

One-Stop Web Services 개발

조희형⁰, 안성수, 안부영, 박재홍, 박형선
 한국과학기술정보연구원
 {choh2⁰, ssahn, ahnyoung, middle75, seonpark}@kisti.re.kr

Development of One-Stop Web Services

HeeHyung Cho⁰, SungSoo Ahn, BuYoung Ahn, JaeHong Park, HyungSeon Park
 Korea Institute of Science and Technology Information

요약

근래 들어 생명과학 및 생명정보에 대한 관심이 높아지면서, 활발한 연구 활동의 결과물로 많은 정보를 얻을 수 있게 되었다. 또한 방대한 생명정보를 분석할 수 있는 많은 알고리즘이 개발되었으며, 웹 기술이 발달함에 따라 생명정보의 활용도가 높아져서 다양한 방법으로 정보서비스가 가능하게 되었다. 특히 XML, SOAP, WSDL, UDDI를 사용하는 Web Services는 분산 컴포넌트 기반의 컴퓨팅 기술로 최선의 기술이다. 이에, 바이오인포매틱스센터에서는 CBB, NCBI, EBI, DDBJ 등의 사이트에서 배포한 Web Services에 접근하여 입력 값과 출력된 결과를 확인 및 유지하고, 객체들의 흐름을 볼 수 있는 프로그램으로 One-Stop Web Services를 개발하였다. One-Stop Web Services는 GUI 환경의 인터페이스로 사용자에게는 용이한 실행 환경을 제공하므로, Web Services 개발자에게 사용자 인터페이스 개발 부담을 줄일 수 있다.

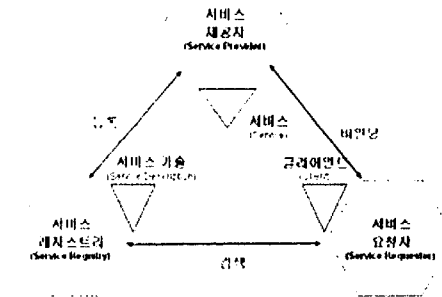
1. 서론

세계적으로 생명과학 및 생명정보에 대한 관심이 높아지고 있으며, 활발한 연구 활동으로 방대한 정보를 얻을 수 있었다. 특히 유전자 프로젝트 수행 후, 많은 유전 정보가 데이터베이스 되었으며, 유전정보를 분석할 수 있는 알고리즘이 개발되었다. 대표적인 데이터베이스로는 GenBank, PDB, dbEST, SWISS-PROT 등이 있으며, BLAST, FASTA, ClustalW, InterProScan 등과 같은 분석 도구를 활용한다. 최근 웹 분야 기술이 발달함에 따라 생물정보의 활용도가 높아지고 다양한 방법으로 정보를 서비스할 수 있게 되었다. 인터넷상에서 지리적으로 분산된 데이터베이스를 검색, 활용하는 기술로 Web Services가 있으며, 이 기술은 데이터 표현방식으로 XML를 사용하고, XML문서를 교환할 때 SOAP을 사용한다. 클라이언트에서 서버의 서비스를 호출할 수 있도록 서비스의 내용을 기술하는 WSDL을 사용하고, 서비스를 UDDI 저장소에 저장하고 검색하는 구조를 가지고 있다. 본 논문에서는 사용자가 Web Services를 활용하는데 손쉬운 방법을 제공하고자 GUI 환경의 응용프로그램을 개발하고, One-Stop Web Services라 한다.

2. Web Services

Web Services는 인터넷 어딘가에 자리 잡고 요청자에게 데이터와 서비스를 제공하는 응용프로그램 로직의 단위로 응용 프로그램들은 각각의 Web Services가 어떻게 구현되었는지 인지할 필요 없이 HTTP, SOAP과 같은 프로토콜로 Web Services에 접근한다. Web Services 이용은 웹사이트에 접속하는 것처럼 간단할 수도 있고, 다

중 조직적인 구조를 가지는 것처럼 복잡할 수도 있다. 현재 많이 사용하는 기술인 CORBA, CGI 스크립트, J2EE 등을 Web Services 기술로 분류할 수 있지만, 표준화된 XML을 사용하는 점에서는 다르다고 말할 수 있다. <그림 1>은 Web Services의 구성요소를 나타낸다.



UDDI 레지스트리

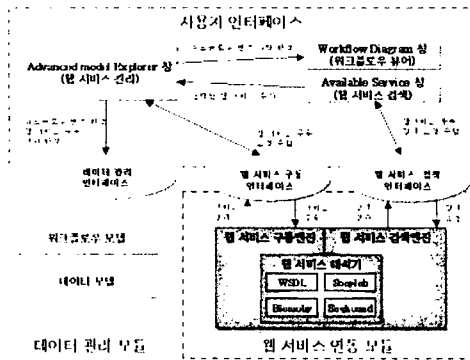
<그림 1> Web Services의 구성요소

3. One-Stop Web Services

바이오인포매틱스에서는 전 세계적으로 배포되어 있는 여러 가지 Web Services의 효율적인 관리를 통하여 사용자에 대한 서비스 강화 및 분산 구축되어 있는 생명정보 자원의 활용성을 높이고자 한다. CBB(Center for Computational Biology & Bioinformatics), NCBI, EBI, DDBJ 등의 사이트에서 배포한 Web Services에 접근하여 입력 값과 출력된 결과를 확인 및 유지하고, 객체들의 흐름을 볼 수 있는 프로그램의 개발이 절실한 실정이다.

3.1 시스템 개요

One-Stop Web Services 관리 프로그램은 Web Services 검색 및 구동을 담당하는 Web Services 연동 모듈, Web Services 워크플로우 및 Web Services 구동 결과 값을 관리하는 데이터 관리 모듈, 그리고 사용자가 Web Services 검색, 워크플로우 생성 작업을 할 수 있도록 지원하는 사용자 인터페이스로 구성되어 있다. <그림 2>는 One-Stop Web Service의 시스템구성도이다.



<그림 2> One-Stop Web Service의 시스템 구성도

3.2 시스템 구성

One-Stop Web Services 시스템은 다음과 같은 기능을 가지고 있다.

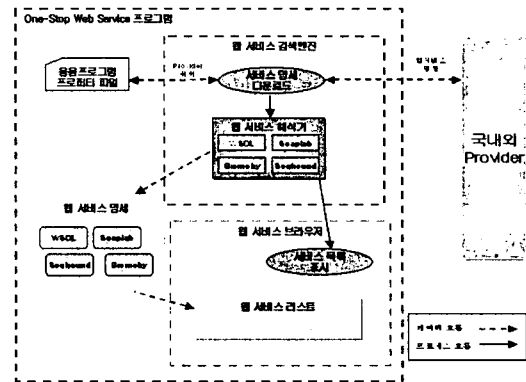
- Web Services 브라우저 - Web Services 검색엔진에 의해 검색된 Web Services 목록을 사용자가 조회/검색하고, 객체 흐름에 포함시킬 서비스를 선택할 수 있도록 지원하는 인터페이스이다.
- Web Services 제어판 - Web Services를 이용한 객체 흐름을 사용자가 생성/수정할 수 있는 인터페이스를 제공하고, 객체 흐름을 구성하는 Web Services에 대한 구동을 Web Services 연동 모듈에 요청하여, 객체 흐름의 결과 값을 얻어오도록 지원한다.
- 객체 흐름 뷰어 - 객체 흐름을 Graphical Diagram 으로 표시하고, Web Services 제어판에서 사용자에 의해 객체 흐름의 내용이 변경될 때마다 이를 Diagram에 반영하여 표시한다.
- Web Services 검색엔진 - 등록된 UDDI 서버에서 제공하는 Web Services 목록을 다운로드하고, 이를 Web Services 해석기에 의해 해석하여 Web Services브라우저에 전달한다.
- Web Services 구동엔진 - Web Services 서버에 서비스를 요청하고 응답을 기다리는 프로세스를 생성하고 실행시킨다. 프로세스 생성시에는 Web Services 해석기를 통해 해당 Web Services의 종류를 판별하여 이에 적합한 요청/응답 프로토콜을 사용하도록 한다.
- Web Services 해석기 - Web Services 검색/구동엔진에 의해 검색/구동이 일어날 때에 해당 Web Services의 명세를 해석하여, 그 결과를 검색/구동엔진에 전달한다.

- 객체 흐름 모델 - Web Services 제어판에서 사용자가 생성/수정하고 있는 객체 흐름을 저장하는 데이터 모델이다.
- 데이터 모델 - 객체 흐름을 구성하는 Web Services들이 구동 결과 값을 저장하는 데이터 모델이다.

3.3 One-Stop Web Services 기능

3.3.1 Web Services 검색

CCBB, NCBI, EBI, DBJ, SEQHOUND 등 국/내외 생명정보 Web Services Provider로부터 Web Services 명세를 다운로드 받고, 이를 각각의 Web Services 종류 (WSDL, Soaplab, Biomed, Seqhound)에 따라 해석하여 사용자가 확인할 수 있게 한다. <그림 3>은 Web Services 검색 기능 흐름도를 나타낸다.



<그림 3> Web Services 검색 기능 흐름도

Web Services 검색엔진에 의해 다운로드되어 해석된 Web Services 목록이 Provider 별로 트리구조로 표시되고, 트리구조로 표시된 Web Services 목록 중 사용자가 원하는 함수를 찾기 위한 기능이 있다. 각각의 Web Services 함수들은 함수명과 함께 해당 함수에 대한 설명을 표시하고, 사용자가 이를 선택하여 워크플로우에 해당 함수를 추가할 수 있다.

3.3.2 워크플로우 생성/수정

서비스 브라우저에 표시된 Web Services 목록에서 사용자가 원하는 서비스를 자유롭게 선택하여 등록/삭제함으로써 생명정보를 가공/처리할 수 있는 객체흐름을 생성/수정할 수 있는 기능을 제공한다. 워크플로우 데이터 모델은 Web Services 함수 및 각각의 함수를 구동하기 위한 입/출력 값 그리고 여러 Web services 함수들을 연결하여 데이터 객체들의 흐름을 나타내기 위한 링크정보 등으로 구성한다.

3.3.3 워크플로우 뷰어

Web Services 검색 창에서 사용자가 생성/수정 중인 워크플로우를 다이어그램으로 표시하며, 작업 중인 워크

플로우에 대한 변경사항을 반영하는 기능을 가지는 뷰어를 구현한다.

3.3.4 워크플로우 실행

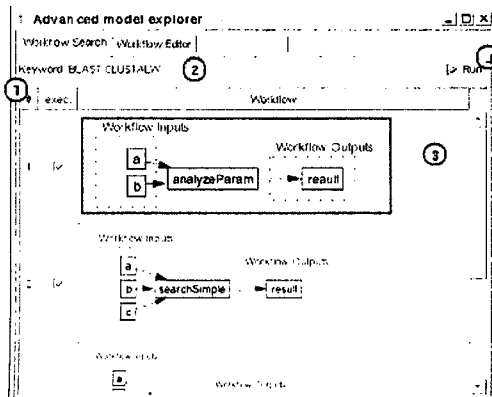
객체흐름을 구성하는 각각의 Web Services 함수들을 구동할 수 있는 프로세스들을 생성하여 Web Services로부터 결과를 얻어오는 기능으로 프로세스 생성 시에는 Web Services 해석기를 통하여 해당 Web Services의 종류(WSDL, Soaplab, etc)를 파악하고 적합한 프로토콜을 사용하는 프로세스를 하고, 각 프로세스들을 객체흐름 순서에 맞도록 수행시키는 Web Services 스케줄러에 의하여 프로세스를 실행한다.

3.3.5 외부 뷰어 연동

워크플로우 실행 결과 값 중 일부 Graphic 타입의 데이터는 외부 뷰어를 구동하여 해당 뷰어에서 내용을 확인할 수 있도록 한다. 향후 확장성 및 다양한 외부 뷰어를 지원하기 위하여, MicroSoft Windows 운영체제의 “연결프로그램정보”를 통하여 연동하도록 한다. 이를 위해서 본 응용 프로그램에서는 연동 대상인 데이터의 Mime Type 및 연동할 외부 뷰어와 “연결프로그램정보”에 의해 연관되어있는 확장자를 프로퍼티 파일에 등록하고, 결과 데이터가 해당 Mime Type에 속하는 경우 그에 해당하는 확장자를 적용하여 해당 데이터를 임시파일로 저장하고 Windows 운영체제에 이의 실행을 요청하는 방식을 사용한다.

3.3.6 워크플로우 라이브러리 검색 및 활용

컴퓨터에 친숙하지 않은 사용자의 경우 Web Services 검색에서 직접 워크플로우를 생성/수정하고 이를 실행시켜 결과를 얻어내는데 어려움을 겪을 것으로 예상된다. 그래서 워크플로우 라이브러리에서는 자주 사용될 것으로 예상되는 Web Services를 선정하여 워크플로우를 생성한 후, 라이브러리로 제공하므로 사용자가 쉽게 찾아 실행할 수 있도록 지원한다.



<그림4> 워크플로우 라이브러리 검색

워크플로우 데이터 모델의 구조를 그대로 따르는 문서 구조를 가지는 XML Element로 각각의 워크플로우를 표현하고, 이러한 Element를 여러개 포함하는 XML 문서의 형태로 워크플로우 라이브러리를 구성한다. <그림4>는 워크플로우 라이브러리를 검색한 결과이고 결과화면에 대한 설명은 다음과 같다.

- ① Web Services 검색의 Workflow Search에서 그림과 같이 라이브러리의 워크플로우 목록을 검색할 수 있다.
- ② Blast/Fasta/Clustalw 등 Web Services의 종류를 나타내는 단어를 입력한다. 각 단어들은 공백으로 구분되어 여러 개를 입력할 수 있으며, 샘플 워크플로우 중 해당되는 종류의 Web Services가 포함된 워크플로우가 검색된다.
- ③ 목록에 표시된 각각의 워크플로우를 선택하면, Workflow Editor와 Workflow Diagram 창에 해당 워크플로우가 반영되어 자세한 내용을 확인할 수 있다.
- ④ Run 버튼을 누르면 목록의 워크플로우 중 exec 체크박스가 선택된 워크플로우를 실행하게 된다.

4. 결론

Web Services는 분산 컴포넌트 기반의 컴퓨팅 기술로 산업 전반에 이용할 수 있다. 본 논문은 바이오인포매틱스 분야에서 제공하는 Web Services를 사용자가 쉽게 실행할 수 있도록 워크플로우 라이브러리를 개발하여 검색 및 활용할 수 있게 하였다. CCB, NCBI, EBI, DDBJ 등의 사이트에서 배포한 Web Services에 접근하여 입력값과 출력된 결과를 확인 및 유지하고, 객체들의 흐름을 볼 수 있는 One-Stop Web Services는 GUI 환경의 인터페이스로 사용자에게는 용이한 실행 환경을 제공하고 있으며, Web Services 개발자에게는 사용자 인터페이스 개발 부담을 줄일 수 있다.

참 고 문 헌

- [1] 안성수, 생물다양성데이터 검색포털 구축, 한국콘텐츠학회, 2005.
- [2] 이경하, 이규철, 웹 서비스의 향후 발전 방향, 한국정보처리학회지, 2002.
- [3] Paul Lipton, Composition and Management of Web Services, SYS-CON, 2004.
- [4] Wang.L.Riethoven.J.J and Robinson.A, XEMBL: distributing EMBL data in XML format, Bioinformatics, 2002.
- [5] Lee.j.M and Sonhammer.E.L, Genomic gene clustering analysis of pathways in eukaryotes, Genome Res, 2003.
- [6] Tom Oinn, Taverna:a tool for the composition and enactment of bioinformatics workflows, Bioinformatics, 2004.