- Choreography) that complement use of WS and extend further syntax of WSDL.

For majority of mineral chemical formulas is enough to use upper and lower index. Both have adequate support in HTML language as well as in cascading styles. Task therefore rests in algorithmic solution of standard record conversion (number without sign means amount of atoms in the molecule and plus sign "+" after the number means valence of ion) stored e.g. in database table in ASCII form to the form supplemented with <sup> and <sub> elements to insure correct index depiction. Conversion process is ensured by stand-alone function which takes chemical formula as an argument and returns formula in supplemented HTML form. Algorithm processes chemical form character by character and analyses actual character as well as one preceding and two succeeding characters. Algorithm evaluates definition of the upper or lower index by incorporating <sup> or <sub> elements at analysis process round about actual character. After this manner adjusted formula is understandable (properly interpreted) by all browsers.

Web service for chemical formula conversion consist of two functions *getChemForm($formula)* and *getChemFormStr($formula)*, each with one argument and one return value – formula in HTML form. Function *getChemForm($formula)* returns value in standard XML form, function *getChemFormStr($formula)* returns value in the form of string. Principle of these two functions is conversion algorithm mentioned in preceding paragraph.

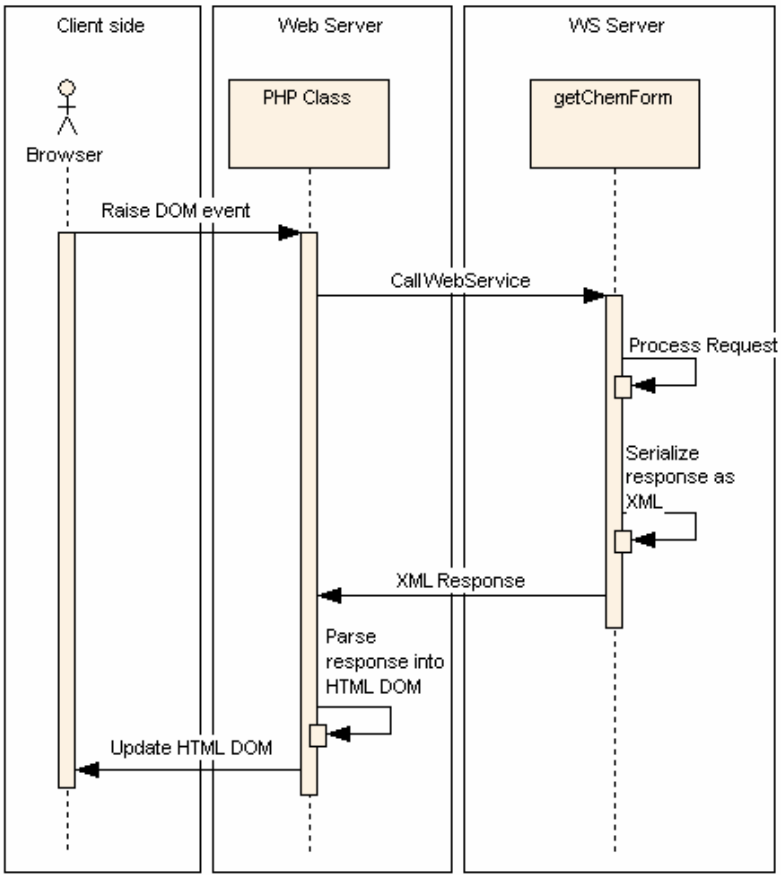

Fig. 1. Connection between service provider and consumer.

Mashup

A mashup (by Wikipedia web application hybrid) achieved in modern web creation a markedly progressive tendency. A mashup word meaning (mash – mixture, compound, and mixed bag) is transferred to web mashup meaning. „A mashup is a website or web application that uses content from more than one source to create a completely new service" (Next, 2007) and presents them in a new way (Shu-wai, 2008). In the last couple of years a "new breed" of consumers has emerged, enabling consumers to be aggregated and consumed by additional consumers. This "new breed" of consumers is usually referred to as a "mashup" (Evdemon, 2007). According this definition a mashup is a set of services, websites or applications that combine content from multiple resources into a new integrated user experience. Mashups are already becoming a mainstay of Web 2.0 (Plummer, 2006). Content used by mashups is typically sourced from a third party (such as a service or website) via a public interface or API (web services). Alternative methods

of sourcing content for mashups include newsfeeds and JavaScript constructs (JSON - JavaScript Object Notation, a lightweight computer data interchange format). There are three types of mashups, consumer mashups (the most common), data mashups, and business mashups.

Chemical formula depiction

Formula depiction (on client side) is solved as mashup that means it makes use of web service calling. The simple and effective possibility of WS calling besides other environments brings also PHP in version 5 with highly-developed SOAP (Simple Object Access Protocol) support in form of extension enabling simple service client creation based on WSDL (a standard method) as well as in non-WSDL mode (low-level access to WS). Interface to web service ensures WSDL file which specifies particular service's functions names as well as input and output parameters of these functions and their types. For web service calling by its consumer (client) it is necessary to know corresponding WSDL file therefore its URL. Principled schema of web service components communication illustrates figure 1. The web service calling on client side requires the service instantiation creation and subsequently calling of appropriate function in form:

```
$urlwsdl='http://lipko.tuke.sk/~horovcak/php_ws/wsdl/chemform.wsdl';
$ws = new SoapClient ($urlwsdl);
$result = html_entity_decode($ws->getChemForm($formula));
```

where *$formula* contains formula input string in an ASCII form and *$urlwsdl* is a service's WSDL file address.

For demonstration of created web service utilization was built-up a simple test page (http://www.ispavilon.tuke.sk/chemform test.php). This page enables entering of own chemical formula or selection one of prepared demo formulas. By button „Convert" clicking web service is called and realized formula conversion into the HTML form which is immediately after that displayed below in a box labeled „Result" (fig. 2).

Access statistics to test page chemformtest

At each access to web service test page chemformtest (http://www.ispavilon.sk/chemformtestr. php) is automatically created a record in script specifying this access, which consists of page language selection (SK-EN), time of page access (hh:mm), date of access (day.month.year), page operation mode (info – information about web service, test – own manual fill in of chemical formula into form input field, without indication – some of prepared chemical formula figures selection). This monitoring was in operation from 24.9.2008 to 10.11.2008, when IP address was appended. In respect of IP address absence in evidence there are on figure 3 (left) presented access numbers as "Entirely" and "Different". As different access we thought such one which access time difference to foregoing access time was greater then 1 minute.

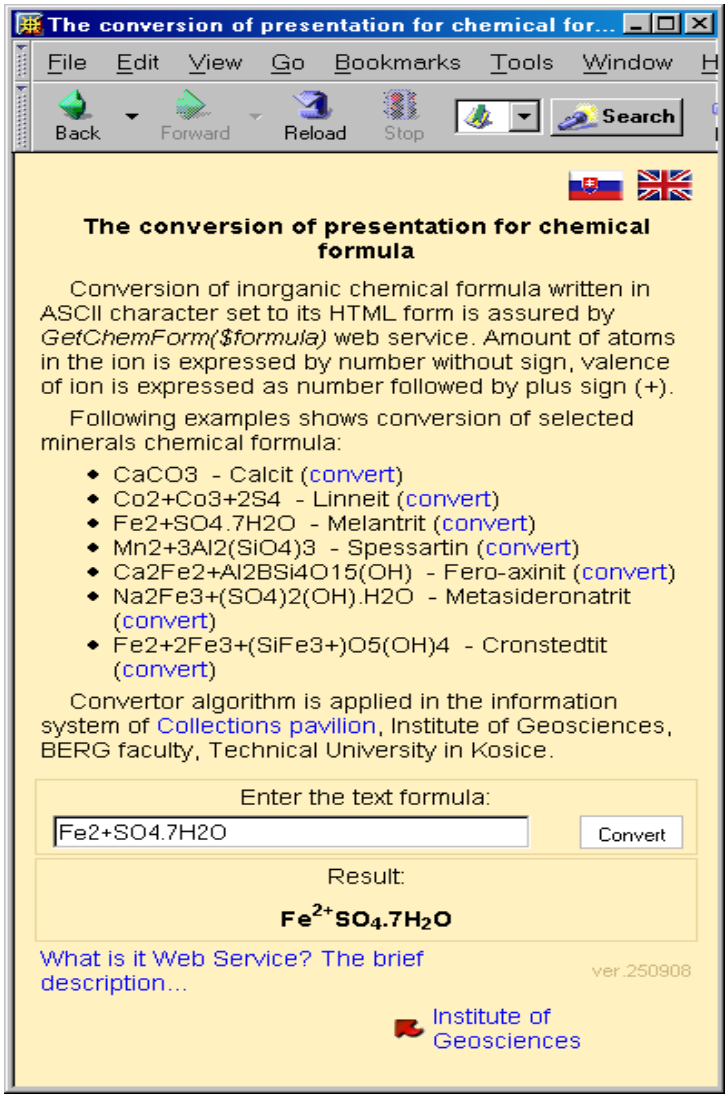

Fig. 2. Web service test page.

In the following time period will be this valuation based on IP address more precise. Figure 3 (right) illustrates access flow regarding page language selection.

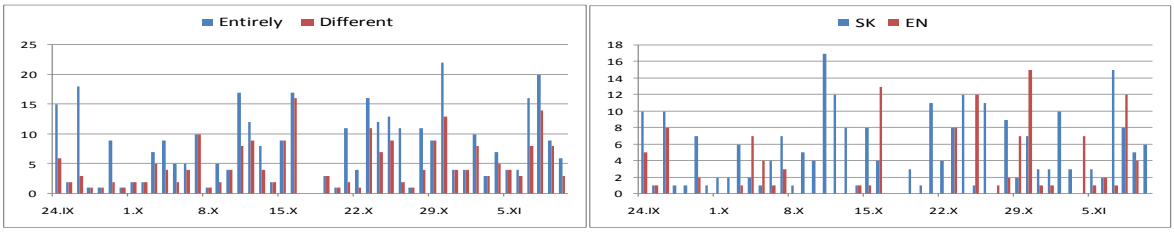

Fig. 3. Access page monitoring in first time period - entirely and different, and page language selection.

Within the time from 11.11.2008 to 9.12.2008 it was recorded besides already mentioned parameters also page visitor's IP address. Figure 4 illustrates page access number entirely and from different IP (left) as well as from page language selection point of view (right).The course is similar to figure 3 even though the accesses number is little lower.

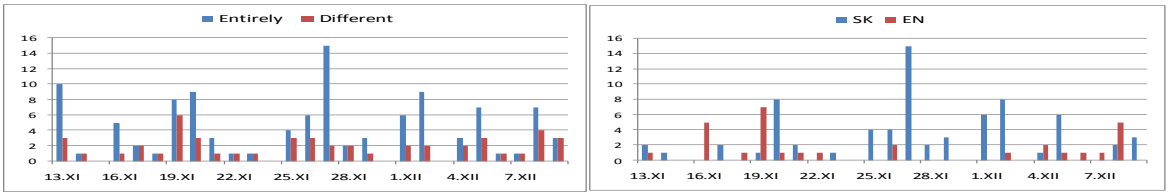

Fig. 4. Access page monitoring in following time - entirely and different, and page language selection.

Conclusion

Designed web service is intended for chemical formula conversion from standard ASCII form into the HTML form that presents the amount of atoms or valence of ions in form of subscript or superscript text (lower or upper indexes). Web service's WSDL file is located at above mentioned address ('http://lipko.tuke.sk/~horovcak/php_ws/wsdl/chemform.wsdl'). Functionality of created web service is demonstrated by means of test page available from URL http://www.ispavilon.tuke.sk/chemformtest.php. This test page enables selection one of prepared formulas or formula's direct entering into form's input element. On the server side there is required a PHP support in version 5 with activated SOAP extension (in the MS Windows environment php_soap.dll). WS can be a part of another web service or a building component of other web application. Web service utilization is real everywhere where is necessary to present chemical formula on web site, first of all if this formula comes from some external source (for example database table, xml or text file). By web services integration and utilization is coming up also certain modification and transformation of standard practices of web application design and development (Smutný, 2006).

Designed web service is disposable free of charge for the possible persons concerned.

Acknowledgement
The contribution was solved within the framework of projects VEGA 1/4194 /07 (L), VEGA 1/0194/08 (S) and VEGA 1/0365/08 (T) from the Slovak Grant Agency for Science.

References

Bos, B.: W3C Math Home, 2008[online] [cited 1.10.2008] available from http://www.w3.org/Math/

Burian, P.: Architektúra a počítanie orientované na služby a multiagentné systémy. In *AT&P Journal priemyselná automatizácia a informatika* č. 8, 2008, ročník 18, HMH s.r.o. Bratislava, ISSN 1335-2237, str. 53 – 57 [online] [cited 1.10.2008] Dostupné na http://www.atpjournal.sk/casopisy/atp_06/online.php?id=25

Design Science, MathPlayer. 2008. <Display MathML in your Browser> [online] [cited 1.10.2008] available from http://www.dessci.com/en/products/mathplayer/

Evdemon, J.: Service Oriented Architecture (SOA) in the Real World, 2007, [online] [cited 1.10.2008] Available from http://www.microsoft.com/downloads/details.aspx?FamilyID=cb2a8e49-bb3b-49b6-b296-a2dfbbe042d8&displaylang=en

EXFORSYS 2008. SOA Service Oriented Architecture [online] © 2008 Exforsys.com [cited 1.10.2008] available from http://www.exforsys.com/tutorials/soa.html

Champion, M., Ferris, Ch., Newcomer, E., Orchard, D.: Web Services Architecture. W3C Working Draft 14, [online] November 2002 [cited 1.10.2008] dostupné z http://www.w3.org/TR/2002/WD-ws-arch-20021114/

Murray-Rust, P.: Chemical Markup Language. 1997, [online] [cited 1.10.2008] available from http://www.xml.com/pub/a/w3j/s3.rustintro.html

Murray-Rust, P.: CML (Chemical Markup Language), 2002, Update. [online] [cited 1.10.2008] available from http://www.ccl.net/chemistry/resources/messages/2002/05/17.002-dir/index.html

Next-Generation Applications Using Ajax and OpenAjax - An OpenAjax Alliance White Paper - 7 February 2007 [online] [cited 1.10.2008] Available from http://www.openajax.org/whitepapers/Next-Generation%20Applications%20Using%20Ajax%20and%20OpenAjax.php

Plummer, D. et all: Gartner's Top Predictions for IT Organizations and Users, 2007 and Beyond. [online] 8.12.2006 [cited 1.10.2008] Available from http://www.gartner.com/DisplayDocument?id=498768

Samtani, G.:unjan: Top five Web service myths. 2002, Builder – Architect - WServices [online] Aug. 2002 [cited 1.10.2008] available from http://builder.com.com/article.jhtml;jsessionid=4VPOR43OSAE21TQQACQSFFA?id=u00320020820GXS01.htm&page=2

Shu-wai, Ch.: *Programujeme Mashup aplikace pro Web 2.0 v PHP*. Computer press, a.s. Brno 2008, ISBN 978-80-251-2057-6, 280 pp.

Smutný, P.: The Implementation Of Advanced Technologies In Intranet And Internet Web Portals. In *Proceedings of 7th International Carpathian Control Conference*. Czech Republic: Rožnov p. Radhoštěm. May 29-31, 2006, pp. 529-532. ISBN 80-248-1066-2

Štumpf, J: Proč a jak zavádět architektúru SOA, In *IT Systems*, CCB s.r.o Brno, ročník 9, 2007, ISSN 1212-4567, č.10, str. 30 – 32, č. 9, str. 12 - 15, č. 9, str. 46 - 48