
공공데이터를 활용한 국가정보화 전략연구 - 시나리오플래닝을 적용하여 -

이상윤* · 윤희주**

The Study on Strategy of National Information for Electronic Government
of S. Korea with Public Data
analysed by the Application of Scenario Planning

Sang-Yun Lee* · Hong-Joo Yoon**

요 약

최근의 웹에서 유비쿼터스로의 지식정보화사회의 급속한 진행에 의한 IT와 컴퓨팅기술에 있어서의 빅데이터 시대라는 새로운 패러다임 도래는 한국 정부 및 각국 정부에 있어, 전자정부 및 국가정보화 추진에 있어서의 주목할 만한 전환점이 되고 있다. 따라서 본 연구는 미래예측방법으로 많이 활용되고 있는 시나리오 플래닝 방법론을 적용하여 한국 전자정부 및 국가정보화 발전의 바람직한 미래상을 도출하였다. 곧 빅데이터 시대에 합당한 한국 전자정부 및 국가정보화의 상대적 미래우위전략을 찾고자, 정부의 빅데이터인 공공데이터의 활용에 대한 발전방안을 모색하였으며, 그 결과 한국의 전자정부 및 국가정보화에 부합하는 공공데이터를 활용한 대국민공개강화 전략을 도출하였다. 또한 ‘인간의 이해’보다는 ‘기계의 이해’를 지향하는 시맨틱웹 기술과 함께하는 링크드데이터 기술의 적용 아래에서의 개발을 제안하였다.

ABSTRACT

As a society of knowledge and information has been developed rapidly, because of changing from web environment to ubiquitous environment, a lot of countries across the world as well as S. Korea for national information with electronic Government have a variety of changes with big data. So this study is about development for national information and e-government of S. Korea with public data as big data analysed by the application of scenario planning. And then this research focused on the strategy consulting of national information with e-Government of S. Korea for utilization of public data as big data analysed by the application of ‘scenario planning’ as a foresight method. As a result, the future policy for utilization of public data as big data for national information with electronic government of S. Korea is to further spur the development of technology for linked data with semantic web for ‘understanding of machine’ than ‘understanding of man’.

키워드

National Information, e-Government, Scenario Planning, Public Data, Linked Data, Semantic Web
국가정보화, 전자정부, 시나리오플래닝, 공공데이터, 링크드데이터, 시맨틱웹

* 교신저자 : 부경대학교 공간정보시스템공학과(waw1313@hanmail.net)

** 부경대학교 공간정보시스템공학과(yoonhj@pknu.ac.kr)

접수일자 : 2012. 08. 28

심사(수정)일자 : 2012. 11. 21

게재확정일자 : 2012. 12. 10

1. 서론

1990년대 중반이후의 인터넷과 컴퓨팅 기술의 급속한 발전은 대량의 정보 확산을 불러왔다. 덕분에 우리는 이제 생산자로서 인터넷을 이용하여 직접 다양한 정보를 만들기도 하고, 소비자로서는 대량의 정보에 쉽게 접근하기도 한다. 나아가 최근에는 인터넷과 컴퓨팅 기술의 급속한 발전에 따른 모바일 인터넷과 소셜미디어의 등장으로 데이터가 경제적 자산이 되는 ‘빅데이터(Big data)시대’로까지 발전하고 있다[1].

정보 혹은 데이터의 이러한 급속한 증가는 이제 이를 처리하는 컴퓨터 기술과 시스템에도 영향을 끼쳐 기존의 단순한 개별적인 데이터 보관방식대신 클라우드 컴퓨팅(Cloud computing)으로의 변화를 불러오고 있다. 곧 단순히 정보를 전달하던 웹1.0(Web 1.0)에서, 다시 참여와 소통을 통해 집단지성을 창출하는 웹 2.0 방식으로 진화했고, 이제 데이터 분석을 통한 융합지식을 창출하는 웹3.0시대까지 진화하고 있는데, 이에는 클라우드 컴퓨팅방식이 합당하고, 여기서 스마트화·개인화·모바일화·실감화·지능형UI(User Interface)와 같은 빅데이터, 즉 공간, 시간, 관계, 세상을 담는 데이터 시대로 더욱 발전하기 때문이다.

이러한 빅데이터 시대에서는 지능형 유비쿼터스(Intelligent ubiquitous) 정보화를 추구하여, IT의 일상화 속에서 데이터가 정보사회를 움직이는 핵심연료가 되고, ICT주도권이 데이터로 이전되어 미래경쟁력과 가치창출의 원천역할을 하게 된다[1].

한국정부는 이상의 데이터 중심의 ICT사회로의 진입을 앞두고 2008년에 국가정보화기본계획을 발표하였다. 안에는 첨단 ICT기술을 활용한 전통산업의 활성화 필요성과 네트워크, 서버 등 ICT 산업 간의 불균형 문제해결, 수요자 중심의 고도화된 공공서비스를 창출한다는 새로운 정보화 정책패러다임의 필요성을 밝히고[2], 이에 대한 준비와 추진을 천명하였다. 또한 2011년에는 국가정보화전략위원회를 통해, ‘빅데이터를 활용한 스마트정부구현안’을 내고, 융합지식의 창출을 통한 정부운영의 효율화 및 투명성을 제고하여 맞춤형 국민서비스의 제공을 천명하였다.

한편 행정서비스 제공에 있어 신속정확한 자동화와 정보화실현이 전자정부의 목표인 점에서, 사실상 한국의 전자정부 추진 및 구축 역시 이러한 방향 아래에

서, 수요자 중심의 고도화된 공공서비스를 창출한다는 국가정보화 추진과 함께 발전했다고 할 수 있다[3]. 곧 현재 한국정부는 효율적으로 정보화를 추진한다는 필요에서 「전자정부법」 제50조(표준화) 및 동법시행령 제59조(표준화)에 의거하여 국가정보시스템의 상호 운용성 확보 및 효율적 개발 등을 위한 다양한 표준화 정책을 추진 중에 있다[4].

이에 본고는 한국의 국가정보화 추진과 함께하는 전자정부에 특히 중점을 두고, 한국정부의 빅데이터 시대에 적합한 공공데이터를 활용하는 국가정보화의 미래전략을 고찰한다. 그리고 그 분석을 위해 미래에 대한 정합성 있는 견해를 찾는 것으로 유용한 시나리오플래닝 기법을 적용하여, 향후 한국의 국가정보화 및 전자정부 구축의 전략방향을 고찰하고, 이를 위한 추진방안을 모색한다.

II. 한국 전자정부와 빅데이터

전자정부특별위원회의 전자정부백서에 나온 전자정부개념에 따르면, 전자정부는 행정서비스 체계를 일원화하고 공개함으로써 정부의 생산성과 투명성을 획기적으로 높일 수 있다는 목적 아래에서, 언제 어디서나 고객의 접근과 이용이 가능한 서비스형 정부를 말하는 것으로 다양한 행정서비스의 온라인화(化)를 통해 이를 실행한다[5].

한국정부는 2010년 UN에 의해 세계 최고수준의 전자정부를 구축한 것으로 평가받았다[3]. 표1[3]을 보면, 80년대 후반부터 90년대 중반까지는 선투자와 후정산, 탑다운 방식을 택하여 전자정부의 물적 토대를 확립하였고, 90년대 중반부터 2000년까지는 정보촉진기금을 통한 안정적 재원 확보 및 정부와 민간의 역할분담의 단계적 추진을 내세우고 부처별 핵심 업무의 정보화 및 부처 간 제한시스템 연계, 행정생산성 향상과 대국민서비스 수준상승을 목표로 했다. 2001년부터 2002년까지는 부처 간 시스템 연계, 기반조성 및 국가 핵심 업무의 통합서비스토대마련을, 2003년에서 2007년까지는 부처 간 시스템 연계확대, 행정효율성 및 투명성 향상과 국민주권실현을 추진배경으로 했으며, 가장 최근인 2008년 이후는 전자정부의 성숙기로서 ‘창의와 신뢰의 선진지식정보사회’를 새로운 국가

표 1. 한국 전자정부 발전과정별 추진동향
Table 1. Trend of electronic government of S. Korea

단계	전자정부 태동기	전자정부 기반조성/ 기반마련	전자정부 착수기	전자정부 본격추진기	성숙기
시기	80년대 후반 ~90년대 중반	90년대 중반 ~2000년	2001년~2002년	2003년~2007년	2008년 이후
주요 정책/ 계획 및 로드맵	국가기간전산망 구축	정보화촉진기본계획 (1996)/전자정부종합 실천계획(1999)/초고 속정보통신 기반구축 종합계획 (1995 ~ 2010)	전자정부11대과제 (2001)	전자정부로드맵 31대 과제(2003~2007)	국가정보화기본계획 (2008)/스마트 전자 정부추진계획(2011)
주요 법/ 비전	전산망이용촉진과 보 급 확장에 관한 법률 제정(1986)	정보화촉진기본법제 정(1995),전자서명법 제정(1999),SW산업 진흥법제정(2000)	전자정부법 제정 (2001),정보격차해소 법제정(2001),정보통 신 마보 호 법 개 정 (2001)/e-코리아비전 2006(2002년 수립)	전자정부법 개정 (2007)/브로드밴드IT 코리아비전2007(2003 년 수립), u-코리아 기본 계획 (2006 ~ 2010)	국가정보화기본법 (2009),전자정부법개 정(2010)/"창의와 신 뢰의 선진지식정보사 회"를 새로운 국가정 보화 비전으로 제시
전자정부단계/ 추진배경	전자정부 목적 토대 확립	부처별 핵심 업무 정 보화 및 부처 간 제 한시스템 연계/행정 생산성 향상과 대국 민서비스질 상승	부처 간 시스템 연계 기반조성 및 국가 핵 심 업무의 통합서비 스토대마련	부처 간 시스템 연계 확대/행정효율성 및 투명성 향상과 국민 주권실현	범정부적 통합 인프 라 구축/시스템 호환 성 상승과 공동 활용 통한 비용절감고조 및 성과창출, 스마트 폰 등 급속한 모바일 환경 변화에 따른 능 동적 대응
주요 전략	선투자와 후정산,타 다운 방식	정보촉진기금을 통한 안정적 재원 확보, 정부와 민간의 역할 분담 단계적 추진	선택과 집중의 원칙, 협력적 거버넌스 체계, 체계적인 점검전략	전자정부와 정부혁신 강조, 이용자 중심 성과관리, 정보산업 육성	선택과 집중 강조,녹 색성장건인,과제이행 체계정립,민관협업
추진체계 (부처/기구/기관)	체신부,총무처/전산 망조정위원회/한국전 산원(1987)	정보통신부,행정자치 부/정보화추진위원회 /한국전산원	정보통신부,행정자치 부/전자정부특별위원 회/한국전산원	행정자치부/전자정부 전문(특별)위원회/한 국정보화사회진흥원 (2006)	행정안전부/국가정보 화전략위원회/한국정 보화진흥원(2009)

정보화 비전으로 제시하며, 범정부적 통합 인프라 구축, 시스템 호환성 상승과 공동 활용을 통한 비용절감 고조 및 성과창출, 스마트폰 등 급속한 모바일 환경 변화에 따른 능동적 대응을 목표로 세웠다.

한편 빅데이터는 표2[6]와 그림1과 같이, 최근의 스마트시대의 IT핵심이슈 중 하나로서, 그 가공과 분석에 따라 상황인식, 문제해결, 미래전망이 가능해진다 는 점에서, 경제적 자산과 경쟁력의 척도로 부각되고 있는 것[1]이며, 사회영역에 활용되는 의미에서는 가치 있는 정보를 추출하기 위한 목적에서 대용량 데이터를 활용 및 분석하여 생성된 지식을 바탕으로 능동

적으로 대응하거나 변화를 예측하기 위한 정보화 기술을 의미한다[6].

최근의 SNS활성화와 스마트 단말의 확산, 사물네트워킹(M2M)의 확산은 데이터의 소위 폭발적인 증가를 더욱 활성화하여 빅데이터 기반은 향후 확대될 것으로 보이며, 그림에서 보듯이, 2020년에 이르면 제타바이트(ZB)시대가 본격화될 것으로 예상된다.

한편 한국정보화진흥원은 2011년 12월에 가트너의 발표를 참고하여 빅데이터의 4가지 구성요소를 정리하였다[6]. 곧 빅데이터에 대해, 먼저 기술적인 발전과 IT의 일상화가 진행되면서 해마다 디지털 정보량이

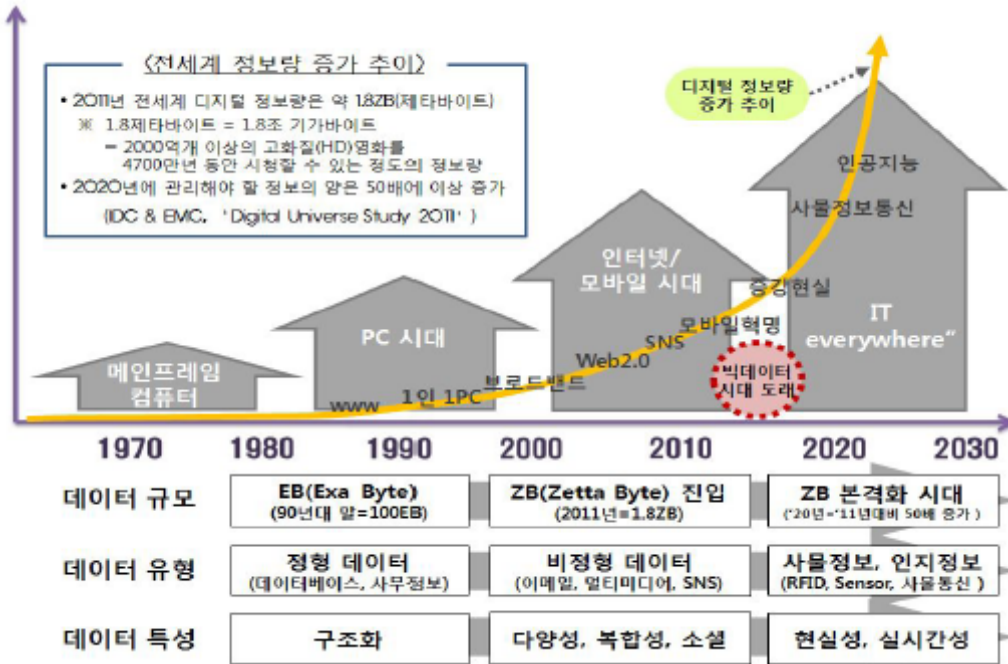


그림 1. ICT발전에 따른 데이터의 변화방향[6]
 Fig. 1 Change of ICT development and data

표 2. 정보사회의 패러다임 변화와 힘의 이동
 Table 2. Change of paradigm of information society

	PC 시대	인터넷시대	모바일 시대	스마트시대
패러다임 변화	디지털화/전산화	온라인화/정보화	소셜화/모바일화	지능화/개인화/사물정보화
IT 이슈	PC/PC통신/데이터베이스	초고속인터넷/WWW/웹서버	모바일 인터넷/스마트폰	빅데이터/차세대PC/사물네트워킹(M2M)
핵심 분야 (서비스)	PC/OS	포털/검색엔진/Web 2.0	스마트폰/앱서비스/SNS	미래전망/상황인식/개인맞춤형 서비스
대표 기업	MS/IBM 등	구글/네이버/유투브 등	애플/페이스북/트위터 등	?
IT 비전	1인 1PC	클릭 e-Korea	손안의 PC, 소통	IT everywhere, 신가치 창출

기하급수적으로 폭증하여 제타바이트시대로 진입하는 '규모(Volume)의 증가', 텍스트 이외의 멀티미디어 등 비정형화된 데이터 유형의 다양화와 로그기록, 소셜, 위치, 소비, 현실데이터 등의 데이터 종류가 증가하는 '다양성(Variety)의 증가', 데이터 종류의 확대나 외부 데이터의 활용으로 관리대상이 증가하며, 데이터 관리 및 처리의 복잡성이 심화되면서 새로운 기법이 요구

표 3. 정형화 정도에 따른 빅데이터의 데이터 종류
 Table 3. Sort of big data

구분	내용
정형 (Structured)	고정된 필드에 저장된 데이터로서 관계형 데이터베이스 및 스프레드시트 등을 들 수 있다.
반정형 (Semi-Structured)	고정된 필드에 저장되어 있지는 않지만, 메타데이터나 스키마등을 포함하는 데이터를 말하며 XML이나 HTML 텍스트 등이 그 예이다.
비정형 (Unstructured)	고정된 필드에 저장되어 있지 않은 데이터로서 텍스트 분석이 가능한 텍스트 문서 및 이미지/동영상/음성 데이터 등이 해당된다.

표 4. 빅데이터의 경제적인 가치전망
Table 4. Economic value of big data

기관명	주요 내용
Economist (2010)	- 데이터는 자본이나 노동력과 거의 동등한 레벨의 경제적 투입자본, 비즈니스의 새로운 원자재 역할 - 비즈니스 트렌드 파악, 질병예방, 범죄해결 등 효과
MIT Sloan (2010)	- 데이터 분석을 잘 활용하는 조직일수록 차별적 경쟁력을 갖추고 높은 성과창출
PwC(2010)	- 빅데이터는 이전까지는 다루지 못하고 시도하지 못했던 데이터의 활용을 가능하게 하며 잠재적 가치와 영향력이 높음 - 빅데이터의 중요성에 대해 기업들이 주목하고 있으며 새로운 비즈니스 가치창출의 핵심 키가 될 것
Gartner (2011)	- 데이터는 21세기 원유, 데이터가 미래 경쟁 우위를 좌우 - 기업은 다가올 '데이터 경제 시대'를 이해하고 정보고립을 경계해야 성공가능 - 빅데이터는 향후 주목해야 할 이머징 기술(2-5년 후 성숙)
McKinsey (2011)	- 글로벌 비즈니스 지형을 뒤바꿀 기술 트렌드의 3가지 핵심은 '클라우드', '빅데이터', '스마트 자산' - 빅데이터는 혁신, 경쟁력, 생산성의 핵심요소 - 의료, 공공행정 등 5대 분야에서 6천억 불 이상 가치창출

되고, 구조화되지 않은 데이터 및 데이터 저장방식의 차이와 중복성 문제가 발생하는 '복잡성(Complexity) 증가', 실시간성으로 인한 데이터생성 및 이동(유통) 속도의 증가와 사물정보 및 스트리밍 정보 등의 실시간 정보가 증가하고, 대규모 데이터 처리 및 가치

있는 실시간 현재정보활용을 위해 데이터 처리 및 분석속도가 중요한 '속도(Velocity)증가'가 그 주요요소였다.

표3은 정형화 정도에 따른 빅데이터의 데이터 종류를 나타낸 것이다. 먼저 정형(Structured)데이터는

표 5. 빅데이터의 특성과 효과
Table 5. Characteristic of big data

빅 데이터의 특성	효과
대규모 (Huge Scale)	- 기술발전으로 데이터를 수집, 저장, 처리능력 향상 - 현실세계 데이터를 기반으로 한 정교한 패턴분석 가능 - 데이터가 많을수록 유용한 데이터, 전혀 새로운 패턴의 정보를 찾아낼 수 있는 확률도 증가
현실성 (Reality)	- 우리사회 일상에서의 데이터 기록물의 증가 등 현실정보, 실시간 정보의 축적이 급증될 전망 - 개인의 경험, 인식, 선호 등 인지적인 정보유통 증가
시계열성 (Trend)	- 현 시점 뿐만이 아니라 과거 데이터의 유지로 시계열적인 연속성을 갖는 데이터의 구성 - 과거, 현재, 미래 등 시간 흐름상의 추세 분석 가능
결합성 (Combination)	- 의료, 범죄, 환경, 안보 등 타분야, 이종 데이터간의 결합으로 새로운 의미의 정보발견 - 실제 물리적인 결합 이전에 데이터의 결합을 통한 사전 시뮬레이션, 안전성 검증 분야 발전가능

표 6. 각국의 빅데이터 관련 전략과 추진방향
Table 6. Strategy of big data of Nations

기관명	주요 내용
대한민국, 국가정보화전략위원회 (2011)	<ul style="list-style-type: none"> - ‘빅데이터를 활용한 스마트정부구현(안)’제시(‘11.11) - 빅데이터 시대에 대비하여 융합지식과 분석행정을 통한 스마트 정부 실현을 위한 세부과제 도출
미국 대통령 과학기술자문위(2010)	<ul style="list-style-type: none"> - ‘Designing a Digital Future’ 보고서 발표(‘10.12) - 대규모 데이터의 가치에 주목하고 데이터 취합 및 관리, 분석의 중요성 제시 - 미국 정부기관들이 데이터를 지식으로, 지식을 행동으로 변환하는 전략에 집중해야 함을 주장
싱가포르 RAHS 시스템	<ul style="list-style-type: none"> - 데이터를 기반으로 싱가포르를 위협하는 리스크에 대한 평가와 환경 변화를 탐지하는 국가위험관리 시스템(RAHS) 구축 - 호라이즌 스캐닝 시스템 : 데이터 수집, 분류, 분석, 관계성, 예측지원
OECD	<ul style="list-style-type: none"> - 빅데이터를 비즈니스 효율성을 제공하는 새로운 자산으로 인식 - 제15차 WPIIS회의에서 빅데이터의 경제학 측정을 의제로 채택 : 무형 자산인 빅데이터의 경제적 가치를 실증하고 정책적 함의 제공이 목적

관계형 데이터베이스 및 스프레드시트 같은 것으로 고정된 필드에 저장된 데이터를 말한다. 다음 반정형(Semi-Structured)데이터는 메타데이터나 스키마 등을 포함하는 데이터로서 고정된 필드에 저장되어 있지 않은 것이며, 그 예는 XML이나 HTML 텍스트 등이라 할 수 있다. 비정형(Unstructured)데이터는 텍스트 분석이 가능한 텍스트 문서 및 이미지/동영상/음성 데이터 등이 그 예인데, 고정된 필드에 저장되어 있지 않은 데이터를 의미한다.

한편 빅데이터 처리는 기존 데이터 처리와 다음과 같은 차이가 있다. 첫째, 빅데이터는 빠른 의사결정이 상대적으로 덜 요구된다는 것인데, 즉각적인 처리속도가 가능한 기존의 데이터 처리와 달리 대용량 데이터에 기반을 둔 분석위주로, 장기적/전략적 접근이 필요하기 때문이다. 둘째, 빅데이터는 처리(Processing) 복잡도가 높고, 이러한 이유로 대체로 분산 처리 기술이 필요하다. 그 이유는 다양한 데이터 소스, 복잡한 로직 처리, 대용량 데이터 처리를 하기 때문이다. 셋째, 빅데이터는 처리의 복잡성을 증대시키는 요인인 비정형 데이터의 비중이 높는데, 소셜 미디어 데이터, 로그 파일, 클릭스트림 데이터, 콜 센터 로그, 통신 CDR 로그 등이 대부분 비정형 데이터기 때문이다. 넷째, 빅데이터는 처리 및 분석의 유연

성이 높은 편으로서 이는 새롭고 다양한 처리방법의 수용을 위해서 필수적이기 때문이다. 다섯째, 대용량 및 복잡한 처리가 특징인 점에서, 동시에 처리가 필요한 데이터량이 기존의 데이터처리에 비해 낫다는 것이다.

표4[6]는 빅데이터의 경제적인 가치에 대한 전망이다. 가트너는 21세기는 원유와 데이터가 미래 경쟁우위를 좌우한다고 하며, 빅데이터의 가치에 대해 향후 추구해야 할 이머징기술로서 기업은 다가올 ‘데이터 경제시대’를 이해하고 정보고립을 경계해야

성공이 가능하다고 하였고, 맥킨지는 빅데이터가 혁신, 경쟁력,생산성의 핵심요소로서 의료와 공공행정 등의 6대 분야에서 6천억 불 이상의 가치를 창출할 것이며, 미래의 글로벌 비즈니스 지형을 뒤바꿀 3가지 기술트렌드로 클라우드, 스마트자산, 그리고 빅데이터라 하였다.

표5[6]는 빅데이터의 특성과 효과이다. 먼저 특성으로는 대규모(huge scale), 현실성(reality), 시계열성(trend), 결합성(combination)이 있는데, 각각의 효과를 보면, 대규모(huge scale) 효과는 데이터가 많을수록 유용한 데이터가 많을 가능성이 높으며, 따라서 전혀 다른 새로운 패턴의 정보를 찾아낼 수 있는 확률도 높다는 것이다. 현실성(reality)효과는 개인의 경

표 7. 빅데이터와 연계된 기술들
Table 7. Technology with big data

용어	설명
빅테이블 (Big Table)	구글 파일시스템 상에 구축된 상용 분산 데이터베이스시스템으로서 H베이스에 영향을 미치는 것
카산드라 (Cassandra)	분산시스템에서 방대한 분량의 데이터를 처리할 수 있도록 디자인된 오픈소스 데이터베이스 관리시스템
데이터웨어하우스 및 분석 어플라이언스	데이터웨어하우스를 위해 서버, 스토리지, 운영체제, 데이터베이스, BI, 데이터마이닝 등 기타 여러 가지 소프트웨어가 최적화되어 설치된 통합제품
분산시스템	네트워크로 연결된 컴퓨터들의 집합으로서 동시에 일을 처리하기 위한 목적에서 단일 혹은 다수의 컴퓨터 리소스를 부분적으로 활용함으로써 시스템의 가격대비 성능비, 안정성, 확장성을 향상시킬 수 있는 것
하둡 (Hadoop)	구글이 개발한 맵리듀스를 오픈소스로 구현한 결과물인데, 야후에서 처음 개발되었고 현재 야콥치 소프트웨어 재단에서 프로젝트의 하나로 관리되고 있으며, 분산시스템 상에서 대용량 데이터처리분석을 지원하는 오픈소스소프트웨어 프레임워크를 말하는 것
H베이스 (Hbase)	구글의 '빅테이블'을 참고로 개발된 오픈소스 분산 비관계형 데이터베이스
맵리듀스 (MapReduce)	분산시스템 상에서 대용량 데이터 세트를 처리하기 위해서 구글이 제안한 소프트웨어 프레임워크
비관계형 데이터베이스	관계형 데이터베이스와 대조적인 개념으로서 데이터를 테이블에 저장하지 않는 데이터를 의미
구글 파일 시스템	하둡과 관련이 있는 것으로 구글에서 개발한 분산파일시스템

협이나 인식, 선호 등의 인지적인 정보유통이 증가하면서 일상의 데이터 기록물과 같은 현실정보가 늘어나 실시간 정보가 축적될 전망이 급증한다는 것이다. 시계열성(trend)효과는 과거 데이터가 유지되면서 시계열적인 연속성을 갖는 데이터가 구성되어 과거/현재/미래 등의 시간흐름상 분석이 가능해진다는 것이며, 결합성(combination)효과는 이종 데이터간의 결합으로 새로운 의미의 정보가 발견된다는 것이며 나아가 실제 물리적인 결합을 해보지 않고서도 데이터를 통한 사전시뮬레이션을 통해 안전성 검증분야의 발전이 가능하다는 것이다.

표6[6]은 각국의 빅데이터 관련 전략과 추진방향을 나타낸 것이다. 먼저 한국은 2011년에 대한민국 국가정보화전략위원회를 통해 '빅데이터를 활용한 스마트 정부구현(안)'을 제시하고 빅데이터시대에 대비하여 융합지식과 분석행정을 통한 스마트정부실현을 위한

세부과제를 도출하였다. 미국은 2010년 대통령과학기술자문위를 통해 'Designing a Digital Future' 보고서를 발표하고 미국정부기관들이 데이터를 지식으로, 지식을 행동으로 변환하는 전략에 집중해야 함을 강조하였다. 싱가포르는 데이터에 기반을 두고 싱가포르를 위협하는 리스크에 대한 평가와 환경변화를 탐지하는 국가위험관리시스템(RAHS)을 구축하였다. 한편 OECD는 빅데이터를 비즈니스 효율성을 제공하는 새로운 자산으로 인식하고 무형자산인 빅데이터에 대한 경제적 가치를 실증하여 정책적 함의를 제공하고자 했다.

III. 공공데이터기술과 전자정부 전략방향 설정

공공데이터(public data)는 정부 또는 공공기관이

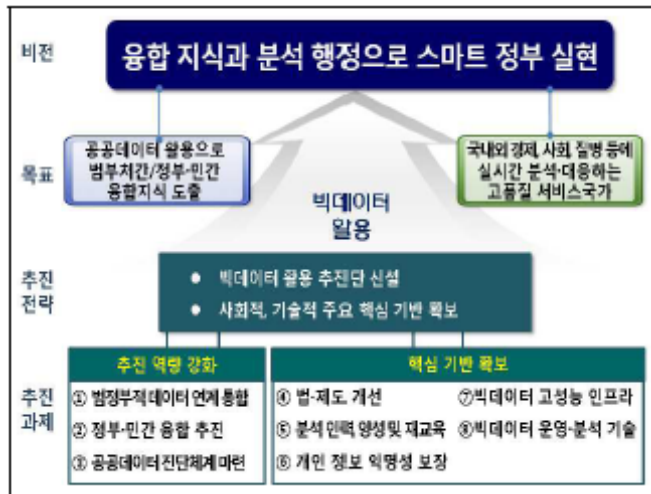


그림 2. 빅데이터 활용 추진방안의 비전 및 목표[1]
Fig. 2 Vision with big data

보유하고 있는 데이터를 말한다. 곧 정보 내지 데이터가 급증하면서 발생하는 현상을 빅데이터 문제라 한다면, 같은 이유로 공공데이터 문제는 정부 혹은 공공기관에서 보유한 정보나 데이터가 급증하는 것이라 할 수 있다. 2012년 7월에 새누리당 김을동 의원이 대표발의한 ‘공공데이터의 제공 및 이용 활성화에 관한 법률’에는 이 공공데이터에 대해, ‘공공기관이 생성 또는 취득하여 관리하고 있는 광(光) 또는 전자적 방식으로 처리되어 부호·문자·도형·색채·음성·음향·이미지 및 영상 등(이들의 복합체를 포함한다)으로 표현된 모든 종류의 자료 또는 정보’로 정의한다[7].

또한 국가정보화전략위원회는 공공데이터에 대해 정부의 정보화와 관련된 빅데이터로서 ‘대용량 데이터를 활용 분석하여 가치 있는 정보를 추출하고 생성된 지식을 바탕으로 능동적으로 대응하거나 변화를 예측하기 위한 정보화기술’로 정의하고 있다[1].

따라서 빅데이터에 연계된 기술은 공공데이터에 연계된 기술이라 할 수 있는데, -물론 빅데이터는 비정형 데이터의 비중이 높고, 반면 공공데이터는 정형데이터의 비중이 높은 차이는 있지만, 대체로 관련 기술들은 함께 사용가능하다- 표7[6]은 이러한 공공데이터에 관련되는 빅데이터에 연계된 기술들을 정리한 것이다.

먼저 빅데이터는 구글 파일시스템 상에 구축된 상용 분산 데이터베이스시스템으로서 H베이스에 영향을 미치는 것이다. 카산드라는 분산시스템에서 방대한 분량의 데이터를 처리할 수 있도록 디자인된 오픈소스 데이터베이스 관리시스템이다. 데이터웨어하우스 및 분석 어플라이언스는 데이터웨어하우스를 위해 서버, 스토리지, 운영체제, 데이터베이스, BI, 데이터마이닝 등 기타 여러 가지 소프트웨어가 최적화되어 설치된 통합제품이다. 분산시스템은 네트워크로 연결된 컴퓨터들의 집합으로서 동시에 일을 처리하기 위한 목적에서 단일 혹은 다수의 컴퓨터 리소스를 부분적으로 활용함으로써 시스템의 가격대비 성능비, 안정성, 확장성을 향상시킬 수 있다. 컴퓨터들의 집합으로서 동시에 일을 처리하기 위한 목적에서 단일 혹은 다수의 컴퓨터 리소스를 부분적으로 활용함으로써 시스템의 가격대비 성능비, 안정성, 확장성을 향상시킬 수 있다. 하둡은 구글이 개발한 맵리듀스를 오픈소스로 구현한 결과물인데, 야후에서 처음 개발되었고 현재 아파치 소프트웨어 재단에서 프로젝트의 하나로 관리되고 있으며, 분산시스템 상에서 대용량 데이터처리분석을 지원하는 오픈소스소프트웨어 프레임워크를 말한다. H베이스는 구글의 ‘빅데이터’를 참고로 개발된 오픈소스 분산 비관계형 데이터베이스다. 맵리듀스는

분산시스템 상에서 대용량 데이터 세트를 처리하기 위해서 구글이 제안한 소프트웨어 프레임워크이다. 비관계형 데이터베이스는 관계형 데이터베이스와 대조적인 개념으로서 데이터를 테이블에 저장하지 않는 데이터를 의미한다. 구글파일시스템은 하둡과 관련이 있는 것으로 구글에서 개발한 분산파일시스템이다.

이상에서 고찰한 데이터연계 기술들은 대체로 대용량 데이터와 비정형화된 데이터의 수집 및 검색, 데이터 전처리 및 분석기술, 시각화기술에 대한 것인데, 지향하는 바는 역시 정보의 공개강화라 할 수 있다.

즉 최근의 급격한 데이터 증가인 정보의 총량증가는 ‘정보의 홍수’ 혹은 ‘정보폭발’이 되었고, 이러한 현상과 인간사이의 갭을 줄이기 위한 노력이 바로 정보의 공개강화(대량의 정보에서 쓰레기정보와 유용한 정보를 구별하여 유용한 정보에 대한 대민접근성 강화 등)라는 점에서, 데이터연계기술들은 이 문제를 해결하기 위한 것이기 때문이다.

전자정부에 있어서도, 역시 정보공개(공공데이터의 대국민 공개강화와 정부기관 간의 데이터 공개 모두에서)가 중요하다. 곧 정부공개데이터운동(Open Government Data Movement)은 미국의 경우, 오바마 정부는 출범하자마자 연방정부를 더욱 투명하고, 참여적이며, 협력적으로 만들기 위한 Open Government Initiative initiative를 선포하고, 2009년 5월에는 이를 실현하기 위하여 웹 기술(cyberspace, e-mail)을 이용한 국민의 정책 결정 참여 프로세스와 웹사이트(regulation.gov)를 구축하였고, 공공기관에서 보유하고 있는 공공 정보(non-personal information)를 국민이라면 누구나 바로 이용할 수 있도록 하기 위해 ‘data.gov’를 오픈하였다.

영국의 경우는 미국처럼, 정부데이터를 공개추진(data.gov.uk)하였는데, 2009년 2월의 ‘Power of Information Taskforce Report’에서 정부가 정보에 대한 최대한 접근과 유용화를 통해 국민의 창의적 혁신을 지원하고, 개인과 커뮤니티가 삶의 질을 개선하기 위해 더 나은 정보에 의한 판단을 통해 솔루션을 만들어 낼 수 있도록 도와야 한다는 내용을 담았으며, ‘Digital Britain report’에는 영국의 인터넷 접근성(connectivity)을 높이기 위한 목적에서, 광대역 네트워크 및 무선 네트워크에 대한 인프라 투자뿐만 아니라 모든 공공 데이터의 공개와 접근을 위한 작업에

착수해야 함을 주장하였다.

호주의 경우는 정보관리청을 통해, 시간과 자원을 절감하기 위한 목적에서, 자동화된 툴을 활용하여 정부 2.0을 통한 정보 개방과 방대한 양의 정보를 검색하고 분석 및 재사용할 수 있도록 자동화된 툴을 활용하여 시간과 자원을 절감하였으며, AGIMO 산하의 정부 2.0 전략/서비스 팀에서는 정부 데이터에 대한 리포지터리 및 검색 툴을 서비스하는 ‘data.gov.au’ 웹 사이트를 운영하고 있다.

한편 한국의 2011년에 나온 국가정보화전략위원회의 ‘빅데이터를 활용한 스마트정부구현(안)’에는 빅데이터를 활용한 정책프로세스 혁신을 강조하면서 빅데이터 활용 시나리오로 5개의 예를 들고 있다. 곧 재난 전조감지, 구체역예방, 맞춤형 복지실현, 물가관리, 과학기술 및 의료 선진화 등을 통해 미래 빅데이터의 유용한 활용시나리오를 전망하고 있다.

또한 그림2와 같이, 공공데이터인 정부의 빅데이터 활용추진방안에서 융합지식과 분석행정으로 스마트정부의 구현을 비전으로 두고, 그 추진과제로 ‘범정부적 데이터의 연계 통합, 정부 및 민간 융합 추진, 공공데이터진단체계 마련’을 주요 이슈로 들고 있다. 따라서 이러한 추진과제가 달성되면 앞에서 말한 대국민 혹은 정부기관 간 정보공개는 더욱 강화되어 한국의 전자정부는 빅데이터 시대에 합당한 완전한 스마트정부를 구현하게 된다.

IV. 공공데이터를 활용한 국가정보화 전략방향 설정

1. 시나리오플래닝을 통한 한국 전자정부 및 정보화 미래상 도출

시나리오플래닝은 통계적 예측도 아니고 단일한 예측도 아니지만 미래를 알 수 있는 방법이며, 그러한 점에서, 시나리오플래닝은 미래의 불확실성을 제한적으로 보다 잘 이해할 수 있는 방법론으로서, 현재에서 미래까지의 미래 이슈의 진행과정을 서술하는 스토리라 할 수 있다[3].

따라서 본 연구는 미래예측방법으로 많이 활용되고 있는 시나리오플래닝 방법론을 적용하여, 공공데이터 활용에 있어서 특히 전자정부에 중점을 두고,

한국의 국가정보화의 바람직한 미래상을 도출한다.

전략의 방향성 탐색을 위한 가장 적합한 방법은 시나리오에 기반을 둔 전략설정인 점에서, 본 연구에서는 기존 시나리오플래닝 방법론에 따라 다음의 순서로 공공데이터 활용에 있어서 특히 전자정부에 중점을 두고, 한국의 국가정보화의 바람직한 미래상을 도출하고 향후 전략방안을 도출한다.

따라서 다음 그림3의 순서로 공공데이터 활용에 있어서 특히 전자정부에 중점을 두고, 한국의 국가정보화의 바람직한 미래상을 도출하고 전략방향 및 전략실행방안을 건설한다.

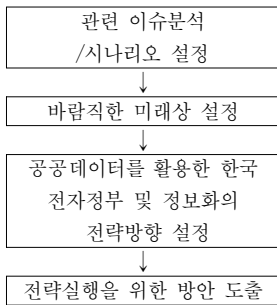


그림 3. 시나리오플래닝을 통한 공공데이터를 활용한 한국 (전자정부) 정보화의 전략방향설정 프로세스
Fig. 3 Strategy process for national information with electronic government of S. Korea by the application of scenario planning

앞 그림1에서 보듯이, 2010년 이후부터 급속한 ICT 관련 기술의 변화인 스마트폰과 같은 모바일혁명으로 이제 빅데이터 시대를 맞이했다. 그리고 2020년 무렵에는 본격적인 ZB시대가 예고되고 있다. 한국의 전자 정부는 2010년 UN평가 세계 1위 수준을 달성하였지만, 앞으로도 지속적으로 세계 최고수준을 유지하기 위해서는 도래한 빅데이터 시대에 적합한 전자정부 구축과 추진전략을 마련해야 한다. 또한 한국의 전자 정부는 그 동안 국가 정보화 계획과 함께 진행되어 왔다는 점에서 결국 제대로 된 국가 정보화 전략구축과 추진은 한국 전자정부의 세계 최고 수준을 유지하게 할 것이다.

본고는 그러한 점에서, 공공분야의 빅데이터인 공공데이터를 활용한 한국의 전자정부를 세계 최소수준으로 미래에도 지속적으로 유지할 수 있는 한국의 국

가정보화전략을 고찰하고 있다.

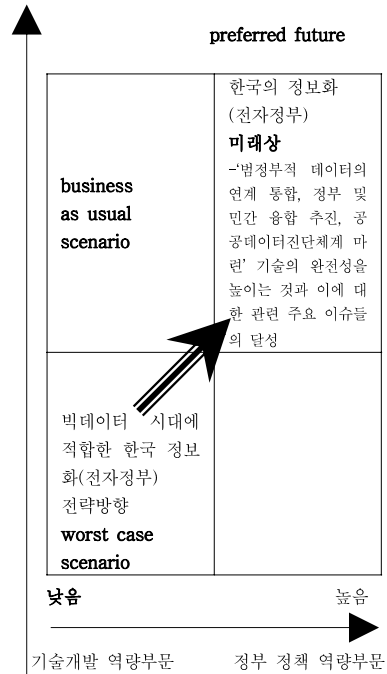


그림 4. 시나리오플래닝을 통한 전략방향 도출
Fig. 4 Strategy of national information with electronic government of S. Korea by the application of scenario planning

따라서 이를 위한 미래예측형 시나리오플래닝을 위해서는 현안에 대한 이슈도출, 그 이슈에 대한 분석, 이를 통한 중심축 설정은 필수적이다. 본고에서는 2011년에 나온 국가정보화전략위원회의 ‘빅데이터를 활용한 스마트정부구현(안)’에서 제시된 추진과제인 ‘범정부적 데이터의 연계 통합, 정부 및 민간 융합 추진, 공공데이터진단체계 마련’을 주요 이슈로 선정한다.

또한 이러한 이슈 등을 고려한 시나리오플래닝을 위하여 크게 공공데이터를 활용한 한국 전자정부의 정보화 전략방향을 ‘범정부적 데이터의 연계 통합, 정부 및 민간 융합 추진, 공공데이터진단체계 마련’ 등의 기술개발 역량부문과 분야 초기부터 정부의 정책적 적극지원 부문으로 2x2 Matrix를 그려서 바람직한 미래상을 설정한다. 또한 미래예측의 시간축은 ZB시대가 본격적으로 예고되는 ‘2020년 중반기’로 설정하였다.

이러한 시나리오플래닝을 통하여 그림4와 같이, 공공데이터를 활용한 한국 전자정부의 정보화에 있어서의 지향해야할 미래상(preferred future)을 도출할 수 있다. 공공데이터를 활용한 한국 전자정부의 정보화의 미래상은 ‘범정부적 데이터의 연계 통합, 정부 및 민간 융합 추진, 공공데이터진단체계 마련’ 기술의 완전성을 높이는 것과 이로 인한 관련 주요 이슈들의 달성이다.

이러한 시나리오 도출 로직을 통해 가장 바람직한 미래방향(preferred future)인 시나리오1을 도출하였다. 시나리오 1 전개(writing)의 주요내용은 다음과 같다.

시나리오 1(바람직한 미래상): 2020년 한국의 전자정부의 정보화 수준은 세계 최고 수준이다. 한국은 2010년부터 2020년 사이의 ‘정보폭발’의 빅데이터 시대에 합당한 미래전략을 세우고 정책을 추진하여 성공적으로 이를 달성하였다. 곧 한국은 ‘범정부적 데이터의 연계 통합’을 달성하여, 이제 정부 각 부처와 공공기관은 정보를 지속적으로 연계 및 수집할 수 있는 체계가 완전히 확립되었고, 덕분에 국가 전반의 데이터의 수집과 분석체계까지 마련하였다. 또한 ‘정부 및 민간 융합 추진’도 성공적으로 달성하였는데, 민간데이터와 공공데이터 사이의 연계 및 활용을 위한 체계와 기술을 완전히 확립하였으며, 공공데이터의 개방으로 민간의 가치창출과 기업활동 역시 강화되어 세계적인 경쟁력을 갖추었다. 한편 ‘공공데이터진단체계 마련’도 성공적으로 달성하여 다수의 시스템에서 활용되면서 동시에 중앙에서 통제 및 관리가 되는 표준 참조데이터인 MDM(Master Data Management)체계를 완전히 구축하였고, 덕분에 공공데이터는 효율적 관리가 되고 있으며, 품질관리 기준과 체계가 마련되어 범부처 데이터 공동활용도 성공적으로 이루어지고 있으며, 나아가 이러한 구축된 데이터에 대해 범정부 차원에서 집중적인 품질관리와 정비 사업 역시 차질 없이 진행되고 있다. 이제 한국의 전자정부와 국가정보화 수준은 세계 최고 수준에서 국제적인 경쟁력을 갖추고 각국의 모범사례가 되고 있으며, 관련 기술들은 세계 각국에 수출까지 되고 있다.

이와 대조적인 최악의 시나리오(worst case scenario) 4 전개(Writing)의 주요내용은 다음과 같다.

시나리오4(최악의 상황): 2020년 한국의 전자정부의 정보화는 완전히 국제 경쟁력을 상실하였다. 한국은 2010년부터 2020년 사이의 ‘정보폭발’의 빅데이터 시대에 합당한 미래전략을 세우고 제대로 된 정책을 추진하지 못했다. 한국정부는 세계 각국이 발 빠르게 빅데이터 시대에 적합한 전자정부와 관련 국가정보화 구축과 추진에 나섰을 때, 주요 이슈를 인식했음에도 불구하고 관련 이슈를 달성하는 기술개발과 전략추진에 완전히 실패하였다. 이제 한국의 전자정부 및 정보화는 빅데이터 시대와 ZB시대 이전의 수준에 여전히 머물러 있으며, 세계의 대부분의 주요국들과의 격차는 회복되기 힘든 실정이다. 이제 한국은 ZB시대의 전자정부와 국가정보화 관련 초기 시장 형성국에 진입하지 못해 선점자의 효과를 누리지 못하고 있으며, 수입국으로 전락하였다.

별다른 정책적 개입이 없이 현재의 상황으로만 유지되는 일상적 상황(Business as usual Scenario)은 다음과 같다.

시나리오3(일상적 상황): 2020년 한국정부의 전자정부와 정보화 수준은 빅데이터 및 ZB시대의 전자정부와 국가정보화에 있어, 선진국들을 추격하고 있다. 빅데이터 및 ZB시대를 앞두고 당시 세계 최고 수준이었던 한국은 ‘범정부적 데이터의 연계 통합, 정부 및 민간 융합 추진, 공공데이터진단체계 마련’과 같은 주요이슈를 선정하고 그 추진에 나섰지만, 초기부터 관련 분야에 대한 체계적이고 전략적인 추진방안을 도출하지 못한 이유로 정책수립에 실패하여, 그 동안 많은 시행착오를 겪었다. 이제 한국은 빅데이터 및 ZB시대에 적합하지 않은 예전의 전자정부에 머물러 있는 것은 아니지만, 그렇다고 해서 국제경쟁력을 갖춘 것도 아니며, 관련 기술을 선점하지 못한 이유로 현 2020년 시대에 적합한 국제 경쟁력을 갖춘 세계 최고 수준의 전자정부 및 정보화 수준을 달성하지 못한 실정이다.

2. 한국 전자정부 및 정보화 추진의 전략방향 설정

한국 전자정부 및 정보화의 바람직한 미래상인 시나리오 1을 달성하기 위한 주요한 전략은 다음과 같다.

본고의 주요 이슈인 ‘범정부적 데이터의 연계 통합, 정부 및 민간 융합 추진, 공공데이터진단체계 마련’의 목적의 지향점은 정부데이터인 공공데이터의 효과적

이고 성공적인 공개에 있다.

앞 그림2에서 보듯이, 빅데이터를 활용하여, 정보공개를 통한 공공데이터 활용으로 -앞에서 말했듯이 빅데이터는 비정형 데이터의 비중이 높고, 반면 공공데이터는 정형데이터의 비중이 높은 차이는 있지만, 대체로 관련 기술들은 함께 사용이 가능한데, 본고에서는 빅데이터인 공공데이터에 대해, 여기서의 빅데이터를 단순히 ‘데이터 크기’의 문제로 보지 않고, 그 ‘분석할 크기’의 문제로 보기 때문이다- 범부처간 그리고 정부와 민간 사이의 융합지식을 도출하여, 실시간 분석 및 대응 가능한 고품질의 서비스 국가인 스마트 정부를 실현하고자 함이다.

결국 이상의 목표를 달성하기 위해서는 공공데이터가 목적에 부합하게 성공적이며 효과적으로 활용될 수 있도록 이슈를 달성하는 정보공개관련 기술들이 초기부터 국제 경쟁력을 갖도록 체계적인 정책 차원에서 제대로 발전되어야 한다.

본고에서는 이를 위해 시맨틱웹(Semantic Web)기술과 함께하는 링크드데이터(Linked Data)기술에 대한 초기부터의 집중적인 투자와 개발을 그 전략방안으로 제안한다.

먼저 시맨틱웹은 실세계의 사물들을 가리키는 리소스 단위를 기본으로 하여, 데이터들 간의 의미 있는 링크들로 구성된 데이터의 웹(the Web of Data)이며, 웹 문서를 만들기 위하여 사용하는 기본적인 프로그래밍 언어의 한 종류인 HTML(Hyper Text Markup Language)로 만들어진 -문서(document)들 간의 의미 없는 링크들로 연결된- 기존의 웹에서 진화한 것이다.

특히 시맨틱데이터는 그림 5처럼, 데이터의 의미를 표현하기 위해 ‘주어부-술어부-목적어부’라는 트리플 형식으로 기술하여 이를 마크업하기 위한 표준을 제공하는 RDF(Resource Description Framework)형식에 따라 데이터가 작성된다. 따라서 이렇게 작성된 RDF데이터는 의미 있는 링크를 통해 데이터연계와 통합(federation)이 가능하며, 데이터의 상호간 운용성이 보장되는데, 기계가 데이터를 이해하고 처리하기 때문이다. 결국 RDF데이터는 데이터의 융복합 및 매쉬업 그리고 협업서비스를 제공하는 기반이 된다.

“이상윤은 학교에 다닌다”

트리플 형식의 의미 표현	이상윤	다닌다	학교
	주어부	술어부	목적어부
기계 가독형 표현	<pre><rdf:Description rdf:about="#이상윤"> <rdf:type rdf:type="http://www.w3.org/2002/07/owl#Thing"/> > <rdf:label rdf:data-type="http://www.w3.org/2001/XMLSchema#string">이상윤 </rdf:label> <다닌다 rdf:resource="#학교"/> </rdf:Description></pre>		

그림 5. RDF데이터 방식
Fig. 5 Method of RDF data

한편 링크드데이터(Linked Data)는 WWW(World Wide Web)을 처음 개발한 Tim Berners-Lee에 의해 2006년에 최초로 제안된 것으로, 시맨틱데이터 상호간의 연결을 통해 데이터를 공개하고 연계하며 공유하는 방법인데, 시맨틱데이터에 웹으로 접근 가능한 이름(URI:Uniform Resource Identifier, 사물(thing)을 기술(description)하는 데이터에 대한 웹에서의 유일식별자를 지칭)을 붙이는 방식을 통한다.

곧 링크드데이터는 시맨틱웹이 구현하고자 하는 데이터의 웹(the Web of Data)을 달성하는 구체적인 수단 혹은 방법이라 할 수 있다.

한편, 데이터를 서로 연결(link)하는 이러한 링크드데이터는 다음과 같은 유용성이 있다. 첫째, URI를 이용하여 링크만 해주면 자신이 만든 것이 아니더라도 누구나 이용할 수 있어 데이터의 재사용이 가능하다. 둘째, 링크드 데이터를 통해 공개된 데이터를 이용하면 내가 원하는 데이터의 존재성 유무와 함께, 어디에 존재하는지 알 수 있으므로, 데이터의 중복이 감소된다. 셋째, 시맨틱웹 표준인 RDF 형태의 데이터로 발행(publish)되므로 마치 하나의 글로벌 데이터베이스처럼 질의하고 이용가능한 점에서, 데이터의 통합성이 높고 상호운용성을 극대화할 수 있다. 넷째, URI로 구별되는 데이터 리소스의 자유로운 접근 및 이용이 가능하므로 손쉽게 데이터의 매쉬업이 가능하여 부가가치가 창출된다. 다섯째, 데이터 네트워크화를

가속(Network of Data, Cloud of Data)할 수 있는데, 초창기의 웹과 같이 데이터의 자유로운 연결과 이용이 가능하기 때문이다.

결국 정부데이터인 공공데이터에 이상과 같은 시맨틱웹(Semantic Web)기술과 함께하는 링크드데이터(Linked Data)기술이 체계적으로 적용되면, 주요 이슈들인 ‘법정부적 데이터의 연계 통합, 정부 및 민간 융합 추진, 공공데이터진단체계 마련’이 달성되어, 이상과 같은 유용성을 가져올 수 있는데, 공공데이터에 대한 재사용이 증대되고, 데이터의 중복 역시 감소되며, 데이터의 통합성을 높여 상호운용성이 극대화되고, 데이터 매쉬업을 통해 부가가치가 창출되며, 공공데이터의 네트워크화를 가속(Network of Data, Cloud of Data)할 수 있게 된다.

곧 공공데이터에 대한 링크드데이터 기술의 적용은 공공데이터에 대해 RDF/XML과 같은 표준 포맷을 이용하게 하여 다양한 어플리케이션에서의 공공데이터에 대한 접근성을 높이고, 데이터가 리소스 단위로 표현되므로 다양한 매쉬업과 데이터 통합이 가능하여 모듈성(Modular)을 높일 수 있으며, 높은 확장성으로 쉽게 링크드데이터를 추가할 수 있다는 점과 함께, 시간이 지나도 지속성을 보장할 수 있게 된다는 점에서, 정부데이터인 공공데이터의 효과적이고 성공적인 공개를 가능하게 하여, 정부기관의 데이터 간의 연계와 통합이 용이해지고, 정부 및 민간 간에도 융합이 추진되며, 나아가 합리적인 공개를 위한 공공데이터에 대한 진단체계도 마련되게 된다.

V. 한국전자정부 정보화와 기계의 이해

현재 한국의 모든 공공기관은 국가안보와 같은 법으로 정한 비공개 대상정보를 제외하고는 모두 공개하도록 ‘공공기관의 정보공개에 관한 법률’로 이를 규정[8]하고 있는데, 작성된 정보목록에 따라 정보통신망을 활용하여 정보공개시스템을 통하여 공개하도록 되어있어, 전자정부 차원에서는 인터넷 접속만으로 정부가 가진 정보에 대해 어떤 것을 공개하고 있는지 확인이 가능한 상태이다.

표 8. HTML웹과 Semantic웹 상에서의 한국 전자정부의 정보화 위상 비교
Table 8. HTML and semantic web for national information with electronic government of S. Korea

구분	Web of Document (Web 1.0)	Web of Data (Semantic Web 혹은 Web 3.0)
목적	사람의 이해	기계의 이해
링크 단위	문서(document)들 간 단순링크	사물(또는 리소스 간)의 의미적 링크
표현 형식	HTML(Hyper Text Markup Language)	RDF(Resource Description Framework)
현재 위상	현재의 한국 전자정부 정보화 수준	한국 전자정부 정보화 미래상

다만 현재의 한국 전자정부의 정보화 문제는 표8과 그림6과 같이, 문서(document)들 간의 단순링크인 Web of Document(Web 1.0) 수준에 머물러 있어, 곧 국가지식포털을 통해 공공데이터와 중복되는 여러 보고서에 접근할 수 있는 정도의 수준이라는 점에서, 사실상 ‘사람의 이해’만 가능한 정도(그림6의 b: 현 정부의 부처간 정보교류방식으로서 완성된 보고서 형태의 인터랙션만 가능한 상태)이지, 본고에서 제안한 시맨틱웹과 함께하는 링크드데이터가 가능한 RDF (Resource Description Framework)표현 방식의 사물(또는 리소스 간)의 의미적 링크가 되어, ‘기계의 이해’가 가능한 Web of Data 수준(그림6의 a: 본고의 발전방안으로서 데이터 간의 인터페이스가 가능한 상태)은 아니며, 그런 점에서 대부분의 정보는 기계에 의한 접근이 불가능한 문서로 제한되어 있는 상태이다.

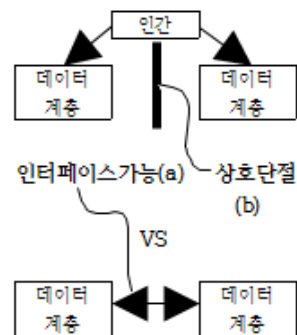


그림 6. 공공데이터 접근방법 비교
Fig. 6 Comparison of public data approach

한 예로 그림6에서 보듯이, 한국정보화진흥원이 운영하는 국가지식포털은 기관의 작성한 보고서에 대해, 직접적인 다운로드기능이나 링크를 제공하고 있는데, 이는 이러한 문서에 대해 기본적으로 목적의식을 갖고 찾는 국민에게만 개방되는 정도(‘인간의 이해’ 수준)에 불과한 실정으로서 데이터 계층에서의 직접적인 접근(‘기계의 이해’)은 가능하지 않다.

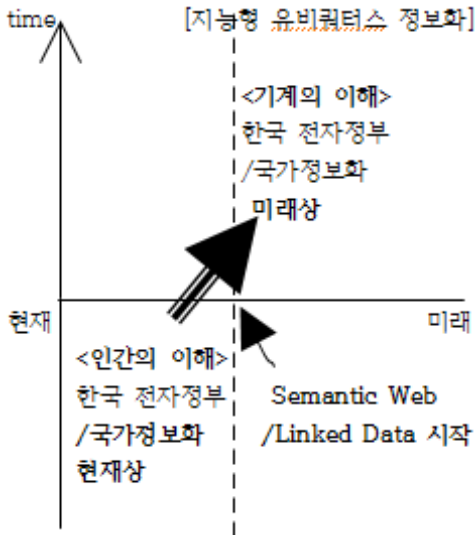


그림 7. 한국 전자정부 및 국가정보화의 세계 최고 유지를 위한 전략
 Fig. 7 Strategy of national information with electronic government of S. Korea

따라서 그림7처럼, 한국 전자정부의 지속적인 세계 최고 수준을 유지하기 위해서는 이러한 문제점을 해결하여 ‘기계의 이해’가 가능한 수준으로 성장해야 할 것이다. 또한 이를 통해, 본고의 ‘범정부적 데이터의 연계 통합, 정부 및 민간 융합 추진, 공공데이터진단체계 마련’의 주요 이슈들은 성공적으로 달성될 것이다.

VI. 결 론

‘언제 어디서나 존재한다.’는 유비쿼터스(Ubiquitous) 컴퓨팅[9]과 함께하는 -최근의 지능형 유비쿼터스 정보화를 지향하는- 빅데이터 시대 도래는 정부의 정보화와 관련된 빅데이터인 공공데이터에도 많은

영향을 끼치고 있다.

미국과 영국 등 선진국들은 데이터를 경제적 가치로 인식하여 -클라우드 환경에 연계되는 대용량의 데이터 관리[10] 관점에서- 공공데이터의 활용을 통해 정부의 대민서비스를 더욱 증대시키고자 이를 정책적으로 적극 추진(미국의 data.gov와 영국의 data.gov.uk 구축 등)하고 있다. 곧 정부가 갖고 있는 데이터를 공개하여 더욱 투명한 정부를 구현하고 부가가치를 창출하고자 노력 중인데, 앞으로의 국가 정보화는 이러한 공공데이터의 성공적인 대국민개방여부에 달려 있다고 해도 과언이 아닐 것이다.

따라서 현재의 한국의 세계 최고 수준의 전자정부 수준 유지를 위해서는 공공데이터의 대국민개방의 폭을 넓히는 방향의 국가정보화 추진이 성공적으로 달성되어야 하는데, 본고에서는 컴퓨터와 같은 기계 스스로 웹에 접근하여 데이터를 분석하는 기술인 시맨틱웹과 함께, 이의 구체적인 방법인 링크드데이터(Linked Data)기술에 대한 초기부터의 집중적인 투자와 개발을 통한 공공데이터의 공개추진을 그 전략방안으로 제안했다.

결국 본고의 제안처럼, 시맨틱웹을 구현하는 링크드데이터를 통해 기계간의 공공데이터의 상호교환이 자유롭게 가능해지고 의미처리가 성공하면, 한국의 전자정부와 국가정보화는 지능형 유비쿼터스 정보화를 완전히 달성하게 된다.

참고 문헌

- [1] 국가정보화전략위원회, “빅데이터를 활용한 스마트정부구현(안)”, pp. 1-27, 2011.
- [2] 정보화추진위원회, “국가정보화기본계획”, pp. 21-24, 2008.
- [3] 이상윤, “공간정보시스템을 활용한 인터넷전자투표연구: 시나리오플래닝을 중심으로”, 기술혁신학회지, 15권, 3호, pp. 604-626, 2012.
- [4] 행정안전부·방송통신위원회·지식경제부, “2012 국가정보화백서”, pp. 1-780, 2012.
- [5] 전자정부특별위원회, “전자정부 백서”, pp. 1-230, 2003.
- [6] 한국정보화진흥원, “신가치창출 엔진, 빅데이터의 새로운 가능성과 대응전략”, pp. 1-34, 2011.
- [7] 김을동, “공공데이터의 제공 및 이용 활성화에

- 관한 법률(가)안”, pp. 1-50 ,2012.
- [8] 법률 제10012호, “공공기관의 정보공개에 관한 법률”, 2010.
- [9] 노영학, 박종철, “유비쿼터스 주거시설 선호도 분석”, 한국전자통신학회 2011년 추계종합학술대회 논문집, pp. 73-77, 2011.
- [10] 김치연, “클라우드 환경에서 확장성을 지원하는 트랜잭션 처리 방법”, 한국전자통신학회논문지, 7권, 4호, pp. 873-879, 2012.

저자 소개



이상윤(Sang-Yun Lee)

2002년 부산대학교 조선해양공학과 졸업(공학사)

2009년 부산대학교 대학원 정치외교학과 졸업(정치학석사)

2011년 부산대학교 대학원 융합기술정책 박사수료

2012년~현재 부경대학교 공간정보시스템공학과 겸임교수

※ 관심분야 : R&D기술개발, 과학기술정책



윤홍주(Hong-Joo Yoon)

1983년 부경대학교 해양공학과 졸업(공학사)

1985년 부경대학교 대학원 해양공학과 졸업(공학석사)

1997년 프랑스 그르노블 I 대학교 대학원 위성원격탐사 전공 졸업(공학박사)

2010년 부산대학교 대학원 융합기술정책 박사수료

1999년~2002년 전남대학교 해양공학과 교수

2002년~현재 부경대학교 공간정보시스템공학 교수

※ 관심분야 : 원격탐사 & GIS, 공간정보정책학