

과학기술 지식인프라 통합 아키텍처 설계방법론

*¹강윤희, ²김영철

*¹백석대학교, yhkang@bu.ac.kr

²홍익대학교, bob@hongik.ac.kr

Design Methodology of Integrated Architecture for S&T Knowledge Infrastructure

¹Yun-hee Kang, ²Young-chul Kim

*¹Baekseok University, yhkang@bu.ac.kr

²Hongik University, bob@hongik.ac.kr

요약

최근 오픈 사이언스를 위한 플랫폼 개발이 국내외적으로 이루어지고 있으며, 과학기술 지식인프라의 데이터 처리 및 서비스 운영을 분석하여 과학기술 지식인프라의 성과확산을 위한 시스템 개발에 대한 요구가 증가하고 있다. 이 논문은 과학 기술 분야의 지식인프라를 위한 통합 아키텍처 설계 방법론을 도출하고 통합 아키텍처에서 데이터와 서비스의 통합수준 및 연계를 설정하는 그 방안을 제시한다. 연구 전과정에서 오픈 사이언스를 활용하기 위해서는 개별 서비스를 모듈 및 독립적인 기능 블록으로 구성하여 재사용 및 상호 운영될 수 있어야 한다. 이를 위해 제안된 설계 방법론은 과학기술 지식 인프라의 통합 서비스를 위한 서비스 통합을 다루기 위해 사용하며, 이는 다분야의 협력 참여연구자 및 엔지니어의 계산 수행 결과 및 산출물의 슬기없는 접근을 지원할 수 있다.

Abstract

Recently, the platforms for open science are being developed for helping researchers and scientists who are interested in convergence studies. This paper is focused on deriving a design method for an integrated architecture for knowledge infrastructure in science and technology and to set up the integration level of details and association between data and services in the integrated architecture. To facilitate the open science in research lifecycle, reconstructing individual services into modular and independent function blocks needs to be reusable and interoperable. The proposed design methodology is used for manipulating a service integration for the convergence service of S&T(Science and Technology) knowledge infrastructure. It can support to seamlessly access computing resources and the research results including artifacts digitally enabled scholars, researchers, and engineers participating in multidisciplinary collaborations.

Keywords: Platform, Open Science, Convergence, Design Method, Integrated Architecture, Knowledge Infrastructure

1. 서론

최근 오픈 사이언스를 위한 플랫폼 개발이 국내외적으로 이루어지고 있으며, 과학기술 지식인프라의 데이터 처리 및 서비스 운영을 분석하여 과학기술 지식인프라의

* Corresponding Author

Received: Aug. 27, 2018, Revised: Sept. 27, 2018, Accepted: Sept. 28, 2018

성과확산을 위한 시스템 개발에 대한 요구가 증가하고 있다[1-2]. 국내외의 공공데이터 공유 개방 흐름에 따라 데이터 기반의 다양한 비즈니스 기회가 창출되며 데이터를 효과적으로 관리하기 위한 필요 및 노력이 공공 분야의 과학기술분야에서 확산되고 있다. 국내 과학기술 분야에서도 연구자료를 공개, 공유하는 환경인 오픈사이언스의 흐름은 과학자 연구자 간의 지식과 정보를 개방 및 공유하고 협업하는 과정을 활성화하고 있으며 연구 전과정에서의 연구개방을 가능케 하기 위한 도구로서 ICT, 플랫폼 및 인프라 기술들이 결합된 플랫폼 개발을 요구하고 있다.

국내 과학기술 지식인프라는 과학기술 정보 수집, 가공 활용을 위한 포털의 형태로 제공되고 있으나 거버넌스를 위한 오픈 사이언스 아키텍처 적용을 통한 개선이 필요하다. 이를 위해서는 성공적인 시스템 통합 전략과 추진 방법을 제시되어야 하며, 각 조직이 시행착오를 겪지 않고 바람직한 통합이 방안이 제시되어야 한다. 이에 따라 지속가능한 과학기술 인프라 서비스 및 데이터의 융합서비스 지원 소프트웨어 아키텍처의 설계방법론 마련이 요구된다.

현행 S&T 지식정보 인프라는 개별 정보서비스를 통해 서비스 결과 전달이 사용자에게 제한적으로 제공됨으로써 정보섬(information island) 문제를 갖을 수 있으며, 연구활동 전주기 관리에 필요한 서비스 연동을 통한 신규지식 획득과 융합연구 지원의 제약점을 갖는다.

이 논문은 과학 기술 분야의 지식인프라를 위한 통합 아키텍처 설계를 방법론을 도출하고 통합 아키텍처에서 데이터와 서비스의 연계를 설정하는 그 방안을 제시한다. 연구 전과정에서 오픈 사이언스를 활용하기 위해서는 개별 서비스를 모듈 및 독립적인 기능 블록으로 구성하여 재사용 및 상호 운영될 수 있어야 한다. 이를 위해 제안된 설계 방법론은 과학기술 지식 인프라의 통합 서비스를 위한 서비스 통합을 다루기 위해 사용하며, 이는 다분야의 협력 참여연구자 및 엔지니어의 계산 수행 결과 및 산출물의 슬기없는 접근을 지원할 수 있다.

2. 관련연구

과학기술 분야의 오픈사이언스 환경을 위한 플랫폼은 연구자료를 공개, 공유하는 개방적 연구규범으로서 정의할 수 있으며 연구 전주기에 필요한 연구 아이디어 개발, 연구수행에 필요한 컴퓨팅 자원 제공, 연구수행에 따른 연구성과물에 대한 확산이 중요해지고 있다. 일 예로 미국 NIH에서는 공공자금 지원으로 수행한 연구프로젝트의 성과물을 온라인으로 공개하도록 한 Public Access Policy 가 시행된 이후, 오픈 액세스 형태로 공개된 논문의 비중이 증가했다 [3]. 2008 년 NIH 지원을 받은 연구논문의 34%만이 오픈 액세스 형태로 공개되었으나 7 년이 경과한 2015 년에는 그 비중이 두 배 이상 증가하여 약 71%의 NIH 연구논문이 오픈 액세스 형태로 공개하고 있다.

미국의 오픈 사이언스 데이터 클라우드(OSDC, Open Science Data Cloud)는 테라바이트나 페타바이트 규모의 과학데이터 집합을 저장, 공유, 분석할 수 있도록 과학 커뮤니티를 제공한다 [4]. OSDC 의 목적은 과학자들이 손쉽게 데이터를 공유하고 분석할 수 있도록 컴퓨팅 인프라스트럭처와 과학적 훈련을 통하여 핵심 데이터 집합을 제공함으로써 과학데이터의 발견에 병목현상 문제를 해결한다.

리서치 게이트(Research Gate)는 과학자와 연구자를 위한 사회관계망 사이트로 논문을 공유하고 질의응답하고 협력자를 발견하기 위한 목적으로 운영된다. 유럽과 북미에 대규모의 이용자 기반을 두고 있으며 다양한 분야의 연구자가 참여하고 있다.

디지털기술과 함께 가능해진 새로운 유형의 연구협력소통방식을 포함하는 과학계 전반의 개방화 추세로서 과학기술분야 연구 성과를 오픈액세스 정책을 기반으로 성과확산을 통한 연구수월성 제고가 가능하다. 연구데이터의 공개·공유 및 온라인 연구협력과 같은 개방형 연구협력을 지원하는 디지털 인프라 및 서비스 확대가 필요하다. 융합연구의 필요성이 증가됨에 따라 연구데이터의 공개·공유가 활발해지면서 이를 지원하는 디지털 인프라와 서비스, 커뮤니티가 중요해지고 있다.

이를 위한 분산형 데이터 플랫폼은 보다 다양한 연계활용을 중심으로 특정도메인에서 사용자들의 편익을 위한 분석형 서비스를 기획하거나, 도메인을 고려하지 않고 연구자 또는 개발자가 클라우드 환경에서 자유로이 대량의 데이터를 분석하고 목적에 맞는 학습 알고리즘을 구현할 수 있는 서비스를 개발할 수 있는 환경을 제공한다.

3. 통합 아키텍처 설계방법론

3.1 개요

과학기술 분야 사용자는 통합서비스 플랫폼을 통해 S&T 지식인프라의 다양한 융합서비스를 제공받기를 원한다. 이를 위해 컴퓨팅, 정보자원 및 협력환경을 제공하기 위한 아키텍처 설계가 요구된다. 본 절에서는 아키텍처 설계를 지원하기 위해 통합아키텍처 설계 방법론을 제안한다. 참조 아키텍처 모델은 통합 아키텍처 설계를 위한 전사의 모든 자원을 관점과 시각으로 분류한 매트릭스로 표현한 것이다. 참조 아키텍처 모델의 종류는 다양하며, 참조모델은 개념을 추상화한 아키텍처를 제공하며 구성요소 간의 인터페이스를 정의한다 [5]. 참조모델의 목적은 사용자 요구사항을 만족시킬 수 있도록 시스템 규격에 대한 개념적인 모델을 추상화하고, 구성요소를 재사용 가능한 컴포넌트로 생성하여, 여러 이해당사자들이 사용할 수 있도록 하는 것이다.

아키텍처 설계를 위한 참조아키텍처 모델은 전사의 모든 자원을 관점과 시각으로 분류한 매트릭스로 표현한 것이다. 아키텍처 모델의 종류는 여러 가지가 있다. 참조모델은 개념을 추상화한 아키텍처를 제공하며 구성요소간의 인터페이스를 정의한다. 참조 모델의 목적은 사용자 요구사항을 만족시킬 수 있도록 시스템 규격에 대한 개념적인 모델을 추상화하고, 구성요소를 재사용 가능한 컴포넌트로 생성하여, 여러 이해당사자들이 사용할 수 있도록 하는 것이다.

여기서는 다양한 시각 중 시스템 설계자(Designer) 관점에서 업무 응용을 운영할 환경을 구성하는 하드웨어, 소프트웨어, 네트워크의 기술기반을 중심으로 아키텍처 설계를 지원하도록 한다. 이를 통해 상위수준에서 정의된 조직기능 모델을 근거로 상세한 명세화 작업을 진행하도록 하며, 조직기능, 엔터티와 프로세스를 묘사하는 데이터 항목 및 기능들을 결정하는 시스템 분석가의 산출물과 대응하게 된다. 설계자 중심에서의 주요한 관점을 살펴되며 개발을 위한 도구를 매핑한다. 기술기반모델의 정의시에는 상호운영성을 위해 개방형 분산컴퓨팅 환경을 채용하여 API 서비스를 정의하며, 두 개의 체계가 외부간섭 없이 정보공유 및 교환이 가능하여 조작성이 이음매 없는(seamless) 체계 수행과 표현에 일관성이 있는 것을 제공하기 위한 통합성을 정의하여 관리한다.

3.2 통합서비스 논리 구성

융합 서비스는 S&T 지식인프라를 구성하는 기존시스템의 서비스를 사용하여 개발된다. 이 과정에서는 다음의 사용자 및 기술적 요구사항을 고려하여 융합 서비스를 설계한다.

- 사용자의 서비스 접근 용이성 및 맞춤형 서비스 제공
- 서비스 시스템 간의 협력을 위한 상호운영성을 포함한 연계 체계 마련[6,7]
- 협력에 참여하는 서비스 시스템의 데이터 및 서비스 수준통합을 위한 방법론 도출

단순히 시스템 통합을 통해 구성된 융합서비스는 추가되어진 서비스로 인해 전체 시스템 복잡도를 증가되게 되며, 시스템 내부와 시스템 간의 경계가 모호해지는 문제점을 가질 수 있다. 이를 위해 융합 서비스는 새로운 인터페이스 기술의 도입, 데이터운영방식 변경, 인터페이스 확장 또는 축소로 인한 변경을 체계적으로 관리할 수 있도록 통합 아키텍처를 설계하여야 한다.

융합서비스는 공통응용환경(CAE, Common Application Environment)를 통해 기존 서비스 기능을 제공받으며, 이를 위해 기본적으로는 비즈니스 규칙에 따라 기존의 응용과 연계할 수 있도록 하며 운영될 수 있는 API 를 활용하다. 외부서비스 연계는 도메인에서 정의한 표준에 따른 외부환경 인터페이스(External Environment Interface)를 정의하여 사용한다. Fig 1 은 S&T 지식 인프라를 위한 통합 아키텍처의 논리적 구조를 보인 것이다. 융합서비스는 통합 포털을 통해 제공한다. 융합서비스는 비즈니스 규칙에 의해 동작되며, 접속사용자 또는 사용자로 이루어진 가상조직에서는 서비스를 접근한다. 공통응용환경은 서비스 게이트웨이로 표현하며 내부 및 외부 서비스 연계 수행을 담당한다[8]. 서비스 통합을 담당하는 서비스 게이트웨이는 기존 서비스 전달을 위한 하부구조를 지원하며, 신규 사용자의 융합서비스 접근 시 균일한 인터페이스를 통한 공유 서비스로서 동작할 수 있는 접근 방식이 고려되어야 한다.

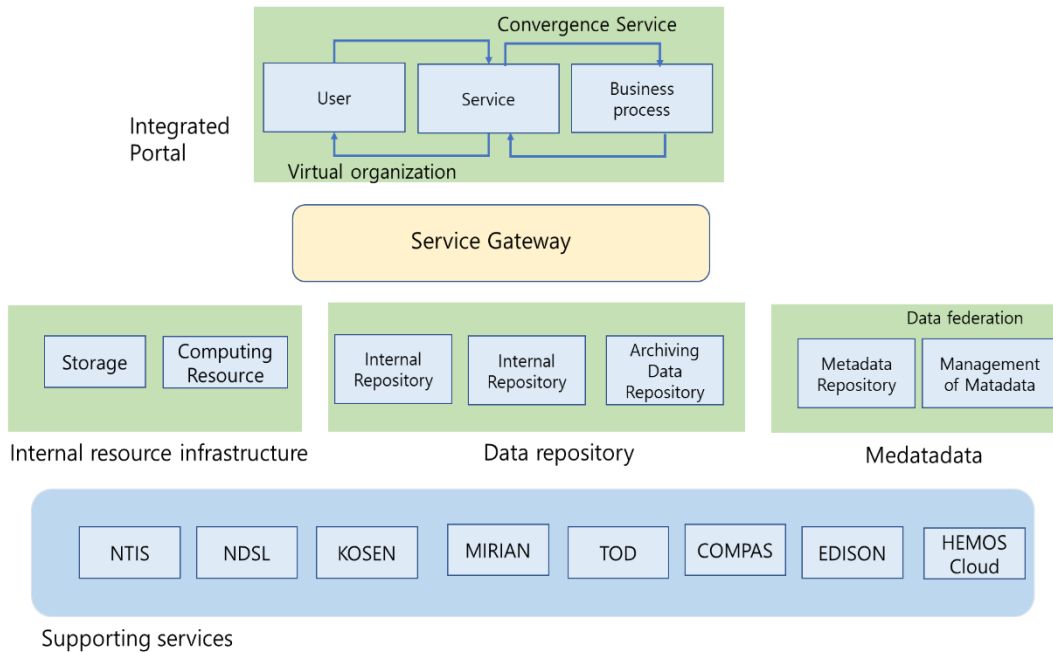


Figure 1. Logical structure of integrated architecture for S& T knowledge infrastructure

3.3 SOA 기반 통합서비스 아키텍처

SOA 는 사용자 애플리케이션의 기능을 서비스 형식으로 전달하는 분산 시스템 구축의 한 접근 방법으로, 느슨한 결합(Loosely Coupled)[9]을 통해 유연성을 높이고 표준 인터페이스를 기반으로 하여 이기종 환경에서의 비즈니스 민첩성을 높일 수 있다[10,12-15]. 서비스는 시스템 간의 통합을 위해 HTTP, SMTP 등의 인터넷 기반 전송 프로토콜을 사용함으로써 편재 요구사항(pervasive requirement)을 만족시킬 수 있으며, SOAP 또는 REST 과 같은 표준화된 프로토콜 및 XML 또는 JSON 확장 가능 메시지 포맷을 제공하므로 상호운용성 및 개방형 기능 제공을 위한 확장성 요구사항(extensibility requirement)을 만족한다.

SOA 기반 서비스통합은 비즈니스 로직의 융합서비스 구성을 위해 요구사항에 따라 서비스 단위로 융합서비스 기능을 수행하도록 설계한다. 설계된 SOA 기반 서비스는 서비스 재사용이 가능하도록 작성되기 때문에 품질(service quality)을 기반으로 동일 기능을 제공하는 서비스로의 대체가 용이하다. 서비스 컴포넌트 재사용 및 서비스 지속적

제공(service continuity)은 다른 서비스체계의 품질과 신규 융합서비스의 관리를 위해 서비스 합의(service agreement)을 활용한다 [11]. SOA 기반 서비스통합은 서비스 결합 및 매쉬업 기반의 논리적인 서비스 구성을 통해 새로운 융합서비스를 제공할 수 있으며, 서비스 운용을 위해 통합 서비스로 개발된 융합서비스의 연계 역시 서비스 인터페이스의 접속을 통해 이루어진다. Fig 2은 SOA 기반 서비스운영 모델을 보인 것이다.

상호운용성은 제약없이 시스템의 인터페이스가 이해하고 상호 협력하여 동작할 수 있도록 구현되거나 접근될 수 있는 시스템의 속성이다[6]. 이를 위해 상호운용성은 특정 목적을 위해 협력하는 모든 원소 간의 상당 수준 호환성의 정도(수준)가 존재하는 경우에 발생한다 [11]. 높은 상호운영성을 갖는 SOA 기반의 서비스통합 관리를 통한 융합서비스 구성의 특징은 Table 1과 같다.

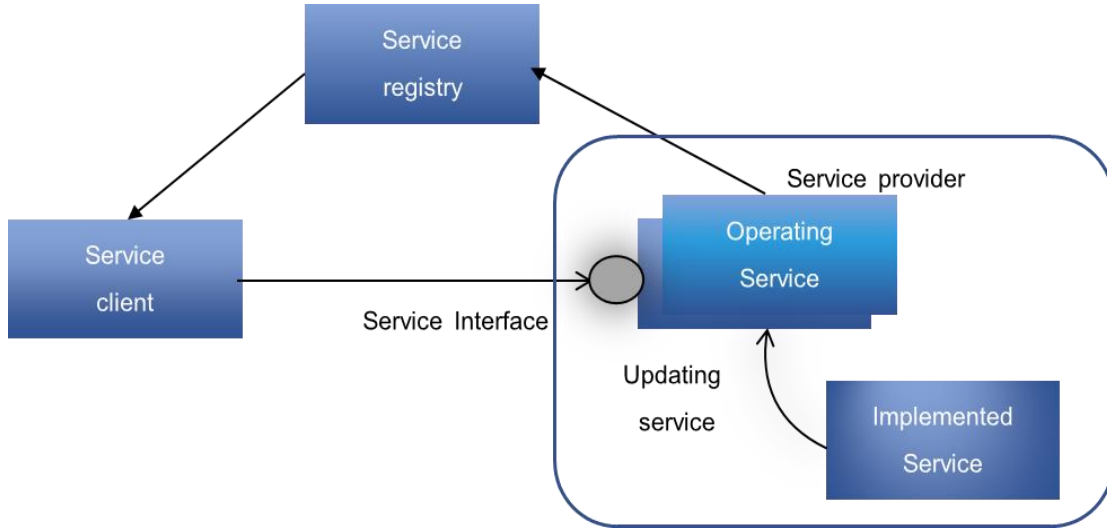


Figure 2. SOA based service operation mode

Table 1. The characteristics of SOA based service management

	Characteristics	Description
1	Service based integration environment	Provide logical integration by functional components based on established services
2	Standardized service	Provide interoperability with service operation with standard message format like SOAP or REST
3	Network based service integration	Provide a convergence service by interacting with scalability independence services

SOA의 실현을 통하여 얻을 수 있는 장점은 비즈니스 프로세스를 웹 서비스를 통해 공개함으로써 전체 S&T 인프라의 데이터를 통합하고, 내부 사용자, 과학기술 연구자, 및 일반사용자에게 유용한 정보를 제공할 수 있다는 점이다. Table 2은 SOA 기반 서비스 통합 체계의 주요 구성요소를 보인 것이다.

Table 2. SOA components of service integration system

	Component	Description of function
1	Service Client	<ul style="list-style-type: none"> ● Playing a role as service consumer ● Discovering and querying a service via service registry
2	Service Provider	<ul style="list-style-type: none"> ● Providing specific service ● Registering service information to service registry
3	Service Registry	<ul style="list-style-type: none"> ● Maintaining information about services provided by a service provider ● Returning information about services by requesting from a service client ● Monitoring a status of a service continuously

서비스 전달 모델에서 융합서비스는 소비자서비스로서 동작되는 것을 가정하며 서비스 전달은 제공자-클라이언트와의 동기 및 비동기 상호작용을 가진다. 서비스는 사용자 또는 이벤트에 의해 작동되며, 비즈니스 프로세스에 의해 생산(production) 정보, 정보소비(consumption) 교환과 데이터 변환(transformation)이 수행된다. 융합서비스의 비대칭 서비스 전달은 소비자 서비스의 요청에 따라 생산자 서비스가 서비스를 전달하는 방식으로 데이터 분석에 필요한 대용량 자료의 파일단위 또는 배치기반 처리에 따른 데이터 전달을 위한 환경에서 적용한다. 요대칭서비스 서비스 계약을 통한 비대칭 서비스가 동작될 수 있다.

4. 결론

과학기술 지식인프라의 데이터 처리 및 서비스 운영을 분석하여 과학기술 지식인프라의 성과 확산을 위해 오픈 사이언스 기반의 시스템 개발에 대한 요구가 증가하고 있다. 이 논문은 과학 기술 분야의 지식인프라를 위한 통합 아키텍처 설계를 방법론을 도출하고 통합 아키텍처에서 데이터와 서비스의 연계를 설정하는 그 방안을 제시하였다. 제시된 방안에서 융합서비스는 공통응용환경인 서비스 게이트웨이를 통해 기존 서비스 기능을 제공받으며, 이를 위해 기본적으로는 비즈니스 규칙에 따라 기존의 응용과 연계할 수 있도록 하며 운영될 수 있는 API 를 활용하다. SOA 기반 서비스통합은 비즈니스 로직의 융합서비스 구성을 위해 요구사항에 따라 서비스 단위로 융합서비스 기능을 수행하도록 설계하였다. 융합서비스의 엔터프라이즈 모델링을 위해 상호운영성을 지원하기 위해 논리 모델과 서비스 전달 모델을 설계하였다.

5. 감사의 글

본 논문은 한국과학기술정보연구원이 지원하는 2018 년도 위탁연구과제(No. P18014)의 연구수행으로 인한 결과물임을 밝힙니다.

6. 참고문헌

- [1] Jeffrey Robert Spies. 2013. The Open Science Framework: Improving Science by Making it Open and Accessible. Ph.D. Dissertation. University of Virginia, Charlottesville, VA, USA. Advisor(s) Brian Nosek. AAI3570855
- [2] Erin D. Foster, MSLS & Ariel Deardorff, MLIS. (2017). Open Science Framework (OSF). *Journal of the Medical Library Association*. 105. 10.5195/JMLA.2017.88.
- [3] National Institutes of Health, "Request for Information: NIH Public Access Policy", available at <https://publicaccess.nih.gov/comments.htm>.
- [4] Ryan Mork, Paul Martin, and Zhiming Zhao. 2015. Contemporary challenges for data-intensive scientific workflow management systems. In *Proceedings of the 10th Workshop on Workflows in Support of Large-Scale Science (WORKS '15)*. ACM, New York, NY, USA, , Article 4 , 11 pages. DOI: <https://doi.org/10.1145/2822332.2822336>
- [5] Philippe Kruchten. 1995. The 4+1 View Model of Architecture. *IEEE Softw.* 12, 6 (November 1995), 42-50. DOI: <https://doi.org/10.1109/52.469759>
- [6] IEEE, "IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries", Institute of Electrical and Electronics Engineers, 1990.
- [7] Sandra Heiler. 1995. Semantic interoperability. *ACM Comput. Surv.* 27, 2 (June 1995), 271-273. DOI=<http://dx.doi.org/10.1145/210376.210392>
- [8] Northrop L, Feiler P, Gabriel RP, Goodenough J, Linger R, Kazman R, et al. Ultra-large-scale systems-the software challenge of the future. Technical report Software Engineering Institute Carnegie Mellon University ISBN. 2006.
- [9] Eric A. Marks and Michael Bell. 2006. *Service-Oriented Architecture (Soa): A Planning and Implementation Guide for Business and Technology*. John Wiley & Sons, Inc., New York, NY, USA.
- [10] Organization for the Advancement of Structured Information Standards (2006), Reference Model for Service Oriented Architecture 1.0 , OASIS .
- [11] Morris, E.J., Levine, L., Meyers, B.C. and Plakosh, D., System of systems interoperability (SOSI): final report. 2004. Software Engineering Institute, Carnegie Mellon University.
- [12] [G. Alonso and F. Casati and H. Kuno and V. Machiraju., *Web Services Concepts, Architectures and Applications*, Springer Verlag, 2004
- [13] T. Erl., *Service-Oriented Architecture (SOA): Concepts, Technology, and Design*, Prentice Hall, 200
- [14] S. Chatterjee and J. Webber, *Developing Enterprise Web Services: An Architect's Guide*, Prentice Hall, 2004
- [15] E. Christensen, F. Curbera, G. Meredith, and S. Weerawarana. *Web Services Description Language (WSDL) 1.1*. W3C, 1.1 edition, March 2001.