

제4차 산업혁명 기술의 기록관리 적용 방안*

Application of 4th Industrial Revolution Technology to Records Management

안대진(An, Dae-jin)** · 임진희(Yim, Jin-hee)**

1. 머리말
 - 1) 연구 목적
 - 2) 연구 배경
 - 3) 연구 범위와 방법
2. 제4차 산업혁명의 주요 기술
 - 1) 클라우드 컴퓨팅
 - 2) 빅데이터
 - 3) 인공지능
 - 4) 사물인터넷(IoT)
3. 제4차 산업혁명 기술에 따른 생산기록의 특징
 - 1) 클라우드 컴퓨팅 환경에서의 기록의 특징
 - 2) 빅데이터 기술 도입에 따른 기록의 특징
 - 3) 인공지능 기술 도입에 따른 기록화 전략의 변화
 - 4) 사물인터넷 환경에서의 기록의 특징
 - 5) 시사점
4. 기록관리 적용 방안
 - 1) 클라우드 컴퓨팅 기술의 기록관리 적용 방안
 - 2) 빅데이터 기술의 기록관리 적용 방안
 - 3) 인공지능 기술의 기록관리 적용 방안
 - 4) 사물인터넷 기술의 기록관리 적용 방안
5. 맺음말

* 본 연구는 2017년 국가기록원 R&D사업 '차세대 기록관리 모델 재설계 연구'의 일환으로 수행된 연구임.

** 명지대학교 기록정보과학전문대학원 박사과정, (주)아카이브랩 대표(주저자).

*** 서울대학교 공학연구원 책임연구원(교신저자).

■ 투고일 : 2017년 10월 16일 ■ 초심사일 : 2017년 10월 16일 ■ 게재확정일 : 2017년 10월 26일

〈초록〉

이 연구에서는 제4차 산업혁명의 신기술을 활용하여 기록관리를 더 잘 할 수 있는 방안을 살펴보았다. 이를 위해 기록의 생산 및 관리에 큰 영향을 미치는 클라우드, 빅데이터, 인공지능, 사물인터넷 등 네 가지 기술들을 선정하고 이 기술들의 개념과 특징, 적용분야를 정리하고 대표적 기술들을 테스트하였다. 이후 각 기술로 인해 변화된 생산기록의 특징을 분석하였다. 신기술로 인해 기록의 생산량은 급증했고 유형은 다양해졌다. 또한 데이터의 품질 확보와 생산 맥락을 설명하기 위한 조치가 필요함을 알 수 있었다. 각 기술을 기록관리에 효과적으로 도입하기 위해서는 기술의 성숙도별로 당장 적용할 기술과 나중에 적용해야 할 기술을 구분하여 로드맵을 설계해야 한다. 생산기록의 특징 변화에 대응하기 위해서는 유연한 데이터 구조를 갖추고 표준화된 형식으로 생산해야 한다. 또한 공공기관이 SaaS 제품을 조달할 수 있도록 하고 디지털 기술을 이용하여 공공서비스의 품질을 높여야 한다.

주제어 : 제4차 산업혁명, 클라우드 컴퓨팅, 빅데이터, 인공지능, 사물인터넷

〈Abstract〉

This study examined ways to improve records management by using the new technology of the Fourth Industrial Revolution. To do this, we selected four technologies that have a huge impact on the production and management of records such as cloud, big data, artificial intelligence, and the Internet of Things. We tested these technologies and summarized their concepts, characteristics, and applications. The characteristics of the changed production records were analyzed by each technology. Because of new technology, the production of records has rapidly increased and the types of records have become diverse. With this, there is also a need for solutions to explain the quality of data and the context of production.

To effectively introduce each technology into records management, a roadmap should be designed by classifying which technology should be applied immediately and which should be applied later depending on the maturity of the technology. To cope with changes in the characteristics of production records, a flexible data structure must be produced in a standardized format. Public authorities should also be able to procure Software as a Service (SaaS) products and use digital technology to improve the quality of public services.

Keywords : 4th Industrial Revolution, Cloud Computing, BigData, AI, IoT

1. 머리말

1) 연구 목적

이 연구의 목적은 제4차 산업혁명의 신기술을 활용하여 기록관리를 더 잘할 수 있는 방안을 살펴보는 것이다. 이를 위해 기록의 생산 및 관리와 관련이 깊은 기술들을 선별하고 일반적인 특징과 기록관리에 미치는 영향을 분석하여 적용방안을 제시하였다. 빅데이터, 인공지능, 사물인터넷, 클라우드 컴퓨팅 등 신기술이 주도하는 새로운 환경에서 차세대 기록관리의 원칙을 수립하고 대비하는 데 시사점을 제공하고자 한다.

2) 연구 배경

2016년 세계경제포럼에서는 처음으로 과학기술 분야가 의제로 채택되었

다. 회장인 클라우스 슈밥(Klaus Schwab)은 인공지능, 자율주행자동차, 유전공학, 뇌과학 등을 총망라하는 변화의 흐름이 제4차 산업혁명을 불러올 것이며 이에 대비하기 위해 ‘파괴적 혁신(disruptive innovation)’이 필요함을 강조했다(World Economic Forum 2016). 제4차 산업혁명이 모든 산업과 비즈니스 모델에 파괴적 수준의 혁신을 초래하고 새로운 가치를 창출하여 우리 삶에 전례 없는 변화를 가져올 것이라는 것이다. 최근 실리콘 밸리에서도 현재의 판을 완전히 뒤엎고 영역 간의 경계를 없앤다는 의미에서 ‘파괴적 혁신(disruptive innovation)’이나 ‘파괴적 기술(disruptive technology)’이란 용어가 유행하고 있다. 또한 파괴적 기술은 저성장 장기화의 위기를 극복할 수 있는 기술로도 주목받고 있다(엄도영, 김동휴, 이희진 2016). 산업통상자원부 역시 제4차 산업혁명이 성공하려면 이러한 파괴적 기술들을 모든 산업, 모든 영역으로 확산하는 것이 관건이라고 강조했다(산업통상자원부, 2017).

클라우스 슈밥은 제4차 산업혁명의 선결과제로 이러한 혁신 기술을 이해하는 것을 꼽았다(Schwab, K 2017). 파괴적 기술들의 특성을 이해해야만 각 영역에 미칠 영향력을 측정하는 등의 실질적인 준비가 가능해질 것이기 때문이다. 제4차 산업혁명은 인공지능(A.I., Artificial Intelligence), 로봇, 사물인터넷(IoT, Internet of Things), 빅데이터(Big data) 등 정보통신 기술이 사회 전 분야에 적용되어 경제·사회 구조의 근본적 변화를 촉발시키는 기술혁명으로 정의된다(Schwab 2017). 2016년 세계경제포럼에서는 제4차 산업혁명을 선도하는 10대 기술로 물리학(무인운송수단, 3D 프린팅, 첨단 로봇공학, 신소재), 디지털(사물인터넷, 블록체인, 공유경제/온디맨드경제), 생물학(유전공학, 합성생물학, 바이오프린팅) 등 3개 분야 기술을 선정했다. 전 세계적으로 ‘사물인터넷’, ‘로봇공학’, ‘3D 프린팅’, ‘빅데이터’, ‘인공지능’ 등 5대 기술이 제4차 산업혁명을 촉발하는 주요 기술 변화 동인으로 꼽히고 있다(한국은행 2016). 이 중 사물인터넷, 빅데이터,

인공지능은 기록의 생산과 관리에도 큰 영향을 미칠 것으로 쉽게 예상해 볼 수 있다. 새로운 유형의 기록 출현이나 보다 지능적인 방식의 기록 활용 등이다.

국가기록원 역시 ‘차세대 기록관리 모델 재설계’ 연구를 통해 제4차 산업혁명 시대를 대비한 기록관리 법과 제도의 개편, 행정정보데이터세트의 관리와 기록시스템 표준모델의 재설계, 지능형 서비스 방안 등을 모색하고 있다. 지난 9월에 열린 2차 워크숍에서는 3대 전략 10대 아젠다의 주요 내용과 클라우드 기반의 모듈화된 기록시스템 모형, 인공지능 기술을 접목한 지능형 서비스 모델이 제시되었다. 한편 서울시는 2019년 서울기록원 건립을 앞두고 국가기록원의 이러한 모형을 토대로 마이크로서비스 방식의 모듈화된 기록시스템을 개발하고 있다. 또한 머신러닝 기술의 적용방안 등 신기술의 도입방안을 연구하고 있다. 이처럼 제4차 산업혁명의 파괴적 기술들은 기록관리 분야에도 큰 변화를 예고하고 있다. 기록의 생산과 관리행위는 업무의 투명성, 설명책임성, 기억, 사회정의 등의 가치를 창출해 왔다. 파괴적 기술들이 기존의 기록 생산, 관리 행위에 어떠한 영향을 미칠 것인지, 기존의 가치를 쉽게 달성하거나 새로운 가치를 창출할 것인지에 대한 연구가 필요한 시점이다.

3) 연구 범위와 방법

이 연구에서는 제4차 산업혁명의 주요 기술 중 기록의 생산 및 관리에 큰 영향을 미치는 네 가지 기술을 중점적으로 분석하였다. 클라우드 컴퓨팅, 빅데이터, 인공지능, 사물인터넷 등이다. 대상을 선정한 이유는 이 네 가지 기술이 최근 공공 업무에 본격적으로 도입되기 시작했고 기술의 성숙도 측면에서 근시일 내에 공무원들의 일상 업무에 적용될 확률이 크기 때문이다. 이 연구는 문헌연구와 실험연구로 진행되었다. 문헌연구는 각 기술

에 대한 논문이나 조사보고서, 기술을 제공하는 기업들의 서비스 홈페이지나 개발문서 등을 조사하여 정리하는 방식으로 수행하였다. 실험연구에서는 각 기술과 관련된 알고리즘이나 상용 서비스를 사용해 보고 작동 방식과 기술 수준을 테스트했다. 테스트 결과는 따로 정리하지 않고 필요한 부분에 기술하거나 제안내용의 근거로 활용하였다. 구글 클라우드, IBM 블루믹스(Bluemix), 아마존 웹 서비스(AWS)에서 제공하는 데모 프로그램이나 클라우드 API 서비스를 주로 이용했다.

이 연구에서는 제4차 산업혁명의 여러 기술 중 4개의 기술만을 대상으로 기록관리에 미치는 영향을 분석하였다. 그리고 실험연구를 통해 얻은 기술 성숙도를 참고하여 공공기록관리 업무에 적용하기 위한 방안을 제시하였다. 따라서 분석 방법론이 정교하지 않으며 다소 주관적인 견해가 포함되었다. 분석의 틀에 가치사슬 모형을 적용하여 기록관리 업무의 주 활동과 하위기능을 나열하고 각 기술이 가치 창출에 미치는 영향을 밝히고자 했으나 국내외 기록관리에 적용된 사례가 부족하여 이 모형을 적용하지 못했다. 향후 신기술들이 본격적으로 활용되는 시점에서는 이러한 방법의 연구로 확장시킬 필요가 있다.

2. 제4차 산업혁명의 주요 기술

1) 클라우드 컴퓨팅

클라우드 컴퓨팅은 네트워크, 서버, 스토리지, 애플리케이션, 서비스 등의 컴퓨팅 자원을 주문에 따라 어디서나 접근 가능하도록 구성하여 제공할 수 있는 컴퓨팅 모델을 말한다(Mell & Grance 2011). 클라우드 컴퓨팅의 특징으로는 유연성, 경제성, 효율성, 신속성, 편의성 등을 들 수 있다(KDB산

업은행 2017). 수요에 따라 컴퓨팅 자원을 할당하여 유연하게 이용할 수 있고 사용한 만큼 비용을 지불하므로 경제적이다. 또한 G-클라우드에서 보듯이 IT자원 공동화로 효율성을 추구할 수 있으며 빠르게 원하는 컴퓨팅 환경을 생성하고 시간과 장소의 제약 없이 편리하게 사용할 수 있다.

클라우드 컴퓨팅은 제조, 서비스, 인프라 등 광범위한 산업 분야에 적용되고 있다. 공공 부문에서도 공공기관 업무환경의 가상화, 교통 및 항공관리 시스템 등의 인프라가 클라우드로 교체되고 있다(KDB산업은행 2017). 서비스 모델별로 보면 서버, 스토리지 등 인프라를 제공하는 IaaS(Infra as a Service), 소프트웨어 개발환경을 제공하는 PaaS(Platform as a Service), 완성된 소프트웨어를 제공하는 SaaS(Service as a Service) 등으로 구분할 수 있다. 우리나라 공공부문에 구축된 G-클라우드는 IaaS 수준의 인프라 통합 관리 단계로 볼 수 있으며, 아마존 등 상용 클라우드는 SaaS 형태의 완성된 서비스를 API로 제공하는 수준에 이르렀다.

IBM¹⁾, 구글²⁾, 아마존³⁾, 마이크로소프트⁴⁾ 등 글로벌 IT기업들은 클라우드 기반의 빅데이터와 인공지능 서비스를 내놓았다. 최근에는 사물인터넷, 블록체인(Blockchain), 마이크로서비스(Microservices)와 도커(Docker), 데브옵스(DevOps) 등의 신기술까지 클라우드 API로 제공하며 클라우드가 명실상부한 제4차 산업혁명의 핵심 플랫폼으로 자리매김하고 있다. 클라우드 컴퓨팅은 위험요인으로 지적되고 있는 보안 및 표준 부재 등에도 불구하고 그 혁신적 특성으로 인해 신기술들의 기반 기술로 각광받고 있다.

1) IBM Bluemix 카탈로그, <https://cloud.google.com/products/>

2) Google Products & Services, <https://cloud.google.com/products/>

3) Amazon Web Service 클라우드 제품, <https://aws.amazon.com/ko/products/>

4) Microsoft Azure Products, <https://azure.microsoft.com/en-us/services/>

〈표 2〉 주요 클라우드 플랫폼별 제공 서비스 현황(2017년 10월 기준)

IBM 블루믹스	구글 클라우드	아마존 웹 서비스	마이크로소프트 Azure
인프라 • 컴퓨팅 • 스토리지 • 네트워크 • 보안 • 컨테이너 • VMWare 플랫폼 • 표준 유형 • API • 애플리케이션 서비스 • 블록체인 • DevOps • 금융 • 기능 • 통합 • Internet of Things • 웹 및 모바일 • 네트워크 • 보안 • Watson	• 컴퓨팅 • 관리도구 • 빅 데이터 • 네트워크 • 인증 & 보안 • 머신 러닝 • Internet of Things • 스토리지 & DB • 개발자 도구 • 데이터 전송 • API 및 에코시스템	• 컴퓨팅 • 스토리지 • 데이터베이스 • 마이그레이션 • 네트워크 및 콘텐츠 전송 • 개발자 도구 • 관리 도구 • 보안, 자격 증명 및 규정 준수 • 분석 • 인공 지능 • 모바일 서비스 • 애플리케이션 서비스 • 메시징 • 비즈니스 생산성 • 데스크톱 및 앱 스트리밍 • 사물 인터넷 • 게임 개발	• 컴퓨팅 • 네트워킹 • 스토리지 • 웹 + 모바일 • 컨테이너 • 데이터베이스 • 데이터 + 분석 • AI + 코그니티브 서비스 • Internet of Things • 엔터프라이즈 통합 • 보안 + 자격증명 • 개발자 도구 • 모니터링 + 관리

2) 빅데이터

2010년 가트너의 애널리스트 더그 레이니는 데이터의 급성장에 따른 이슈와 기회를 데이터의 양(volume), 데이터 입출력의 속도(velocity), 데이터 종류의 다양성(variety)이라는 세 개의 차원으로 정의하였다(Laney 2010). 이 “3v” 모델은 이후 가장 널리 사용되는 빅 데이터의 정의가 되었다. 즉 빅데이터는 아날로그 환경에서 생성되는 데이터보다 그 규모가 페타바이트 수준으로 방대하고, 형태도 수치, 문자, 영상 등으로 다양하며, 생성 주기도

짧은 데이터를 말한다. 이러한 특징에서 보듯이 빅데이터는 IT의 흐름이 하드웨어, 소프트웨어에 이어 데이터로 이동하기 시작하는 정보형태의 변화, 정보관리 및 기대치의 변화에 원인이 있으며, 통계 등 단순 분석보다는 데이터의 의미와 가치에 더 많은 관심을 가지기 시작하면서 발전하기 시작하였음을 알 수 있다(김성태 2013).

빅데이터 분석을 위해서는 광범위한 데이터를 수집하는 단계부터 저장, 처리, 분석, 활용하는 각 단계의 기술이 요구된다. 하둡(Hadoop)⁵⁾, 아파치 스파크(Apache Spark)⁶⁾ 등 주요 빅데이터 분석 프레임워크들은 몽고디비(MongoDB)⁷⁾ 등 각 단계의 오픈소스 소프트웨어를 제공하는 벤더와 고객사들로 구성된 독자적인 생태계를 구축하고 있다. 또한 글로벌 IT기업들은 자사의 클라우드 플랫폼을 통해 빅데이터 분석의 전 과정을 지원하는 제품군들을 묶어 제공하고 있다. 최근에는 머신러닝이 점차 빅데이터 애플리케이션과 서비스에 탑재되고 있으며, 사물인터넷과 오픈소스 기술의 발전으로 인해 실시간 데이터를 처리하는 스트리밍 분석 서비스까지 출현했다.

빅데이터 분석은 업무 효율성 개선, 고객 관계관리, 신규 가치 창출 등을 목적으로 전 산업 부문에서 이루어지고 있다. 가트너는 빅데이터 분석의 가치 사슬⁸⁾ 모형을 통해 빅데이터 분석으로 창출되는 가치를 설명하기도 했다(Gartner 2015). 고객 인사이트 획득(Customer Insight), 제품과 프로세스 효율성(Product and Process Efficiency), 디지털 제품 및 서비스(Digital Products and Service), 운영 효율성(Operational Excellence), 디지털 마케팅(Digital Marketing), 위험관리 및 법규준수(Risk Management and Compliance) 등

5) Apache hadoop, <http://hadoop.apache.org/>

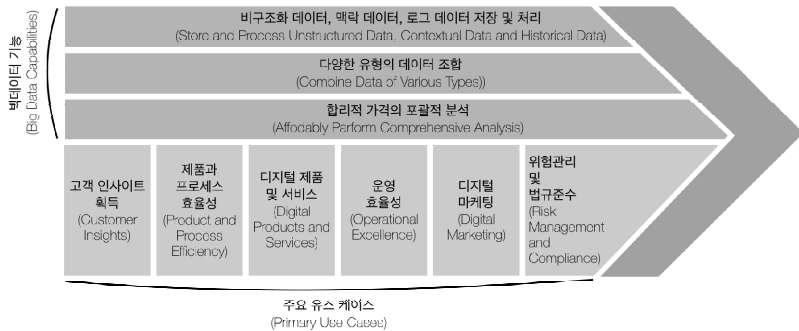
6) Apache Sprak, <https://spark.apache.org/>

7) MongoDB, <https://www.mongodb.com/>

8) 가치사슬(Value Chain)이란 기업이 고객에게 가치(Value)를 전달하기 위해 수행하는 주요 활동, 기능, 프로세스들을 주활동(Primary Activities)과 지원활동(Support Activities)로 나누어 정리한 기업가치 분석 모델이다(Porter 1991; Porter 2008).

여섯 가지이다. 어떠한 분야에서든 빅데이터 분석을 업무에 도입하고자 할 때 이 여섯 가지 중 어디에 해당되는지를 인식하고 도입효과를 가늠해 보는 것이 좋을 것이다. 산업별 적용사례로 보면 제조 부문에서는 발전소 운영 효율 시스템, 스마트공장 품질관리, 서비스 부문에서는 고객 성향분석을 통한 맞춤형 서비스 제공, 물류 효율화, 공공 부문에서는 전력수요 예측, 재난예측, 교통량 예측 등을 들 수 있다(KDB산업은행 2017).

〈그림 1〉 빅데이터 가치사슬 모형



※출처 : Gartner 2015

빅데이터 도입 관련 주요 쟁점으로 플랫폼의 선정, 정보보안, 개인사생활 침해, 정보의 범위와 깊이 등이 논의되고 있다(조영임 2013). SNS, 공공데이터 등에까지 빅데이터 수집, 분석의 범위가 미치다 보니 민감정보나 개인 정보가 유출되어 사생활을 침해할 위험이 발생하고 있다. 따라서 빅데이터의 윤리 정책에도 많은 관심이 집중되고 있다.

3) 인공지능

인공지능은 인간의 사고능력을 모방한 기술로 정의할 수 있다. 철학적 관

점에서는 인간의 일을 얼마나 수행할 수 있는지에 따라 강인공지능(Strong AI)과 약인공지능(Weak AI)으로 구분하기도 한다(Searle 1980). 강인공지능은 알고리즘을 설계하면 AI가 스스로 데이터를 찾아 학습하여 어떤 문제를 실제로 사고하거나 해결할 수 있는 인공지능으로 다양한 분야에 보편적으로 활용할 수 있다. 약인공지능은 미리 정의된 규칙을 이용해서 지능을 흉내내는 컴퓨터 프로그램으로 챗봇 등 특정 분야에서만 활용 가능하다. 미래학자 레이 커즈와일(Ray Kuzweil)은 이런 강인공지능을 ‘싱귤래리티(Singularity)’라는 용어로 설명했다. 싱귤래리티는 기술 발전이 이어지면서 인공지능이 인간을 뛰어넘는 순간을 뜻한다. 강인공지능의 출현 가능성에 대해 인공지능 전문가들은 “기우일 뿐”이라며 선을 긋지만 약인공지능만 해도 인간의 일자리를 상당수 대체할 것으로 예상되고 있다. 2016년 세계경제포럼에서는 인공지능의 발전에 따라 앞으로 5년간 15개국에서 710만 개 일자리가 사라질 것이라는 분석이 나왔다.

인공지능은 제조 부문에서 가전, 자율주행자동차, 스마트공장의 맞춤형 제품생산, 서비스 부문에서 투자예측, 건강진단, 드론을 이용한 무인배송, 창고관리 로봇, 자동 통번역, 무인콜센터, 인프라 부문에서 농산물 출하량 및 병충해 예측, 공공 안전관리, 환경오염 및 교통·기상 모니터링, 전력·수도·가스 자동조절, 발전소 이상징후 감지 등에 적용되고 있으며 가시적 효과가 매우 큰 특징을 보인다((KDB산업은행 2017). 기록관리 부문에서도 대량의 기록물을 빠르게 처리하여 업무 효율과 정확도를 향상시키거나, 기록의 내용을 분석하여 관련된 맥락 정보들을 연결해 주는 지능화된 서비스가 가능해질 것으로 예상되고 있다(국가기록원 2017).

인공지능 기술 중 자연어 처리(Natural Language Processing), 비전(Vision), 음성 인식(Speech) 등 기술의 성숙도가 높은 분야는 이미 다양한 상용 서비스들이 활용되고 있다. IBM,⁹⁾ 구글,¹⁰⁾ 마이크로소프트,¹¹⁾ 아마존 웹 서비

9) IBM Cloud – Watson, <https://www.ibm.com/cloud-computing/bluemix/ko/watson/>

10) Google Cloud AI, <https://cloud.google.com/products/machine-learning/>

스,¹²⁾ 페이스북,¹³⁾ 바이두,¹⁴⁾ SK C&C,¹⁵⁾ 네이버¹⁶⁾ 등은 국내외 IT기업들은 인공지능 플랫폼을 오픈소스로 공개하고 클라우드 서비스를 제공하고 있다. 또한 대부분의 클라우드 서비스들은 자사의 PaaS 플랫폼 안에서 빅데이터나 인공지능의 요소기술을 개별 API로 제공하고 이를 조합하여 제품을 만들도록 개발환경을 제공하고 있다. 이처럼 인공지능 기술은 향후 빅데이터, 사물인터넷 등과 결합하여 스스로 학습데이터를 선별하여 학습하고, 가전 및 스마트폰에 탑재되어 지능화된 서비스로 발전될 가능성이 매우 크다.

4) 사물인터넷(IoT)

사물인터넷은 기존 센서 기술을 기반으로 사물에 다양한 IT 기술을 적용하고 유·무선 네트워크 환경을 구성하여 사람-사물-서비스가 상호 지능적 협력 관계를 형성하는 네트워크를 의미한다(강정호 외 2015). 사물인터넷은 M2M(Machine to Machine), 유비쿼터스, NFC(Near Field Communication) 등 기존의 기술로부터 파생되었다. 교통카드, 바코드, ATM 기기 등에 활용되다가 최근 네트워크 인프라 발달과 소형 기기 증가에 따라 사람-사물-서비스 간 통신이 가능한 사물인터넷 기술로 진화하였다. 수년 간 스마트홈과 스마트팩토리에 많은 투자가 이루어지며 비용절감, 생산효율성 효과를 입증하였고, 이를 통해 제4차 산업혁명의 도래에 일조하고 있다.

사물인터넷의 핵심 기술은 센서, 네트워크 인프라, 데이터 공유 플랫폼, 빅데이터 처리 및 보안 기술이다(김재생 2015). 센서 분야는 반도체, MEMS 등의 제조기술이 접목되며 스마트센서 개발이 활발히 진행 중이다. 센서의

11) Microsoft AI, <https://www.microsoft.com/en-us/ai/>

12) AWS AI, <https://aws.amazon.com/ko/amazon-ai/>

13) Facebook AI Research(FAIR), <https://research.fb.com/>

14) Baidu-research, <https://github.com/baidu-research/>

15) AIBRIL, <https://www.aibril.com/>

16) Naver Clova, <https://clova.ai/ko/ko-platform.html/>

종류는 위치, 차원센서, 메카니컬, 다이내믹 센서, 압력, 밀도, 부피 센서, 플로우&레벨 센서, 화학&환경 센서, 전기&자기 센서, 온/습도, 날씨 센서, 비전, 이미지, 식별 센서, 신체&사람, 메디컬 센서, 광&방사선 센서, 바이오 센서, 음향&초음파 센서 등 보통 10~11개 분야로 구분된다.¹⁷⁾

사물인터넷은 특히 전통적 제조 공장에 적용되어 비용절감 및 생산효율화를 달성하고 있다. 맞춤형 주문생산을 통한 공장 내 재고의 최적화, 안전 사고 예방 등이 그 사례이다. 서비스 부문에서는 택배나 유통의 물류운송 최적화, 자율주행, 네비게이션, 보안 등에 활용되고 있다. 인프라 부문에서는 가정이나 빌딩, 도시의 에너지관리, 보안·안전관리, 교통통제 등에 활용될 수 있다.

맥킨지는 2025년까지 사물인터넷 데이터를 통해 얻은 인사이트로 창출 가능한 경제적 가치를 11조 달러로 전망했다(McKinsey & Company 2015). 가트너의 전망에 의하면 2020년에는 사물인터넷 하드웨어 시장 규모가 3조 달러에 육박할 것이라고 한다(Gartner 2017). 이처럼 사물인터넷은 거의 모든 영역에 적용되어 산업을 발전시키고 일상생활과 업무환경을 변화시키고 있다. 앞으로는 센서로부터 수집한 데이터를 인공지능, 빅데이터, 로봇틱스 등과 연계하는 형태로 발전하며 지능형 서비스의 핵심 인프라가 될 확률이 크다.

3. 제4차 산업혁명 기술의 기록관리 영향

1) 클라우드 컴퓨팅 환경에서의 기록의 특징

우리나라의 공공기록물 생산 인프라와 업무 프로세스는 이미 클라우드

17) 센서 인사이트. “센서 분류”. <https://www.sensor-insight.com/-sensor-classification->

컴퓨팅 환경으로 옮겨가고 있다. 정부는 2014년부터 자체(Private) 클라우드인 G-클라우드 전환을 가속화하고 있고, 민간 클라우드 이용을 확대하기 위해 클라우드 친화적으로 제도를 개선하고 있다. 행정자치부와 국가기록원은 2016년부터 정부업무관리시스템인 온나라 시스템과 RMS의 소프트웨어를 통합하는 SaaS 프로젝트를 진행하고 있다. 공공기록의 생산과 관리시스템이 클라우드 기반으로 변화하고 있는 것이다.

클라우드 환경에서 업무를 수행할 경우 변화되는 기록 생산의 양상은 다음과 같다.

첫째, 기록물의 양이 급증하고, 대용량 기록이 많아지며 유형이 다양화될 것이다. 이는 종이에서 전자기록으로 생산환경이 변화할 때의 영향과 비슷한 맥락이다. 전자기록 생산환경이 레거시(Legacy) 시스템에서 클라우드 시스템으로 한번 더 변화하면서 이전보다 좀 더 신속하고 편리하게 기록을 생산하고 유통하게 될 것이다. 출장 중에도 클라우드 시스템에 접속하여 기록을 생산할 수 있게 되고 문서유통센터 용량이 확장되면서 이전보다 기록의 생산량이 많아지고 대용량 기록의 생산이 증가할 것이다. 여기에는 클라우드 스토리지 비용이 낮아지는 것 또한 영향을 미친다. 동영상 등 멀티미디어 형태의 정보 이용을 선호하는 대중들의 요구 또한 점차 반영하여 이러한 기록들을 더 많이 생산하게 될 것이다. 기록의 유형이 다양화해진다는 의미는 기록의 포맷 뿐만 아니라 소통한 기록, 분석한 기록 등 공무원들의 업무형태 변화를 반영한 새로운 기록이 생산된다는 것이다. 향후 클라우드 업무환경이 정착되면 클라우드 상의 빅데이터 분석, 인공지능 서비스, 사물인터넷 등 다양한 PaaS, SaaS 서비스들을 업무와 연계하여 활용하게 될 것이다. 이럴 경우 각 서비스의 결과물로 생산된 분석 로그, 시각화 결과물, 센서 데이터 등 새로운 유형의 기록이 생산된다. 또한 기관 내외의 업무협력을 위해 과제별 업무공간 등에서 소통한 기록 또한 추가될 것이다. 자료 공유를 위한 블로그, 소모임의 토론 메모 등이다.

둘째, 개방형 문서포맷의 생산이 증가할 것이다. 온나라 시스템에서 붙임 첨부문서의 형식은 업무PC에서 작성한 파일포맷 그대로 첨부하도록 기존 방식을 유지하지만, 기안문과 본문은 ODF(Open Data Format) 형식으로 작성하도록 되어 있다. 향후 부처간 협업과 문서 유통이 활발해지는 시점에서는 보다 원활한 활용을 위해 ODF 뿐만 아니라 PDF, HTML 등 표준 기반의 개방형 문서포맷 생산이 붙임 첨부문서까지 확대될 가능성이 있다.

셋째, 링크 유통이 활발해질 것이다. 클라우드 온나라 시스템에서는 문서, 메모보고 등을 파일로 첨부하지 않고 콘텐츠별 공유 URL을 전송하여 정보를 공유할 수 있게 되었다. 따라서 기록과 메타데이터의 링크가 지속적으로 작동되도록 하는 것이 중요해진다. 생산 측면에서는 등록 과정에서 고유한 URL 체계를 마련하여 자동으로 생성하거나 깨진 URL을 기록관리시스템의 원본 기록으로 리디렉션하는 프로세스가 마련되어야 할 것이다.

넷째, 사본이 많아질 것이다. 기록 생산자나 관리자, 클라우드 서비스 제공자는 사고에 대한 복구의 일환으로 사본을 만들어 여러 개의 저장소에 분산시켜 저장한다. 또한 공공데이터 개방 확산에 따라 서비스용 사본을 별도의 저장소에 두고 관리하기도 한다. 이러한 경우 실제로 위기가 발생했을 때 복구에는 유리하겠지만 기록의 진본성, 무결성, 신뢰성에는 혼선이 오게 된다.

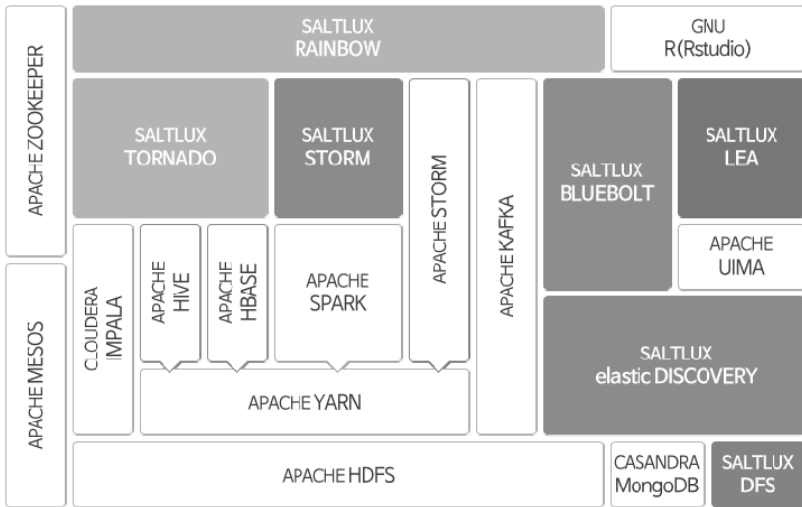
2) 빅데이터 기술 도입에 따른 기록의 특징

공공 영역에서도 빅데이터 분석을 업무에 활용하기 시작했다. 2013년부터 행정자치부는 공공 빅데이터 사업을 준비하였고 2015년부터 아파트 관리비 비리 적발 등 국가현안 및 사회적 이슈에 대해 선제적으로 대응하려고 하고 있다. 정부통합전산센터는 2016년 5월 빅데이터 전담조직을 신설하고 ‘데이터활용공통기반시스템 운영지침’을 수립하였다. 현재 공공기관의 공무원들은 빅데이터 분석을 운영책임관에게 의뢰하여 분석결과물을 제공받거나

‘데이터활용공통기반시스템(이하 해안시스템)’을 통하여 직접 분석을 수행할 수 있다.

해안 시스템은 국내 기업의 빅데이터 분석 플랫폼과 하둡(Hadoop), 아파치(Apache) 빅데이터 생태계의 오픈소스 제품들을 결합하여 만들어졌다. 이 솔루션은 시맨틱 마이닝 엔진(Discovery), 지식자산 수집엔진(Tornado), 스트림 빅데이터 분석 엔진(Bluebolt), 지식 저장소 엔진(Storm), 지식 시각화 엔진(Rainbow), 자연어 처리 엔진(LEA) 등으로 구성되어 있다. 핵심 기반기술인 자연어 처리, 기계학습, 분산병력/인메모리 분석 등을 이용하여 80억 건 이상의 소셜 빅데이터와 공공 빅데이터를 수집하고 지식 베이스를 구축하고 있다.

〈그림 2〉 해안 시스템(Big O) 구성도



※출처 : 솔트룩스 홈페이지

빅데이터 기술 도입이 국내 기록관리 환경에 미치는 영향을 거시적으로

파악하기 위해서는 데이터 관련법의 흐름을 살펴볼 필요가 있다. 해안 시스템 또한 데이터 관련법을 근거로 만들어진 것이기 때문이다. 현재까지 국가기록원은 행정정보데이터세트의 관리와 활용에 대한 주도권을 형성하지 못하고 있다. 그러는 동안 시민 등 외부의 요구에 의해 ‘공공데이터의 제공 및 이용활성화에 관한 법률(이하 공공데이터제공법)’이 시행되었고, 공공기관 내부의 의사결정방식 개선 요구에 의해 ‘데이터 기반 행정활성화에 관한 법률(안)(이하 데이터기반법)’이 입법예고되었다. 이처럼 데이터세트 기록에 대해서만은 별도의 법이 만들어져 여러 주체가 대응하고 있다. 방대한 양의 공공기록이 데이터세트로 누적되어 다양한 목적으로 활용될 가능성이 있기 때문이다.

우선 공공기록물관리법과 데이터 관련법들의 ‘데이터’에 대한 표현이나 언급하고 있는 기술, 공개 및 공유 방식 등을 통해 차이점을 인식할 필요가 있다. 두 데이터 관련법은 ‘행정정보’나 ‘데이터세트’란 표현 대신 ‘공공데이터’, ‘데이터베이스’, ‘메타데이터’ 등의 표현을 사용하고 있다. 기술 측면에서 공공데이터제공법은 Open API, 표준데이터(CSV, JSON, RDF, XLS, XML), LOD 등에 기반을 두고 있다. 데이터기반법은 ‘인터넷 데이터의 수집’, ‘데이터분석기반시스템’ 등의 표현으로 미루어 볼 때 빅데이터 분석 기술과 밀접하게 연관되어 있음을 알 수 있다. 데이터기반법이 다루는 데이터의 범위는 공공데이터와 인터넷 상의 외부 데이터를 모두 포괄하고 있다. 이와 같이 두 데이터법은 공공과 민간의 데이터 자원을 효율적으로 이용할 수 있도록 품질을 향상시키고 수집, 분석, 제공기반을 마련하는 것을 목표로 하고 있음을 알 수 있다.

이처럼 빅데이터 기술과 국내 데이터 관련법들의 흐름은 공공기록물 생산과 관리, 활용에 새로운 관점이 필요함을 시사하고 있다.

첫째, 공공기록관리의 범위가 민간 빅데이터까지 확장된다. 이제 공공영역에서도 업무상 필요 때문에 더욱 많은 빅데이터를 생산하고 활용할 것이다. 여기에는 민간의 산업별 오픈데이터와 통계지표, 소셜 빅데이터, 해

외의 다국어 언어자원까지도 포함된다. 정부가 지향하는 데이터 기반 행정을 지원하기 위해서는 기록관리의 영역이 민간과 해외의 빅데이터 관리까지 확장되어야 한다. 공공 영역에 들어온 빅데이터를 기록으로서 잘 유지하고 관리하여 향후 또 다른 빅데이터 분석에 활용할 수 있게 하는 것이 중요해지기 때문이다.

둘째, 기록의 선별과 평가에 미래의 활용가치 비중을 늘려야 한다. 빅데이터의 속성 중 양(Volume)이나 다양성(Variety) 측면에서는 다양한 기록을 많이 모아서 향후 특정 데이터와의 연계(linkage)될 가능성을 늘리는 전략이 바람직하다. 이 관점을 기록의 평가와 선별에 적용해 보면 기록에 담긴 현재의 가치와 함께 잠재적인 활용가능성을 좀 더 고려해야 한다. 이를 전통적 기록물 평가 업무의 무용론으로 오해하기보다는 기록을 생산하고 활용하는 방식의 변화로 인식할 필요가 있다. 개별 기록에 수록된 가치보다 연결된 데이터를 통해 더 큰 맥락(Insight)을 만들어내는 것이 빅데이터가 추구하는 가치이다. 어찌 보면 빅데이터 기술은 기록의 생산 맥락을 사회적 차원으로 이해하려는 거시평가 이론의 유용한 실행도구가 될지도 모른다.

셋째, 기록의 품질을 높여야 한다. 빅데이터의 정확성(Veracity)이나 가변성(Variability) 측면에서는 메타데이터의 품질이나 정형화된 데이터 구조 등 기록의 품질이 중요해진다.¹⁸⁾ 이 또한 선별평가 정책에서 미래의 활용가치 항목으로 포함될 수 있을 것이다. 기록의 품질을 높이기 위해서는 가급적 분석과 활용에 용이하도록 공공데이터 표준을 준수하여 데이터의 구조를 정하거나 맥락정보를 충분히 기술해야 한다. 또한 행정정보시스템이나 기록시스템에 데이터 품질관리를 위한 기능, 그리고 XML, JSON, RDF 등 공공데이터 표준포맷으로 내보내는 기능이 추가되어야 한다. 향후 이러한 기

18) IBM은 가트너가 정의한 빅데이터의 3대 속성에 정확성(Veracity)을 추가하여 4V를 정의하였고, 브라이언 홉킨스(Brian Hopkins) 등은 가변성(Variability)을 추가하여 4V를 정의하였다(출처: 위키피디아).

술들을 적극적으로 도입하게 되면 부처가 보유한 공공기록이 보다 쉽게 분석, 유통되어 기록물관리기관의 기록관리 활동이 시민, 공무원, 기업들로부터 그 가치를 인정받는 계기가 될 것이다.

넷째, 빅데이터 시대를 위한 기술적 준비가 필요하다. 데이터 관련법들은 현재의 기록시스템으로 포괄되지 않는 데이터베이스 품질관리, 파싱, 빅데이터 수집, 분석, 시각화, 저작권 관리와 분쟁조정 등의 새로운 영역과 기술을 요구하고 있다. 또한 데이터를 수집, 분석, 제공하기 위한 별도의 서버 등 인프라도 마련되어야 한다. 기록시스템에 이 모든 기능들이 필요하다기보다는 기록을 이용하는 환경이 변했으니 받아들일 항목과 연계할 대상을 검토해야 한다는 것이다.

3) 인공지능 기술 도입에 따른 기록화 전략의 변화

인공지능 기술은 이미 일상에 스며들었다. 스마트폰의 음성비서, 인공지능 스피커, 자율주행자동차, 챗봇 등 인공지능 기술은 이제 특정 분야의 기술이라기보다 IT 전반의 인프라 역할을 하고 있다. 공공 분야에도 민원상담이나 위험관리 등 점차 인공지능 기술을 이용한 업무 방식이 도입되고 있다. 기록관리 분야에서는 대량의 기록을 한꺼번에 처리하여 업무 효율성과 정확성을 향상시키는 방향으로 관심과 기대가 모아지고 있다. 하지만 기록관리 측면에서의 핵심적인 변화는 인공지능이 수행하는 업무에 대한 설명 책임성과 투명성을 보장하는 것이 어려워진다는 것이다. 이에 대한 구체적인 내용을 몇 가지 이슈로 구분하면 다음과 같다.

첫째, 인공지능 에이전트의 역할과 책임(R&R)을 설계해야 한다. 최근 유명한 윤리 논쟁인 ‘트롤리 딜레마(Trolley Problem)¹⁹⁾’가 자주 언급되고 있

19) 트롤리 딜레마(Trolley Problem)는 사람들에게 브레이크가 고장 난 트롤리 상황을 제시하고 다수를 구하기 위해 소수를 희생할 수 있는지를 판단하게 하는 문제 상황을 가리키는 말이다(출처: MIT Technology Review, 2015. <https://www.technologyreview.com/s/542626/why-self-driving-cars-must-be-programmed-to-kill/>).

다. 실제로 자율주행차가 사고를 냈을 경우 탑승자와 제조사, 개발자 중 누구에게 책임이 있느냐를 놓고 법적·윤리적 논쟁이 치열하게 이루어지고 있다. 트롤리 딜레마에서 보듯이 사람 대신 인공지능 에이전트가 수행하는 업무가 많아질 수록 기록을 통해 업무 과정을 설명하는 것이 더욱 어려워지고 중요해질 것이다. 예를 들어 챗봇(Chatbot)이 민원 상담을 수행했을 경우 이 상담일지의 내용에 대한 책임이 담당 공무원에게 있는지, 대리한 챗봇에게 있는지부터 혼선이 발생한다. 또한 챗봇의 지위가 공무원인지도 현재로서는 명확하지 않다. 챗봇이 어떤 내용을 상담할 수 있고 책임질 수 있는지, 어떤 과정으로 업무를 수행해야 하는지를 설계하지 않으면 이러한 혼란은 가중될 것이다. 싱귤래리티가 도래하는 시점이 2030년으로 예상되는 현 시점에 인공지능을 단지 업무지원의 도구로만 인식하는 것은 순진한 생각일 수 있다.

둘째, 인공지능 에이전트의 역할과 책임을 알고리즘에 적용시켜야 한다. 트롤리의 딜레마에서 보듯이 매 순간 인공지능 에이전트는 판단을 해야 한다. 인공지능이 판단하는 근거는 내장된 알고리즘이다. 도덕적 행위자(AMA, Artificial Moral Agent)로서의 인공지능이 출현하기 전까지는 파괴적 행위나 비윤리적 행위 등 특정 행위를 수행할 수 없도록 설계하여 알고리즘에 적용하는 것이 바람직하다. 구체적으로 어떤 행위를 할지 말지를 결정하기 위해서는 상당한 기간의 논의와 합의가 필요하다. 실제로 MIT 미디어 랩의 ‘모럴 머신(Moral Machine)²⁰⁾ 사이트에서는 트롤리 딜레마의 다양한 상황에 대해 투표하게 하고 사람들이 원하는 질문을 올리도록 하면서 사회적 윤리 기준을 확인하는 작업을 진행하고 있다. 인공지능 에이전트의 자율성과 인간의 통제권을 어떻게 조화시킬 것인지는 앞으로도 지속적으로 논의되어야 한다.

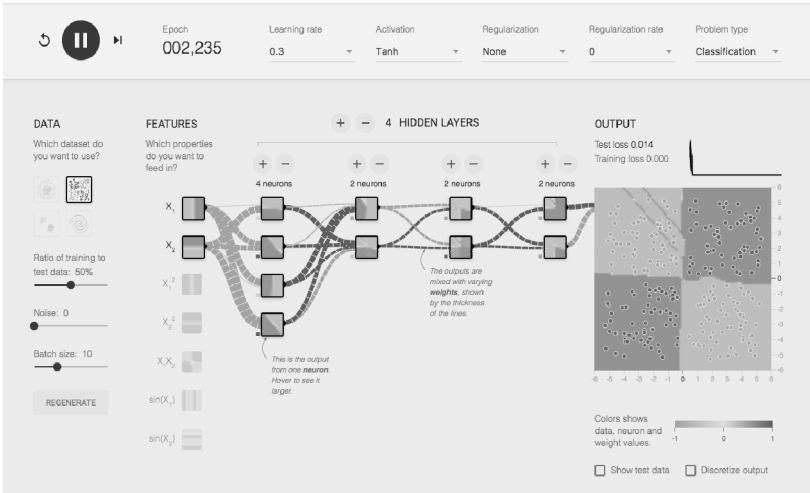
셋째, 인공지능 알고리즘의 기록화가 필요하다. 인공지능이 수행하는 업

20) Moral Machine, <http://moralmachine.mit.edu/hl/kr/>

무가 많아질 때를 대비하여 기록의 생산맥락을 확보하기 위해서이다. 100년 뒤 챗봇이 생산해 낸 상담기록을 이해하기 위해서는 챗봇에 대한 전거 레코드가 필요하다. 이 레코드에는 챗봇의 제조사, 알고리즘 유형과 특징, 학습능력, 학습과정, 학습데이터, 품질과 성숙도 등 챗봇의 정체를 충분히 드러낼 수 있는 정보가 담겨야 한다. 또한 챗봇의 진화 과정에 대한 디스크립션도 필요하다. 각 세대나 버전별 특징, 기술적 성숙도에 대한 설명이 추가되어야 한다. 챗봇은 인스턴스만 계속 생성하면 수천 개의 에이전트가 동시에 업무를 수행할 수 있으므로 확산 가능성이 높고 예상보다 많은 양의 기록을 생산할 수 있다. 따라서 챗봇의 역할과 책임, 전거레코드 등이 DB로 명문화되어 자동으로 획득되어야 한다.

넷째, 인공지능 알고리즘의 투명성을 확보해야 한다. 투명성은 인공지능 알고리즘의 기록화만으로 해결되지 않는 문제이다. 최근 인공지능 알고리즘은 학습한 내용을 이해하는 머신러닝에서 스스로 진화하는 딥러닝으로 진화하고 있다. 딥러닝으로 훈련된 알파고의 바둑 솜씨는 이미 인간을 앞서고 있다. 하지만 딥러닝 알고리즘의 작동 방식을 설명하거나 검증하는 것은 상당히 어렵다. 딥러닝은 뇌의 뉴런과 유사한 복수의 정보 입출력 계층을 활용해 층마다 다른 층위의 데이터를 학습하며 경험적으로 검증된 방법을 사용하기 때문에 종종 블랙박스로 비판받기도 한다. 깊은 신경망일수록 외부로 나타나지 않는 은닉층(hidden layer)이 많기 때문에 작동 방식을 이해하기는 더욱 어려워진다. 아래 그림과 같이 텐서플로우 테스트를 통해 2차원의 단순 분포 데이터셋을 분류(classification)하는 딥러닝 알고리즘을 구현할 때 2개의 입력층과 2개의 출력층, 4개의 은닉층이 필요함을 확인하였다. 복잡한 분포의 데이터셋을 처리하기 위해서는 훨씬 많은 은닉층이 필요해진다.

〈그림 3〉 텐서플로우의 뉴럴 네트워크를 이용한 데이터 분류 테스트 결과



※출처 : <http://playground.tensorflow.org/>

이미 공공 영역에서도 민원상담 업무에 챗봇을 도입하려 시도하고 있다. 대구시는 최근 AI 및 빅데이터 기반 지능형 민원상담 시스템을 구축하고 있으며²¹⁾, 인천국제공항공사²²⁾, 중소기업은행 등도 챗봇 도입 사업을 발주했다. 인공지능 기술이 발전할수록 기록의 투명성과 설명책임성에 대한 위험은 커질 것이다.

21) 매일경제 - “마인즈랩, 대구시와 지능형 민원상담 서비스 개발 실시”. <http://news.mk.co.kr/newsRead.php?year=2017&no=645112>

22) 나라장터 - 인천공항 챗봇 도입 사업. <http://www.g2b.go.kr/pt/menu/selectSubFrame.do?framesrc=pt/menu/frameTgong.do?url=http://www.g2b.go.kr:8101/ep/tbid/tbidList.do?taskClCd=&bidNm=%C3%AA%BA%BF&searchDrType=1&fromBidDt=2017/04/15&toBidDt=2017/10/16&fromOpenBidDt=&toOpenBidDt=&radOrgan=1&instNm=&area=®Yn=Y&bidSearchType=1&searchType=1>

4) 사물인터넷 환경에서의 기록의 특징

사물인터넷은 점차 모든 디바이스나 사물에 적용되고 인공지능 등 신기술과 연계되어 거의 모든 영역의 업무를 변화시킬 것이다. 한 예로 보건복지부는 2016년 말부터 이동통신사와 함께 독거노인들의 집에 에너지미터를 설치하였다.²³⁾ 에너지미터에는 활동감지, 화재감지, 가스감지, 출입감지 등 4개의 스마트 센서²⁴⁾가 달려 있다. 사회복지사는 각 가정에 방문하지 않고도 독거노인의 활동 상태를 모니터링하며 실시간 데이터의 패턴을 통해 심정지나 독거사 등의 위험을 방지하게 된다. 사람이 하던 일을 사물인터넷을 통해 보다 효율적으로 정확하게 수행하게 되는 것이다. 이 경우 공격 서비스 향상에 대한 센서 데이터가 기록으로 축적된다. 이러한 데이터는 향후 주요한 빅데이터로 활용될 수 있다. 기록관리가 기록을 통해 업무의 투명성과 설명책임성을 구현하는 도구라면 이러한 기록의 맥락을 잘 포착해야 할 의무가 있다. 2016년 시점에 공공기관인 보건복지부가 어떤 알고리즘이나 비즈니스 로직을 이용하여 복지 서비스를 향상시켰는지의 맥락을 기록화해야 한다. 예를 들어 전력 사용량 센서 데이터의 패턴에 특정 임계치를 설정하여 그에 도달하면 이상 징후로 인식하게 하는 구현방식과 그 결과물로서의 데이터셋 모두 맥락 정보에 해당한다.

위 사례에서 보듯이 사물인터넷이 본격화되면 스트리밍 데이터의 양이 급증하게 된다. 수천, 수만 개의 센서로부터 작은 크기(KB 단위)의 스트리밍 데이터가 동시에 지속적으로 전송되기 때문이다. 네트워크에 연결된 사물들이 저마다 데이터를 생성하는 이러한 환경은 본격적인 빅데이

23) 경향신문 - “독거노인 돌보기도 IoT시대... 누가 날 지켜주는 것 같아 든든”. http://news.khan.co.kr/kh_news/khan_art_view.html?artid=201612182051005

24) 스마트 센서는 자체적인 처리 기능을 수행하는 마이크로프로세서를 내장하고 있어 신호를 디지털로 변환 후 더욱 지능화된 정보를 추출할 수 있다.

터 시대가 된다는 것을 의미하기도 한다. 따라서 실시간으로 데이터가 이동하는 동안에 유의미한 정보를 추출해 내는 기술 확보가 중요한 과제가 된다. 스트리밍 데이터가 증가하면 좀 더 기술에 종속적인 기록관리 환경으로 변화할 수밖에 없다. 종전의 기록관리에서처럼 아키비스트가 수집할 기록을 선별하고 평가할 시간이 없기 때문이다. 스트리밍 데이터 분석 시스템은 데이터의 흐름을 모니터링하면서 일정한 패턴이나 특이사항을 탐지하고 잠재적인 문제(전력 감지기가 작동하지 않음 등)를 경고하거나 도움이 되는 제안(현재의 속도로 전력량이 감소하면 24시간 후 해당 가정에 방문이 필요함 등)을 하게 된다. 따라서 사물인터넷 환경에서 담당자의 업무 프로세스는 미리 정해진 규칙(전력량 기준, 비상대책)을 분석 시스템의 규칙 엔진(Rule Engine)에 적용하고 모니터링하는 방식으로 변화하게 된다. 센서들이 생산한 스트리밍 데이터는 이를 처리하는 행정정보시스템에 저장되는데, 실시간성이 중시되는 이러한 데이터셋을 얼마나 오래 보존할지, 유의미한 데이터만을 보존할지 등의 판단 또한 필요하다.

또한 사물인터넷은 빅데이터와 클라우드 융합을 통해 가치가 증대된다. 현실 세계의 정보들은 단지 디지털 수치로 변환되는 것만으로는 사용자에게 필요로 하는 정보를 제공할 수 없기 때문이다. 그리고 센서를 통해 축적된 데이터는 로컬에 저장되기보다는 클라우드에 저장될 때 가치가 올라간다. 데이터가 여러 디바이스에서 재활용될 수 있어야 하기 때문이다. 마지막으로 모든 사물이 연결되는 환경은 사생활 침해나 보안의 문제를 발생시킬 수 있다. 일상의 일거수 일투족이 기록되기 때문이다. 사물인터넷이 보다 산업화되고 공공 업무에 정착하기 위해서는 보안 및 정보보호 정책 등이 동시에 진행되어야 한다.

5) 시사점

이상 살펴본 바에 의하면 신기술 환경에서 변화되는 기록의 특성은 다음과 같다.

첫째, 기록의 양이 급증하고 유형이 다양화해진다. 클라우드의 유연한 특성과 공무원들의 업무형태 변화, 사물인터넷 환경, 인공지능 에이전트의 출현은 새로운 유형의 기록 생산을 증대시킬 것이다.

둘째, 데이터 품질 확보가 필요하다. 빅데이터 분석과 인공지능 학습의 효율을 위해서는 표준포맷 구조로 만들고 내보낼 수 있어야 하며 정확하고 신뢰할 수 있어야 한다.

셋째, 생산맥락의 기술이 중요해진다. 기존의 디스크립션 체계로는 인공지능 로봇이나 빅데이터 인사이트, 센서가 생산한 기록의 맥락을 설명하기 어렵다. 이러한 현상은 신기술이 융합하여 가치를 만들어내는 특성으로 인해 가속화되고 있다.

4. 제4차 산업혁명 기술의 기록관리 적용 방안

1) 클라우드 컴퓨팅 기술의 기록관리 적용 방안

공공기관이 우선 도입해야 할 클라우드 기술은 IaaS이다. 인프라 도입을 통해 향후 PaaS, SaaS 환경의 효율적인 기록 생산과 유통 환경으로의 전환을 준비해야 한다. 특히 공공기관의 경우 민간 클라우드를 우선 도입하도록 노력해야 한다. 중앙행정기관은 2016년에 절반 이상인 47개 부처 577개 업무를 G-클라우드로 전환했다. 하지만 중앙부처를 포함하여 지방자치단체와 공공기관의 민간 클라우드 도입은 거의 이루어지지 않고 있다. 현행 서

버를 민간 클라우드 서비스 업체들의 데이터센터로 완전히 분리하기 어렵다는 점이 주요 이유이다. 그 외에도 대기업 참여 제한으로 KT 등 3개의 업체만 보안인증을 받은 상태라서 믿고 쓸 수 있는 솔루션이 부족한 실정이다. 또한 현재 민간 클라우드는 IaaS만을 대상으로 하여 글로벌 IT 기업들이 제공하는 다양한 SaaS 서비스를 이용할 수 없는 실정이다. 행정자치부는 올해 서비스 인증기관을 5개로 늘린다는 계획이지만 아마존 웹 서비스 등 해외 클라우드 사업자들은 불필요한 절차와 시스템 및 데이터의 물리적 위치를 국내로 제한한 것에 반발하며 보안인증을 거부하고 있다.²⁵⁾ 국내와 국외 사업자의 보안인증 기준을 달리 하는 호주의 사례를 참고할 필요가 있다(안대진, 2015). 2017년 10월 기준 호주 재무부에서 인정하고 있는 클라우드 서비스 사업자는 185개에 달한다.²⁶⁾

보안인증 외에 현행 G-클라우드의 유연성이나 상호운용성은 상당히 낮은 수준이다. 올해 진행되고 있는 “2017년 제3차 범정부 정보자원 통합구축 HW1사업”의 제안요청서에 따르면 이전 대상인 8개 기관 17개 업무시스템에 대해 동일한 소프트웨어 규격을 요구하고 있다. 이기종 시스템이 G-클라우드로 쉽게 이전할 수 있도록 다양한 OS와 WEB, WAS, DBMS를 제공하여 클라우드 컴퓨팅의 취지를 살려야 한다.

그 외에 당장 활용할 수 있는 클라우드 기술은 빅데이터 분석 서비스이다. 행정자치부에 구축된 해안 시스템을 업무에 적극적으로 활용해 볼 필요가 있다.

추후 적용할 클라우드 기술은 PaaS와 SaaS 형태의 서비스들이다. 빅데이터, 인공지능, 사물인터넷, 데브옵스(DevOps), 도커(Docker), 마이크로서비스(Microservices) 등 다양한 신기술들을 PaaS나 SaaS 형태로 G-클라우드에

25) 보안뉴스, “클라우드 보안인증 두고 정부 vs. 글로벌 기업 기싸움?”, <http://www.boannews.com/media/view.asp?id=55045>.

26) aus tender, Standing Offer Notice View, <https://www.tenders.gov.au/?event=public.son.view&SONUID=745895FF-E769-50C9-D8607CECECE179B4>

점차 도입하는 것이 바람직하다. 특히 도커와 마이크로서비스 기술은 국가 기록원이 설계하고 있는 차세대 AMS의 모듈화된 구조를 구현하고 클라우드를 통해 효과적으로 배포하기 위해 필요한 기술이다. 참고로 영국 정부는 2017년 5월, 공공 부문 클라우드 서비스의 조달 과정을 단순화하고 혁신을 가속화하기 위해 정부 클라우드인 G-Cloud 9²⁷⁾에 도커 컨테이너 기술을 도입하였다. 또한 이에 대한 교육 및 전문서비스 조직(PSO, Professional Service Organization)을 구성하였다.²⁸⁾ 영국의 공공기업들은 구식의 시스템 아웃소싱 모델을 벗어나 온라인 스토어²⁹⁾를 통해 필요한 클라우드 솔루션을 쉽게 조달할 수 있다. 소프트웨어 기업들은 도커 컨테이너 기술 덕분에 복잡한 추가 개발 없이 자사의 솔루션을 컨테이너 형태로 G-Cloud 9에 등록하여 판매할 수 있게 되었다. 기록관리 업무에 필요한 유용한 소프트웨어들을 손쉽게 도입하여 실무 효율성을 제고해야 한다. 표준RMS의 기능 미비 등으로 인해 보다 편리한 기록관리 서비스를 내외부에 제공하지 못하는 기존의 한계를 극복해야 한다.

2) 빅데이터 기술의 기록관리 적용 방안

모든 공무원은 행정자치부에서 제공하는 빅데이터 솔루션인 헤안 시스템을 활용할 수 있다. 현재 제공하는 분석 결과는 제한적이지만 고도화 사업이 꾸준히 이루어지고 있는 만큼 현 시점에서는 데이터 분석을 통한 문제 해결법을 익히고 점차 발전시켜 나갈 수 있을 것이다. 향후에는 사물인터넷 인프라 기반의 실시간 데이터 분석 등과 연계해 나가야 한다. 빅데이터 분석은 공공기록과 외부 자원을 연결하여 소장 자료의 가치를

27) Crown Commercial Service – G-Cloud 9. <https://ccs-agreements.cabinetoffice.gov.uk/contracts/rm1557ix>

28) Docker blog – “Docker Enterprise Edition Now On G-Cloud 9 Framework.”. <https://blog.docker.com/2017/06/docker-enterprise-edition-now-g-cloud-9-framework/>

29) GOV.UK Digital Marketplace. <https://www.digitalmarketplace.service.gov.uk/g-cloud>

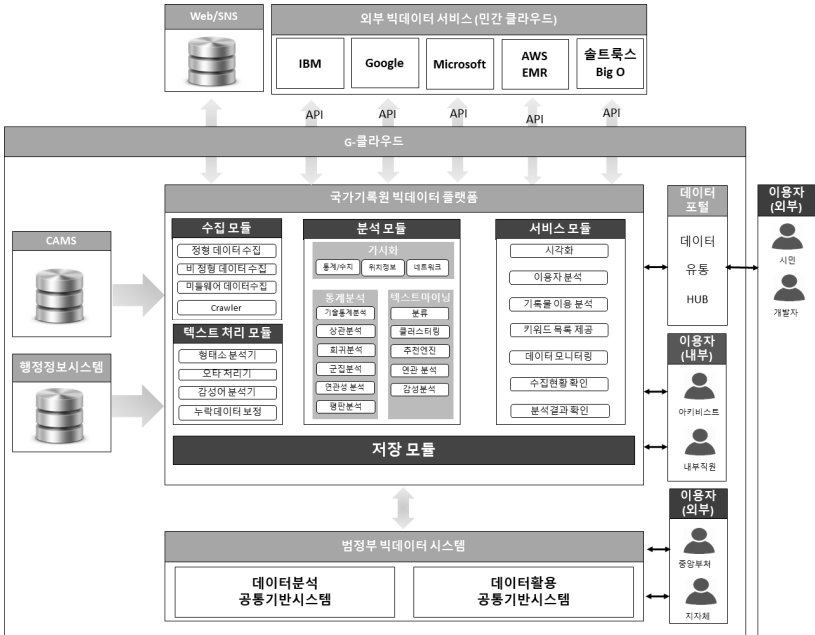
높일 것이다.

이처럼 공무원들이 해안 시스템을 이용하여 정책 입안이나 의사 결정에 빅데이터를 활용하게 되면 분석 결과 기록에 대한 관리 조치가 추가되어야 할 것이다. 해안 시스템은 키워드 분석과 언론 스크랩, 이동인구 분석 등을 제공한다. 분석 결과는 워드클라우드, 분석 대상 간의 관계 그래프 등 시각화 결과물 3종이다. 이 분석 결과물로 인해 실제로 정책이 입안되거나 의사 결정 과정에 참고가 되었다면 이 역시 별도의 기록 건이나 붙임.첨부 기록으로 등록되어야 한다. 또한 이러한 형태의 기록들에 대한 별도의 디스크립션 항목을 만들어야 한다. 예를 들어 분석에 사용된 플랫폼명, 플랫폼의 버전과 기술적 특징, 분석 대상 키워드 및 데이터세트, 분석 방법, 분석 수행자 등이다. 해안 시스템에 대한 디스크립션이나 전거데이터 또한 별도로 필요할지 모른다.

클라우드 컴퓨팅이 PaaS와 SaaS 레벨로 확장되고 본격화되면 국가기록원 등 영구기록물관리기관에서는 독자적인 빅데이터 플랫폼을 구축을 검토할 수도 있다. 만약 국가기록원이 빅데이터 분석 플랫폼을 구축한다면 정부통합전산센터가 운영하는 ‘범정부 빅데이터 공통기반’, 한국정보화진흥원(NIA)이 관리하고 있는 ‘빅데이터 분석 활용센터’와 빅데이터의 수집, 분석, 제공 기반을 연계해야 한다. 이를 통해 공공영역의 주요 빅데이터 프로바이더 또는 매개자 역할을 수행하고 공통기반에 존재하는 웹, SNS, 민간데이터까지를 포괄하여 분석 범위를 넓혀가야 한다. 국토부, 기상청, 서울시, 경기도 등 이미 빅데이터 플랫폼을 보유한 기관들과의 공조나 역할 구분 또한 필요할 것이다.

빅데이터 분석 기술을 통해 행정정보데이터세트를 이관과 관리의 대상이 아닌 연계와 활용의 대상으로 재설정함으로써 기록물관리기관들은 새로운 차원으로 기록관리를 도약시키는 계기를 마련할 수 있을 것이다.

〈그림 4〉 국가기록원 빅데이터 플랫폼 구성(안)



3) 인공지능 기술의 기록관리 적용 방안

인공지능의 연구분야 중 기록관리에 적용 가능한 부문은 비전(Vision), 음성 인식(Speech), 그리고 자연어 처리(Natural language processing)를 이용한 기계번역과 챗봇 등이다. 이러한 기술들은 텍스트, 사진, 영상 등 기록의 내용 분석과 처리에 유용하며 기술의 성숙도 또한 높다³⁰⁾. 구글 비전(Google Vision)과 IBM 왓슨에서 테스트해본 결과 사진 인식(비전), 음성 인식 등은 아주 뛰어난 수준이었다. 특히 구글 비전의 경우 등장인물인 우

30) 구글 클라우드(Google), IBM Bluemix 등 대표적인 클라우드 플랫폼들은 공통적으로 음성 인식, 자연어 처리, 비전 등의 인공지능 기술을 클라우드 서비스로 제공하고 있다.

리나라 대통령과 스포츠선수, 장소, 관련된 키워드를 제공하고 음란물, 속임수, 의료정보, 폭력성 등 유해성 정보까지 분석했다.

우선 비전은 이미지나 비디오 등의 시각 데이터를 처리하는 기술이다. 기록관리에 적용한다면 시청각 기록을 분석하여 인물, 장소 등의 정보를 추출하고 관련 기록과 연결하는 서비스가 가능해질 것이다. 정확도 향상을 위해서는 주요 인물이나 장소의 다양한 샘플 데이터를 선정하여 학습시키는 과정이 병행되어야 한다. 카네기멜론 대학교에서 만든 파이썬(Python)³¹⁾용 딥러닝 SVM(Support Vector Machine) 패키지인 오픈페이스(OpenFace)³²⁾를 이용하여 우리나라 대통령 5인의 사진을 각각 20장씩 학습시켜 테스트한 결과 60~70%의 정확도로 추론하는 것을 확인하였다. 더 많은 양질의 학습 데이터를 이용한다면 현장에서 충분히 활용 가능한 수준의 분석 정확도를 도출할 수 있을 것이다.

다음으로 음성 인식 기술 또한 기록관리에 유용하게 쓰일 수 있다. 이 기술은 녹취 안된 구술기록이나 회의 영상으로부터 텍스트를 추출하거나 시각장애인을 위한 기록 읽어주기 서비스를 제공하는 데 쓰일 수 있다. 음성 인식 기술의 한계는 한글 처리와 노이즈 문제이다. 대부분의 음성 인식 솔루션들이 영어보다 한글 문장의 처리에 상당히 취약하다. 기계 번역이나 음성 인식 등의 정확도 향상을 위해서는 한글 처리를 위한 말뭉치(Corpus) 개발이 지속되고 인공지능 플랫폼이 이를 학습해야 한다. 문화체육관광부와 국립국어원은 2018년부터 22년까지 “2차 세종계획”을 추진한다고 밝혔다³³⁾. 한국어 인공지능 기술의 발전을 위해 154억 7천만 어절의 말뭉치를 구축하는 데 5년 동안 175억 원의 예산을 투입하기로 했다. 이를 통해 국내의 자연어 처리 기반 인공지능 서비스 수준 또한 향상될 것이다.

31) Python, <https://www.python.org/>

32) OpenFace, <https://cmusatyalab.github.io/openface/>

33) 연합뉴스, “인공지능용 한국어 말뭉치 155억어절 구축…5년간 175억 지원”, <http://m.news.naver.com/read.nhn?oid=001&aid=0009591992&sid1=105&mode=LSD>

마지막으로 챗봇 서비스이다. 챗봇은 봇(bot)의 일종으로 명령방식에 따라 챗봇, 음성인식봇, 개인비서로 구분된다. 애플의 시리가 대표적인 음성인식봇이며 구글 나우(Now), 마이크로소프트의 코타나(Cortana)가 잘 알려진 개인비서 서비스이다. 기록관리에 적용한다면 내부 직원들의 기록관 업무에 대한 질의 처리, 기록물 검색에 대한 답변 등을 수행하는 챗봇 서비스를 상상해볼 수 있다. 챗봇의 기술요소는 패턴인식, 자연어 처리, 시맨틱 웹, 텍스트 마이닝, 상황인식 컴퓨팅 등이다. 3장에서 살펴봤듯이 이미 공공기관들이 챗봇 개발 사업을 시작했다. 챗봇 분야는 페이스북, 마이크로소프트, 구글, 텐센트, 네이버, 카카오톡 등 기존 메신저 제공 기업들이 생태계 선점을 위해 챗봇 플랫폼과 API를 오픈소스로 공개했기 때문에 기술장벽이 그리 높지 않은 편이다.

인공지능 기술들은 알고리즘 개발과 하드웨어 구축, 그리고 인공지능 학습에 상당한 자금과 기간이 소요된다. 무엇보다 이를 계획하고 실행할 전문인력을 배치하는 것이 쉽지 않다. 따라서 직접 인공지능 인프라를 구축하는 것과 구글, IBM, SK C&C 등이 제공하는 클라우드 기반 API 서비스를 활용하는 방안을 두고 효율성을 검토해 봐야 한다. 직접 인공지능 인프라를 구축할 경우에는 해당 기관의 업무 효율성이나 서비스 수준을 향상시키는 데 꼭 필요한 인공지능 서비스를 기획하고, 적합한 플랫폼을 신중히 선별해야 한다. 정확도 향상을 위한 학습 계획을 세우고 샘플 기록물 또한 선정해야 한다. 현재로서는 상용 인공지능 API를 적절히 활용하며 조직 내 역량을 키우는 것이 바람직할 것이다. 국가기록원 수준의 기관에서는 인공지능을 다루는 전문인력을 충원하고 지속적인 테스트와 연구를 수행할 필요가 있다. 인공지능 분야는 솔루션을 구매하면 되는 것이 아니라 지속적인 연구와 노력이 요구되기 때문이다.

〈그림 5〉 Google Vision 사진 인식 테스트 결과

The screenshot shows the Google Vision API interface. On the left, a photo of two men shaking hands is displayed with bounding boxes around their faces. Below the photo is the filename '박찬호김영삼.jpeg'. On the right, the 'Web Entities' list is shown with the following items and scores:

Entity	Score
Kim Young-sam	13.358
Chan Ho Park	14.242
Blue House	5.6528
Los Angeles Dodgers	0.5782
KBO League	0.5198
President	0.3735
Seoul	0.33616
Baseball	0.3079
Athlete	0.2941
Pitcher	0.2582
Public Relations	0.22215
November 12	0.2198
	0.2119

〈그림 6〉 IBM Watson 사진 인식 테스트 결과

The screenshot shows the IBM Watson API interface. It features a photo of two men shaking hands with bounding boxes around their faces. Below the photo, there are two sections: 'Classes' and 'Faces', each with a 'JSON' link and a score indicator.

Classes

Class	Score
diplomat	0.64
person	0.80
military attache	0.53
official	0.51
resident commissioner	0.50
reddish orange color	0.70

Type Hierarchy

- /person/diplomat
- /person/military attache
- /person/official
- /person/resident commissioner

Did We Wow You? Yes No

Faces

Face	Score
age 18 - 24	0.39
female	0.00
age 18 - 24	0.39
female	0.00

Did We Wow You? Yes No

4) 사물인터넷 기술의 기록관리 적용 방안

기록관리 분야에서는 스마트 센서를 통한 보존서고 출입 통제나 스마트 항온항습 체계는 근시일 내에 도입할 수 있을 것이다. 스마트 센서를 서고에 부착하게 되면 실시간으로 온습도 데이터가 관리자에게 전송되고 위험 징후가 탐지될 경우 어플리케이션이 경고 메시지를 전달하게 된다. 향후에는 나노센서를 중요 기록물에 부착하여 곰팡이나 산화도를 모니터링할 수 있을 것이다. 기록 매체의 특성상 보존에 취약하거나 처리를 요하는 주요 기록물에 대해 우선순위를 부여하고 체계적으로 관리하는 것이 가능해질 것이다. 보존처리 부서의 산적한 업무를 센싱을 통해 우선순위를 부여하여 처리할 수 있을 것이다. 나노센서 등은 상용화되면 대량생산을 통해 일시에 가격이 낮아지는 특성이 있으므로 지금 현재 비용 대비 효과를 측정할 필요는 없다. 센서는 초경량 초저전력 통신기능을 내장하고 있으므로 사물인터넷 네트워크 인프라 내에서만 통신이 가능하다.

또한 먼 미래에는 아마존 물류창고처럼 서거나 기록물 박스를 옮기는 서고물류로봇이 도입될 수도 있다. 이 경우 종전의 기록물 시설 및 환경기준을 재검토해야 한다. 서가와 서가 사이의 거리는 더 이상 작업자의 편의를 도모할 필요가 없어지므로 좁아져도 되고 효율적인 공간 활용이 가능해질 것이다. 다만 국가기록원과 같은 영구기록물관리기관에서는 1년에 한두 번 수행하는 정수점검에 쓰겠다는 것이므로 효과성이 떨어진다. 사물인터넷 기술은 상시적으로 반복되는 업무에 적용해야 효과가 높다. 단, 센서의 가격이 저렴해지면, 폐기 등 기록의 이동이 많은 기록관 단계에서는 효과성이 높아질 것이다. 수년 전부터 기록물관리기관들은 RFID를 도입해 왔다. 현재까지의 비용효율성은 아주 크지 않았다고 볼 수 있다. 다행히도 기록관들이 그 동안 RFID 시스템을 구축하며 축적한 시행착오와 RFID 인프라는 사물인터넷의 기반으로 활용될 수 있을 것이다. 사물인터넷과 관련하여 강조하고 싶은 것은 국가 차원의 기술 연구와 인프라 지원이 필요하다는

것이다. 신기술이 특정 국가에서 성숙되기 위해서는 긴 시간의 시행착오가 요구된다. 하지만 최근 중국은 넓은 지리적 공간에 동시다발적으로 신기술 개발을 인큐베이팅하여 다양한 성공과 실패 사례를 단기간에 얻어 내며 비약적인 성과를 거두고 있다. 또한 국가가 기술을 구매해 줌으로써 로컬 기업들이 해외 진출까지 성공하고 있다. 중국의 사례에서 보듯이 사물인터넷과 같은 기반 산업은 국가 차원의 전폭적인 기술 투자와 교육을 통해 기술 기업들이 생겨나고 다양한 분야에 적용될 수 있도록 해야 한다.

기록정보 서비스 관점에서도 사물인터넷은 혁신적인 서비스 모델을 제공하는 기반이 될 것이다. 만약 LOD 형태로 기록물 정보를 구축했다면 소장 기록과 모든 디바이스를 연결할 수 있게 된다. 또한 여러 외부 서비스와도 연결할 수 있게 된다. 이렇게 되면 일상적으로 사용하는 스마트폰이나 TV를 통해 기록 정보를 의미 있게 연결시킬 수 있게 된다. 예를 들어 TV 뉴스에서 5·18 민주화 운동을 보다가 국가기록원이 소장한 당시 사망자 명단을 조회해 볼 수 있게 된다. 이러한 서비스는 클라우드와 LOD, 사물인터넷 기술이 복합적으로 연계되었을 때 가능해질 것이다. 이러한 융합을 통해 기록을 통한 새로운 가치를 만들어낼 방안을 점차 모색해야 할 것이다.

5. 맺음말

오스트리아에서 열린 ‘비엔나 비엔날레 2017’의 주제는 “로봇, 노동, 우리의 미래(Robots, Work, Our Future)”였다. 비엔나 전역에서 5개월 동안 전 세계 300명 이상의 디자이너, 건축가, 예술가들이 인공지능과 로봇, 사물인터넷 시대의 유토피아와 디스토피아에 대한 작품을 전시했다. 축제 마지막 날 들렀던 응용미술관(MAK, Museum of Applied Arts)은 방문객이 많아 줄을 서야 했고 전시장의 열기는 대단했다. 비엔날레 소개문에는 “최근의 디지털 혁명이 산업 혁명에 비교할 만하며 서양의 근대주의와 비교했을 때

지금의 디지털 근대주의는 모든 부문에서 삶의 질을 지속적으로 개선할 수 있는 잠재력을 가지고 있다”고 선언하고 있다.³⁴⁾ 전시장 곳곳의 암울하고 과장된 작품 속에서도 관람자들은 무척 진지하며 들떠 있었다.

〈그림 7〉 비엔나 비엔날레 2017 – 응용미술관 전시장 사진



제4차 산업혁명의 파괴적 기술들은 혁신적인 효율성을 발휘해 인간의 노동력을 대체할지도 모른다. 또한 인공지능 에이전트라는 지능을 가진 새로운 객체가 출현하여 종전의 규칙과 윤리에 새로운 화두를 던지고 있다. 로봇과 인공지능에 관한 논쟁은 열정과 비판, 더 나은 첨단 기술의 유토피아와 인간이 소외되는 디스토피아 사이에서 격렬하게 부딪치고 있다. 그러는

34) Museum of Applied Arts, “Vienna Biennale 2017” – http://www.mak.at/en/vienna_biennale_2017_1

사이 IT와 디자인은 이와 상관 없이 구체적인 솔루션을 내놓고 있다. 국내에도 제4차 산업혁명과 신기술에 대한 들뜬 분위기는 대단할 정도이다. 다만 이에 대한 진지한 논의의 틀은 아직 형성되지 않은 듯 하다. 냉소적으로 말하자면 신기술 맹신자와 무관심한 자들의 간극으로 인해 제대로 된 논의가 이루어지기 어려운 것 같다. 현재의 시점은 신기술에 대한 이해를 넓히고 각자의 분야에 긍정적으로 적용하기 위한 방법을 상상해 보는 초입이다.

문헌과 실험을 통해 클라우드, 빅데이터, 인공지능, 사물인터넷 중 이미지 인식 등 개별 기술의 대단한 성능을 확인하였고 이 기술들이 융합되었을 때 우리의 일상과 업무환경을 크게 향상시킬 것이라는 확신이 들었다. 이러한 기술들은 방대한 기록을 정확하게 처리함으로써 기록관리 업무에 큰 효율을 가져다 줄 것이다.

이제 더 연구해 볼 분야는 각 기술들이 자신의 업무 분야에 어떤 가치를 만들어낼 수 있을지를 나열하고 구체적인 로드맵을 수립하는 것이다. 이를 위해서는 치열한 학습을 통해 각 기술의 특징과 파급력을 이해하는 것이 반드시 선행되어야 한다. 분명한 건 신기술들이 업무의 양상을 주도할 것이라는 점이다. 클라우드 컴퓨팅이 주는 유연함과 효율성은 기록의 양을 더욱 증가시킬 것이고 센서 데이터 등 새로운 유형의 기록을 만들어낼 것이다. 더욱 방대해진 기록의 망망대해에서 기록관리자와 아키비스트는 기술 주도로 생산된 기록들의 투명성과 설명책임성을 밝히기 위해 여전히 매진해야 한다. 기술이 발전할수록 기록 담당자의 노력은 더욱 필요하다.

이 연구가 제4차 산업혁명의 신기술들을 이해하고 차세대 기록관리의 틀을 설계하는 데 도움이 되길 바란다.

〈참고문헌〉

강정호, 김형주, 전문석. 2015. 사물인터넷 시장 및 기술 동향. 『한국콘텐츠학회지』, 13(1), 14-17.

- 국가기록원. 2017. '차세대 기록관리 모델 재설계 제2차 워크숍' 발표자료
- 김성태. 2013. 새로운 미래를 여는 빅데이터 시대. 한국정보화진흥원. 빅데이터 전략연구센터, 2, 9-579.
- 김재생. 2015. 사물인터넷의 기술 소개 및 정책 방안. 『한국콘텐츠학회지』, 13(1), 18-24.
- 산업통상자원부. 2017. '5차 산업업 민관협의회 컨퍼런스 : 산업부 장관 기조강연'. Retrieved 2017.10.5. http://www.motie.go.kr/common/download.do?fid=bbs&bbs_cd_n=81&bbs_seq_n=159256&file_seq_n=3
- 안대진. 2015. 「기록관리 민간위탁 해외사례」, RIKAR 월간 브리프 기록 haja 9월호, 12.
- 엄도영, 김동휴, 이희진. 2016. 파괴적 기술 분야에 대한 표준화 전략 연구. 『기술혁신학회지』, 19(3), 483-510.
- 한국과학기술기획평가원(KISTEP). 2017. '2016 기술영향평가 보고서: 가상, 증강현실 기술'. 연구보고 2017-34.
- 한국은행. 2016, '제4차 산업혁명: 주요국의 대응현황을 중심으로'. 국제경제리뷰 2016-24호.
- 한국전자통신연구원(ETRI). 2009. 클라우드 컴퓨팅 기술 동향. 『전자통신동향분석』, 24(4).
- 한국전자통신연구원(ETRI). 2015. 인공지능 기술과 산업의 가능성. Issue Report, 2015-04.
- KDB산업은행. 2017. '4차 산업혁명의 주요 신기술 적용 현황 및 시사점'.
- Gartner. 2015. *Selecting Impactful Big Data Use Cases*.
- Gartner. 2017. "Gartner Says 8.4 Billion Connected "Things" Will Be in Use in 2017, Up 31 Percent From 2016". Retrieved 2017.10.10. <http://www.gartner.com/newsroom/id/3598917>.
- Laney, D. 2001. 3D data management: Controlling data volume, velocity and variety. META Group Research Note, 6, 70. Retrieved 2017.10.5.
- Mckinsey & Company. 2015. Unlocking the potential of the Internet of Things". Retrieved 2017.10.10. <https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/The%20Internet%20of%20Things%20The%20value%20of%20digitizing%20the%20physical%20world/The-Internet-of-things-Mapping-the-value-beyond-the-hype.ashx>
- Mell, Peter. & Grance, Timothy. 2011. "The NIST Definition of Cloud Computing". National Institute of Standards and Technology: U,S. Department of Commerce. Retrieved 2017. 10. 5. <http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf>.
- MIT Technology Review. 2015. "Why Self-Driving Cars Must Be Programmed to Kill". Retrieved 2017. 10. 10. <https://www.technologyreview.com/s/542626/why-self-driving-cars-must-be-programmed-to-kill/>.

- Porter, M. E. 1991. Towards a dynamic theory of strategy. *Strategic management journal*, 12(S2), 95-117.
- Porter, M. E. 2008. *Competitive advantage: Creating and sustaining superior performance*. Simon and Schuster.
- Schwab, K. 2017. *The fourth industrial revolution*. Crown Business.
- Searle, J. R. 1980. Minds, Brains, and Programs. *Behavioral and Brain Sciences*, 3(3), 417-457.
- World Economic Forum, 2016. 『World Economic Forum Annual Meeting 2016 : Mastering the Fourth Industrial Revolution.』, Retrieved 2017.10.2. http://www3.weforum.org/docs/WEF_AM16_Report.pdf.