

클라우드 저장소 장점을 활용한 기록 콘텐츠¹⁾ 관리기능 설계

Designing the Record Management Functions for Record Content Using Advantages of Cloud Storage

임진희(Jin-Hee Yim)

E-mail: yimjh68@seoul.go.kr

서울특별시 정보공개정책과장



논문접수 2019.8.1
최초심사 2019.8.6
게재확정 2019.8.27

초 록

최근 중앙행정기관은 업무관리시스템을 클라우드 기반의 온나라2.0으로 변경하였다. 국가기록원은 클라우드 업무관리시스템의 기록을 이관받아 관리할 수 있도록 클라우드 기반의 기록관리시스템을 개발하여 보급하고 있다. 클라우드 컴퓨팅의 이점을 극대화하여 기록관리가 보다 효과적·효율적으로 이루어지도록 재설계할 수 있는 기회이다. 전자기록 관리의 프로세스와 방법이 종이기록 관리 방식을 단순히 전자화하는 것에서 탈피하여 디지털 기술에 따른 변환(Transformation)의 수준으로 나아갈 수 있는 기회이기도 하다. 첫째, 이관의 방식을 변환해 볼 수 있다. 업무관리시스템과 기록관리시스템이 클라우드 저장소를 공유하게 되면 처리과에서 기록관으로 기록물 이관 시 콘텐츠 파일들을 물리적으로 이동시키지 않고 메타데이터만 복사하는 방식으로 이관할 수 있어 비용이 줄고 무결성 훼손 위험이 줄어들 수 있다. 둘째, 기록물의 저장공간 할당에 대한 전략을 구상해 볼 수 있다. 클라우드 저장소를 업무관리시스템과 기록관리시스템이 공유하는 것을 전제로 한다면 콘텐츠 파일들을 저장할 때 기록의 보존기간에 따라 저장하는 위치를 구분함으로써 이점을 얻을 수 있다. 셋째, 기록의 생산시스템, 기록관리시스템, 정보공개시스템 등 콘텐츠에 접근하는 시스템들이 클라우드 저장소를 공유하게 되면 콘텐츠의 중복을 최소화하는 방향으로 설계를 전환할 수 있다.

© 한국기록관리학회

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

1) 'contents'와 'content'가 영어에서 갖는 의미 차이를 영두에 두면 이 논문의 제목에는 '콘텐츠'가 맞는 단어이다. '콘텐츠'는 기록의 메타데이터와 대비되는 기록의 내용을 자체를 지칭하는 것으로, 디지털 기록의 경우에는 컴퓨터 파일에 담기게 된다. 웹콘텐츠나 '목록'을 의미하는 '콘텐츠'와의 혼동을 피하기 위해 기록의 내용물에 대해 '콘텐츠'라는 단어를 사용하고 싶었으나, 논문 심사의견을 존중하여 표준어인 '콘텐츠'로 표현하기로 한다. 이 논문에서 '콘텐츠'라 지칭하는 대상은 MoReq2010에서 'content'로 지칭하는 것과 동일하다. 공공기록관리 용어를 이용하여 정리하자면, 전자기록 건(item) 하나는 하나 이상의 컴포넌트(component)들로 구성되며, 하나의 컴포넌트는 하나의 콘텐츠(content)를 참조한다. 여기서, 콘텐츠는 기록의 내용을 담고 있는 객체이며, hwp, odt, xls 등과 같은 확장자로 저장 매체에 존재하는 컴퓨터 파일이다.

ABSTRACT

Recently, the central administrative agency changed its business management system to cloud-based On-nara 2.0. To transfer and manage the records of the cloud business management system, the National Archives Service has developed and distributed a cloud-based records management system. It serves as an opportunity to maximize the benefits of cloud computing and redesign the records management to be more effective and efficient. The process and method of electronic record management can be transformed through digital technologies. First, we can change the transfer method for electronic records. When the business and the records management systems share the same cloud storage, it is possible to transfer the content files between the two systems without moving the contents files physically, thus copying only the metadata and reducing the cost and the risk of integrity damage. Second, the strategy for allocating storage space for contents can be conceived. Assuming that the cloud storage is shared by the business and the record management systems, it is advantageous to distinguish the storage location based on the retention period of the content files. Third, systems that access content files, such as records creation, records management, and information disclosure systems, can share the cloud storage and minimize the duplication of content files.

Keywords: 기록관리시스템, 클라우드 저장소, 논리적 이관, 저장 전략, 콘텐츠 중복
Records management system, Cloud storage, Logical transfer, Strategy for storage, Content duplication

<http://ras.jams.or.kr>

1. 머리말

행정안전부는 클라우드 서비스 기반의 업무 관리시스템인 온나라2.0과 클라우드 행정 포털 서비스를 개발하여 2018년 말까지 모든 중앙행정기관에 보급을 마친 상태이다(국가기록원, 2019). 여기서 ‘업무관리시스템’은 일반명사이고 ‘온나라’는 브랜드에 해당하는 고유명사이다. 최초의 공공기관용 업무관리시스템은 참여정부 시절 대통령비서실에서 만든 ‘이지원’ 시스템이다. 이후 이지원을 개작하여 ‘하모니’ 시스템과 ‘온나라’ 시스템이 개발되었으며 2006년 말부터 중앙행정기관에 보급되기 시작했다. 이 때의 온나라시스템을 버전 1.0으로 부르고, 최근 클라우드 기반으로 고도화한 온나라시스템을 버전 2.0으로 구분하여 부른다.

온나라2.0은 문서2.0, 지식2.0, 공유2.0, 포털 등의 세부 서비스로 구성되어 있다. 이 중 온나라 문서2.0이 기존의 온나라 업무관리시스템을 클라우드 기반으로 고도화한 버전이다. 업무관리시스템은 행정효율과 협업추진에 관한 규정에 근거하여 기능이 구성된다. BRM(Business Reference Model)과 연계하여 과제관리를 하고 과제카드 하위에 기안문을 작성하여 결제 받는 기능이 기본이다. 기안문과 시행문의 서식을 표준화하거나 유효한 관인을 부착하는 방법 등이 업무관리시스템을 통해 시스템적으로 지원된다. 협업과제를 진행하고, 정책실명제에 따른 과제 및 문서관리를 수행하며, 조직 변경 시 업무의 인계인수를 전자적으로 수행하기 위한 기능도 업무관리시스템을 통해 수행하도록 한다(국가기록원, 2019).

온나라2.0은 클라우드 컴퓨팅으로 누릴 수

있는 여러 장점을 가지고 있는데 몇 가지를 제시해보면 다음과 같다.

첫째, 온나라2.0은 정부의 클라우드 서비스로 개발되어 기존의 업무관리시스템과는 질적으로 다른 배포 방식과 실행구조를 갖는다. 온나라1.0의 경우 각 기관이 독자적인 서버를 구축하고 행정안전부가 배포해주는 온나라1.0 소프트웨어를 각자의 서버에 설치하여 사용했다. 반면, 온나라2.0의 경우 하나의 정부 클라우드 내에 기관들이 모두 모여서 응용 서비스를 실행하는 구조인데, 행정안전부가 개발한 하나의 단일한 소프트웨어를 다중 서비스로 실행하므로 별도의 배포 절차가 생략되는 구조이다. 클라우드 서비스를 이용하게 되면 동일한 소프트웨어를 사용하는 기관들이 많을수록 하드웨어 자원을 절약하면서 소프트웨어의 배포 비용도 획기적으로 줄일 수 있다. 업무관리시스템 이외에도 기록관리, 정보공개, 재정관리, 인사관리 등 다중 기관이 필수적으로 이용해야 하고 표준화가 요구되는 응용을 함께 클라우드 서비스화하면 갖는 이점이 상당히 크다고 평가된다.

둘째, 온나라2.0은 G-클라우드 저장소를 사용한다. G-클라우드 저장소는 중앙행정 기관들이 함께 공유하는 공간으로 온나라2.0에서 생성되는 문서뿐만 아니라 개인 업무용 PC의 문서도 저장할 수 있도록 운영된다. 저장소 공유로 인해 전체 중앙행정기관 간에 문서 공유가 용이해졌다. 기관들 간 업무 협조를 위해 수많은 문서들이 전자문서유통시스템을 통해 물리적으로 송수신되던 것이 이제는 저장소 내에서 콘텐츠 파일의 URL을 공유하는 방식으로 단순화되었다. 또한, 출장 중에 타 지역의 스마트 워크센터에서 저장소에 올려둔 문서를 열어

처리할 수 있는 등 이동성도 크게 강화되었다. 클라우드 저장소의 도입은 업무관리시스템의 저장 공간과 기록관리시스템의 저장 공간이 별도 서버로 분리되어 있던 온나라1.0 - 표준기록관리시스템 체계와는 완전히 상이한 전략을 구상할 수 있다는 점에서 특히 주목된다.

셋째, 온나라2.0을 개발하면서 결재문서 본문의 포맷이 변경되었다는 점이다. 기존의 업무관리시스템이나 전자문서시스템은 대부분 기안기로 한글 문서편집기를 사용하였으며, 따라서 결재문서의 본문은 hwp 포맷으로 저장하도록 되어 있었다. 그런데, 온나라 문서2.0에서는 개방포맷 ODF(Open Document Format) 종류인 odt로 저장하도록 변경한 것이다. 포맷의 변경은 클라우드와 직접적인 관련이 없어 보이지만 기본적으로 개방적인 플랫폼을 추구하는 클라우드에 걸맞는 발전적 선택이기에 연관이 없다 할 수 없다. 클라우드 플랫폼에 사용되는 소프트웨어들이 대부분 개방소프트웨어(OSS, Open Source Software)로 구성되어 있으므로 일관된 개방형 전략에 따라 저장 포맷을 선택했을 것이다. 과거 전자민원 서식들이 대부분 hwp 문서로 배포되어 시민들이 상용 한글 워드프로세서를 구매할 수밖에 없는 상황이 오랫동안 문제가 되었다. 민원서식과 결재문서의 본문을 개방형 포맷으로 전환하는 것은 뒤늦은 감은 있으나 시대적 요구에 맞춘 전향적 조치라 평가할 수 있다. 현재는 결재문서 본문에만 국한되지만 향후 모든 콘텐츠를 최대한 개방형 포맷으로 전환해가는 정책을 추진할 필요가 있겠다.

업무관리시스템은 기록관리의 관점에서 보자면 기록의 생산시스템이다. 기록관리를 효과

적으로 추진하기 위해서는 생산 과정에서부터 기록관리의 원칙이 적용되도록 적절한 통제를 가하는 것이 필수적이다. 하지만, 그간 우리나라 공공영역에서 기록관리 기관이 기록의 생산 시스템에 통제력을 행사할 만한 권한과 책임을 갖기는 어려웠다. 그 결과로 기록관리 기관은 주로 기록이 생산되는 환경과 시스템의 변화과정을 잘 관찰하고, 변화된 환경에서 생산된 기록의 유형과 품질을 평가하는 것으로 부터 업무를 시작해왔다. 그렇다면, 클라우드 업무관리 시스템이 만들어진 상황에서 기록관리 기관이 해야 할 일은 무엇일까? 클라우드 저장소에 생성되는 기록정보를 효과적, 효율적으로 관리하기 위한 방안은 무엇일까? 클라우드 업무관리 시스템에 조응한 기록관리시스템은 어떤 특징과 기능을 갖추어야 할까? 이 논문은 이러한 질문 중 핵심 쟁점 세 가지에 관해 논의하고 대안을 구상해 보고자 한다. 쟁점 중에는 과거에 이미 논의되었던 것도 있고, 클라우드 도입으로 인해 새롭게 제기되는 것도 있다. 과거에 논의되었던 쟁점의 경우에는 기록관리 분야의 묵은 숙제라 볼 수도 있는데 클라우드 환경으로의 전환을 계기로 다시 한 번 개선안을 만들어 볼 수도 있겠다.

국가기록원은 클라우드 업무관리시스템인 온나라2.0이 개발되는 과정에서 이에 조응하는 클라우드 기록관리시스템인 cRMS(Cloud Record Management System) 개발 완료하였고, 온나라2.0을 도입한 중앙행정기관에 cRMS도 함께 도입하도록 하고 있다(국가기록원, 2019). 하나의 클라우드 저장소와 플랫폼에 업무관리시스템과 기록관리시스템이 함께 서비스로 구동되는 체계인 것이다. 이렇게 기록생산시스템과

기록관리시스템이 클라우드로 묶이게 되면 그동안 처리과-기록관 단계로 나뉘어 있던 기록 관리의 절차와 처리 방식을 완전히 새롭게 구상할 수가 있다. 특히, 지난 10 여 년간 전자기록 관리 실무자들이 비효율적이라 판단했던 물리적 이관 방식이나 장기보존포맷 변환 등에 관해 재검토하고 다시 설계해 볼 좋은 기회가 주어진 것이다. 국가기록원은 cRMS를 개발하면서 개선의 기회를 잘 활용했을까?

긍정적으로 평가할 점은 국가기록원이 요청하여 업무관리시스템에서 결재문서 본문을 odt로 저장하면서 동시에 문서보존포맷인 PDF/A-1로도 저장하도록 했다는 점이다. 기록관리 법령에 따라 콘텐츠 파일을 문서보존포맷으로 변환해 본 경험이 있는 기록관리 기관은 대부분 여러 종류의 문제를 발견했을 것이다. 콘텐츠 파일이 암호화되어 있어서 열어볼 수 없는데 생산자가 암호를 기억하지 못해 복호화에 실패하는 경우가 있다. 압축된 파일의 경우에는 일일이 압축을 풀어 변환을 진행해야 한다. 또한, 결재문서에 첨부할 콘텐츠 파일이 용량이 커서 별송으로 처리한 경우에는 콘텐츠 파일 확보에 실패하는 경우가 대부분이다. 문서보존포맷 변환 후에 제대로 변환이 이루어졌는지 확인하는 과정은 샘플링하여 육안으로 검수를 해야 해서 막대한 비용이 수반되고 있다. 이런 상황을 놓고 본다면 결재문서가 만들어짐과 동시에 관련 콘텐츠 파일들을 모두 문서보존포맷으로 변환하여 작성자가 변환 결과를 확인하도록 하는 것이 장기보존 기록의 품질을 높이는데 결정적이라 할 수 있다. 즉, 문서보존포맷 변환의 시점을 문서의 생산시점으로 당겨옴으로써 많은 문제를 해결할 수 있다는 것이다. 문

서보존포맷 변환 과정을 통해 문서 작성자들에게 본인의 문서가 영구적으로 보존될 수도 있다는 점을 인식시켜줄 수 있다. 반면 아쉬운 점은 업무관리시스템에서 결재문서를 구성하는 본문과 첨부 콘텐츠 파일들이 중복적으로 저장되는 구조가 개선되지 못하고 지속되고 있는 것인데, 파일 중복의 문제는 4장에서 심층적으로 다루도록 하겠다.

클라우드 컴퓨팅 기술이 발전하면서 ‘클라우드 컴퓨팅 발전 및 이용자 보호에 관한 법률’이 만들어지고, 공공영역에서도 클라우드 서비스를 제공하면서 ‘정부 클라우드저장소(G드라이브)이용지침’을 발행하였다. 클라우드 서비스가 공공 기록관리에 어떤 영향을 줄지에 관해서는 서울시가 수행한 서울기록원 정보화전략 계획(ISP) 수립 용역에서 다룬 바가 있다. 영구기록물관리기관인 서울기록원에서는 대량의 전자기록을 안정적으로 저장할 공간을 설계할 필요가 있었다. 매년 이관되어 오는 기록의 증가량을 고려하고, 이관 시점에 품질 검사를 수행하기 위한 작업 공간이 필요하며, 장기보존을 위해 포맷 변환을 할 때 대량의 임시 작업 공간이 필요하다는 점 등을 고려했을 때 클라우드 저장소를 활용하면 여러 가지 장점이 있을 것으로 예상한 바 있다. 우선 값비싼 스토리지 중심의 대용량 저장소를 꾸리는데 드는 비용을 절감할 수 있고, 일시적으로 필요한 저장 공간도 클라우드에서 자동 할당받아 사용 후 반환함으로써 융통성있게 시스템을 운영할 수 있을 것으로 예측한 바 있다(서울특별시, 2016). 국가기록원이 수행한 ‘차세대 기록관리 모델 재설계 연구’에서도 클라우드 관련 내용을 다룬 바가 있다. 이 연구에서는 소위 제4차 산업

혁명 기술이 전자기록 관리에 미치는 영향을 살펴보면, 한편으로는 클라우드를 포함하여 인공지능, 블록체인, 모바일 등의 기술에 의해 새롭게 생산되는 기록의 유형에 유의할 것과 다른 한편으로는 이 기술들을 활용하여 기록관리를 개선할 것에 대해 논의하고 있다(국가기록원, 2017). 연구 결과로 작성된 안대진, 임진희(2017)의 논문에서는 클라우드, 빅데이터, 인공지능, 사물인터넷 등의 기술을 기록관리 업무에 어떻게 적용해야 할지 제안하고 있다. 이 연구가 이루어지던 2017년, 행정안전부에서는 이미 클라우드 기반의 업무관리시스템인 온나라 문서2.0의 개발을 완료하고 보급을 시작하였으며, 기록관리시스템의 클라우드 버전을 국가기록원과 논의하고 있었다. 국가기록원이 행정안전부와 협력하여 cRMS를 만들 때 서울시의 연구결과나 국가기록원의 연구결과를 살펴보고 최대한 이를 반영하고자 했다면 지금 이 논문에서 제기하는 여러 문제들은 이미 해소되고 없었을 지도 모른다. 하지만 아쉽게도 서울시나 국가기록원의 선행연구에서 제안된 개선 모델들은 cRMS를 만드는 과정에서 수용되지 못했고 어쩌면 아예 검토조차 이뤄지지 못했으며, 결과적으로 전자기록 관리 체계를 개선할 기회를 놓치게 되었다.

이 논문에서는 클라우드 업무관리시스템 도입에 조응하여 클라우드 기록관리시스템을 설계할 때 검토해야 할 디지털 변환의 과제를 제시하고자 한다. 이 논문에서 다루는 과제는 앞에서 살펴본 온나라2.0과 cRMS의 도입과 일치하는 환경에서의 고려사항으로 한정한다. 즉, 기관 내의 전자기록 관리로 논의를 한정하고자 하며, 기관에 클라우드 저장소와 플랫폼이 구축

되고 기록생산시스템과 기록관리시스템이 함께 클라우드의 서비스로 전환되면서 클라우드 저장소를 공유하는 경우를 가정한다. 클라우드 환경의 장점을 최대화하여 그간 전자기록 관리 과정에서 드러난 어려움을 완화하거나 최소화하는 과제로 전자기록의 논리적 이관 실현, 생애주기에 따른 기록정보의 저장 전략, 콘텐츠 중복 최소화로 효과적인 관리 등의 개선 방안을 제시하고자 한다. 충분히 개선되지 못한 채로 이미 시스템 보급이 완료된 중앙행정기관에 이러한 개선사항을 적용하기 위해서는 다시 오랜 시간을 기다려야 하겠지만, 서울시나 기타 공공기관 중에 독자적인 클라우드 서비스 체계를 추진하는 기관에서는 여기서 제시하는 과제를 검토하여 좀 더 개선된 모델로 나아갈 수 있으리라 기대한다.

2. 이관방식의 변환

2.1 이관 단순화의 필요성

클라우드 서비스의 기본 중 하나는 인프라로서 저장소를 제공한다는 것이다. 이 때, 클라우드 저장소는 여러 어플리케이션들이 쉽게 공유될 수 있는 체계로 설계된다. 업무관리시스템과 기록관리시스템이 하나의 클라우드 플랫폼에서 서비스된다면 기록이 생산되고 저장된 저장소를 공유하게 되므로 전자기록의 이관 방식에 대해 새로운 접근을 할 수 있다. 소위 '물리적 이관' 대신 '논리적 이관'을 하여 기록정보의 이관과정을 단순화할 수 있게 된다. 그렇다면, 기록의 물리적 이관과 논리적 이관은 어떤 차

이를 갖고 있을까? 물리적 이관이 기록물 객체를 다른 곳으로 이동시키면서 관리권한을 이양하는 방식이라면, 논리적 이관은 기록물 객체는 이동하지 않고 관리권한만을 이양하는 방식이라고 설명할 수 있겠다. 물리적 이관이든 논리적 이관이든 기록의 메타데이터를 복사해 오는 것은 공통적으로 필요한 과정이다.

이생동(2014)은 기록학 분야에서 처음으로 클라우드 기반의 기록관리시스템 프레임워크를 제시하면서 물리적 이관에서 기록물에 대한 관리권 변경으로 이관방식을 변경할 것을 제안한 바 있다. 이후 정부의 클라우드 플랫폼이 본격화되고, 김주영, 김순희(2019)도 클라우드 저장소를 활용하여 기록물을 논리적으로 이관하는 방안을 제안하고 있다. 김주영 등은 논리적 이관이 물리적 이관에 비해 갖는 첫 번째 장점으로 클라우드 컴퓨팅 기술이 이미 확대되고 있으므로 이 환경을 효과적으로 활용하는 것이 필요하다는 점을 강조했다. 클라우드 서비스를 촉진하기 위한 법령이 제정되고 공공영역에 클라우드 서비스를 확장하기 위한 여러 모색이 이루어지고 있는 시점에 해당 기술로 인해 혜택을 볼 수 있는 형태로 업무 방식이 변화될 필요가 있다는 점에서 적절한 지적이라 볼 수 있다. 두 번째와 세 번째 장점으로 물리적 이관 과정에서 전자기록물이 무결성 훼손 등의 위험에 더 노출될 수 있는데 논리적 이관으로 변경함으로써 위험도를 낮출 수 있다는 점을 들고 있다. 디지털 객체는 현재 저장되어 있는 매체 자체가 노후화되거나 보안 상의 취약점이 노출되기 전까지는 그곳에서 보관하는 것이 다른 매체로 이전하여 보관하는 것보다 상대적으로 안전하다. 비트스트림은 동일 매체에 보관 중

일 때 보다 매체 공간을 옮겨가는 동안에 손실이나 왜곡이 일어날 가능성이 높기 때문이다. 네 번째로는 물리적 이관 작업은 대량의 콘텐츠들을 이동시키는 대규모 작업이라 많은 시간과 노력을 요구한다는 점을 지적하고 있다.

전자기록의 이관 방식을 논의하기에 앞서 기록의 이관을 왜 하는가에 대해 정의해보면 클라우드 환경에서 효과적이고 효율적인 이관 방식에 대해 좀 더 합리적인 결론내릴 수 있을 것이다. 이 논문의 논의 범위는 공공기관의 처리과에서 기록관으로 전자기록을 이관하는 것으로 한정한다. 기록의 이관은 “기록이 필요한 기간만큼 안전하고 무결하게 보존되고 이용 가능하도록 한다”라는 목표를 달성하기 위해 수행하는 관리 행위의 하나라 할 수 있다. 처리과에 두어서는 기록이 안전하기 어렵다거나 무결성을 유지하기 어렵다거나 보존여건이 미비하다거나 이용하기 어려워진다고 판단될 때 기록관으로 옮겨 보다 전문적인 관리와 통제를 수행하도록 하는 것이 이관의 목적인 것이다.

종이기록의 경우에는 처리과별 캐비닛에 기록물이 흩어져 있어 전사적인 이용이 어렵고, 현안 업무가 종료되고 나면 해당 기록물을 소홀히 다루게 되어 분실할 수도 있으니 현용 단계가 지난 기록물을 기록관의 통합 서고로 이동하는 것이 필요하다. 즉, 종이기록은 실물이 공간을 이동하는 방식의 물리적 이관이 일반적이다. 하지만, 전자기록의 경우에는 어떨까? 전자기록은 주요한 생산시스템에 집합적으로 모여 있는 것이 일반적이고, 시스템을 기능에 이용하여 전사적으로 활용이 용이하며, 현안 업무가 종료되었다고 하여 시스템에서 함부로 삭제하는 일은 흔치 않은 상황이다. 그럼

에도 불구하고 업무관리시스템 스토리지에 있는 콘텐츠 파일들을 기록관리시스템의 스토리지로 복사해가는 방식의 물리적 이관을 해야만 할까?

여기서 주목해야 할 것은 전자기록의 경우 관리상의 기본 목표가 진본성과 이용가능성을 유지하는 것인데, 물리적 이관을 거치게 되면 그 과정에서 진본성과 이용가능성을 훼손할 위험이 증가된다는 점이다. 따라서, 논리적 이관이 가능한 상황이라면 그 방향을 취하는 것이 더 유리하다고 본다. 그런데, 원한다고 언제든 전자기록의 논리적 이관이 가능한 것은 아니다. 논리적 이관이 가능하려면 공공기관이 업무시스템과 기록관리시스템이 함께 접근할 수 있는 기록물 저장소를 운영하고 있어야 하고, 이관된 기록물과 이관되지 않은 기록물을 구분하여 관리권한을 통제할 수 있어야만 한다.

2.2 전자기록 논리적 이관 논의

현재 전자기록을 물리적으로 이관하게 된 배경이나 역사를 살펴보면 다음과 같다. 2004년 참여정부시절 대통령비서실에서 ‘기록관리체계 재설계 방안 연구’ 용역이 있었다.²⁾ 그 연구에서 다른 과제 중 하나가 EDMS(Electronic Documents Management System)와 ERMS(Electronic Records Management System)를 분리할 것인가, 혹은 EDRMS(Electronic Documents & Records Management System)로 통합할 것인가 하는 문제였다. 여기서 말하는 EDMS는 당시 대통령비서실에서 개발 중인 ‘이지원’ 시스

템을 지칭하는 것이었고, ERMS는 자료관시스템을 고도화하여 구상하던 새로운 기록관리시스템을 지칭하는 것이었다. 당시의 법령에 따르면 전자문서시스템에서 만들어진 전자기록은 종이 기록과 마찬가지로 2년 뒤에 자료관시스템으로 이관하도록 되어 있었고, 메타데이터와 콘텐츠 파일들을 규격에 맞춰 자료관시스템으로 복사해가고 전자문서시스템에서는 삭제하는 방식이었다. 즉, EDMS와 ERMS의 역할이 구분될 뿐만 아니라 별도의 서버로 구축하도록 하고 기록을 물리적으로 이동하는 방식으로 이관하는 것이 기본 설계였다.

연구용역의 과제는 전자정부법의 취지에 맞춰 공공기록의 주류가 더 이상 종이기록이 아니라 전자기록이라는 전제 하에 기록관리 프로세스를 점검하고 절차를 재구성하는 작업이었다. 전자결재 기능이 중심이던 전자문서시스템의 한계를 인식하고, 업무혁신을 하고자 과제관리, 지시사항, 일정관리, 일지, 전자결재, 메모보고, 온라인 회의 등의 관리기능을 대폭 추가한 이지원 시스템을 개발하던 과정에서 진행된 연구용역이었다. 이지원시스템은 단순한 EDMS차원을 넘어서는 BPMS(Business Process Management System)이었고, 결재문서 외에도 다양한 기록이 생산되는 시스템이었으며, 기록 간의 연관관계가 중요한 맥락으로 작동하는 시스템이었다. 대통령비서실의 이지원시스템을 중앙행정기관용으로 수정 보급한 것이 온나라 업무관리시스템이었다. 이지원시스템에 조응하는 새로운 기록관리시스템을 설계하는 연구로서 전자기록의 이관 방식을 포함하여 진본성과 장기보존 요건 등

2) home.mju.ac.kr/yimjh “나의 프로젝트 산출물” 게시판에서 최종보고서를 볼 수 있다. 아쉽게도 EDMS와 ERMS의 통합에 따른 장단점 비교표 등의 중간산출물은 최종보고서에서는 누락되어 확인할 수 없다.

새로운 논의를 진전시켰으며, 연구 결과는 2006년도 공공기록물 관리에 관한 법령 전부 개정 시에 상당부분 반영이 되었다.

2004년 당시, 새로운 전자기록 관리체계에서 이관은 어떤 방식이어야 할가에 관해 논의하면서 콘텐츠 파일들을 물리적으로 이관하는 체계는 IT관점에서 보면 비합리적이며, 이지원시스템에 기록관리 기능을 통합해 개발하거나 이지원시스템과 기록관리시스템 간에 콘텐츠 저장소를 공유하는 방식으로 구축하는 것이 효과적·효율적이라는 쪽으로 의견이 개진되었었다. 하지만, 논의 끝에 이지원시스템에 기록관리 기능을 통합해서 넣는 것은 여건 상 허용되지 않았고, 저장소를 공유하는 것은 업무관리와 기록관리 간에 권한과 책임 소재가 명확치 않다는 등의 이유로 받아들여지지 않았다. 한편 기록관리 내에서도 기록이 생산된 후 가능하면 이른 시점에 기록을 고정시켜 물리적으로 획득해오는 것이 중요 기록의 누락이나 훼손을 막는 최선의 조치라는 주장이 제기되어, 최종적으로 이지원시스템과 기록관리시스템의 서버는 분리되었고 물리적 이관 방식이 채택되게 되었다. 다만 이관시점은 생산 2년 이내가 아니라 1년 이내에 이관하는 것으로 조정하게 되었다(대통령비서실, 2004).

2004년의 연구 내용은 2005년 정부혁신지방분권위원회 산하 기록관리혁신위원회의 기록관

리로드맵에도 반영되었다(정부혁신지방분권위원회, 2005). 대통령비서실에서 만들어진 이지원시스템-기록관리시스템의 모형이 이후 온나라업무관리시스템-표준기록관리시스템 모형으로 진화하게 된다. 이로써 이후 공공기관의 기록관리 담당자들은 콘텐츠 파일들을 물리적으로 이관하는 단 하나의 기록관리 체계만을 경험하게 되었다. 시스템 상에 구현된 이관 기능을 표준적인 것으로 받아들여 현재도 전자기록을 종이기록처럼 물리적 이관을 하는 것이 기본이고 원칙인양 여기고 있는 듯하다. MoReq2010³⁾에서도 제시하고 있고, 국가기록원의 NAK 19-1:2012(v1.0) 전자기록생산시스템 기록관리 기능요건⁴⁾에서도 언급하고 있듯이 기록관리시스템은 기록생산시스템과 하나로 통합되어 있을 수도 있고 기록생산시스템과 별도로 존재할 수도 있다. 콘텐츠 파일의 이관은 각각의 상황에 따라 논리적인 방식과 물리적인 방식을 선택할 수 있어야 정상이다. 물리적 이관이라는 당시의 결정은 특정 조건과 상황을 반영한 것으로 모든 조건에서 최선의 선택은 아님에도 불구하고 법령이나 표준에 이러한 점이 잘 표현되지 못했다. 그 결과 상황의 변화에도 불구하고 그에 조응한 다양한 이관 방식을 상상하는 것조차도 제약받아온 건 아닌지 아쉬움이 든다.

3) <https://moreq.info> 사이트에서 MoReq2010에 관한 문서를 찾아볼 수 있다.

4) 표준의 9-16쪽에 걸쳐 “5 전자기록생산시스템과 기록관리시스템과의 관계 유형”에서 3가지 기록관리 모듈의 유형이 소개되고 있고, 기관마다 복수의 모듈 유형이 필요할 수 있으며 그 경우 여러 유형의 기록 목록을 통합하여 관리하는 통합 기록관리 모듈이 하나 필수적으로 필요하다는 점이 제시되어 있다. 3가지 유형이란, 기록생산시스템에 기록관리 모듈이 포함되어 함께 구축되는 경우, 기록생산시스템과 독립적으로 기록관리시스템을 구축하여 기록물을 물리적으로 이관하는 경우, 기록생산시스템과 독립적으로 기록관리시스템을 구축하되 기록물을 물리적으로 이관하지 않고 기록생산시스템에 둔 채로 관리하는 경우 등이다.

2.3 클라우드 저장소 기반의 이관

콘텐츠 파일들을 물리적으로 이동하는 과정은 쉽게 복사하여 붙이기(copy&paste)로 이해할 수 있다. 그런데, 이동하는 두 시스템의 특성이 다르고 상호운용성이 확보되어 있지 않은 경우에는 복사하여 붙이기가 그리 간단치가 않다. 현재 국가기록원이 배포한 이관규격서를 보면 업무관리시스템에서 표준기록관리시스템으로 이관할 때, 전자문서시스템에서 표준기록관리시스템으로 이관할 때, 표준기록관리시스템에서 중앙기록관리시스템으로 이관할 때 등 각각의 경우 이관할 기록정보를 내보내기(export) 하는 방법과 들여오기(import) 하는 방법을 구체적으로 명시하고 있다. 이 규격에 맞춰 이관에 필요한 소프트웨어를 만들어 사용해야 한다. 온라인으로 기록정보를 송수신할 것인지 오프라인 매체로 기록정보를 이동할 것인지에 따라서 구현방식에 차이가 있을 수 있다. 내보내기를 하고 나서 빠짐없이 정확하게 기록정보가 추출되었는지 검증해야 하고, 들여오기를 하기 전에 임시 저장소에서 기록정보가 오류없이 신뢰할 만한 상태로 입수된 것인지 검증해야 한다. 이처럼 물리적 이관에서는 상호간에 절차와 방법을 맞추고 사전 사후에 검증이 필요하며 이에 소요되는 비용이 상당하다.

클라우드 통합저장소가 제공되고 업무시스템과 기록관리시스템이 동시에 해당 저장소에 접근할 수 있는 환경이 제공된다면 기록의 이관 작업을 단순화할 수 있다. 물리적인 이관 작업에서 거치는 여러 단계를 생략할 수 있기 때문이다. 물론 클라우드 환경에서도 기록의 이관 작업은 필요하다. 왜냐하면, 기록을 이관하

는 이유는 일정한 기간 동안 진본인 상태로 잘 보관하며 활용하다가 적시에 완전하게 폐기해야 하는 임무를 잘 수행하기 위해서이다. 업무담당자들은 업무 과정에서 기록을 생산, 유통하지만 그 이후의 관리에는 큰 관심을 두지 않는 것이 일반적이다. 따라서, 생산된 기록을 잘 분류하여 활용을 촉진하고 생애주기를 관리하는 역할은 기록관리 담당자에게 주어진 것이라 볼 수 있다. 업무시스템에서 생산한 기록이 클라우드 통합저장소에 저장된 후 이관 시점이 되었을 때 기록관리시스템에서 해당 기록의 메타데이터를 획득하고 관련 콘텐츠 파일들의 품질 확인을 한 후 처리과로부터 관리권한을 이양받는 것으로 이관을 종결할 수 있다.

서울시의 경우, 2011년부터 사용 중인 서울형 업무관리시스템에 현재까지 생산한 모든 결재문서가 고스란히 쌓여있다. 그간 업무관리시스템에서 생산된 결재문서를 함부로 삭제하는 일은 없었으며 안전하게 보관되면서 활용되고 있다. 현행 기록물관리 법령에 따르면 2011년 생산된 결재문서부터 2018년 생산된 결재문서까지가 모두 표준기록관리시스템에 이관되어 있어야 하나 여러 가지 사정 때문에 이관이 미뤄져왔던 것이다. 업무관리시스템의 결재문서를 표준기록관리시스템에 이관하기 시작한 것은 2018년 부터인데, 이관을 시작한 이유는 업무관리시스템에 너무 많은 기록이 쌓여있어 스토리지도 부담이 되고 무엇보다 문서 검색의 속도가 현저히 느려져서 문서를 덜어낼 필요가 절실했기 때문이다. 서울시의 경우에는 최근의 결재문서들은 업무관리시스템에서 검색하고, 10년 이상 오래된 기록은 기록관리시스템에서 찾아보는 관행이 자연스레 정착되어 있다. 그렇다

면, 이런 상황에서도 업무관리시스템의 기록을 생산된 다음 해에 기록관리시스템으로 굳이 이관해야 할까?

서울시의 이관 경험에 그 일부의 답이 존재한다. 서울시는 2011년, 2012년도의 결재문서 기록을 기록관리시스템으로 이관하는 과정에서 타 기관에서도 경험하듯이 암호화된 문서나 열리지 않는 파일, 대용량 파일 별도 전송 건들을 다수 확인하였다. 8년 전 문서의 암호를 기억하고 있는 문서 작성자는 거의 없으므로 이관 과정에서 오류가 난 문서들을 보정하기 어려웠다. 만약 직전 년도의 기록물에 암호화 문제가 발생했다면 문서 작성자가 암호를 기억할 가능성이 좀 더 높아질 것으로 기대할 수 있을 것이다. 별송한 대용량 파일 역시 8년 전의 것은 찾을 가능성이 거의 희박한 반면 1년 이내의 것이라면 찾아낼 가능성이 높아질 것이다. 결국 기록으로서의 품질을 확인하고 필요한 조치를 취할 수 있다는 점에서 이관을 가능하면 이르게 하는 것이 기록관리 측면에서 유리하다는 점을 확인할 수 있다. 콘텐츠 파일들을 이동시키지 않더라도 이런 품질 확인 작업이 포함된 이관 작업은 가능하면 이르게 실시하는 것이 진본기록을 이용가능한 상태로 관리하는데 유리할 것이다.

클라우드 저장소를 통해 업무관리시스템과 기록관리시스템이 콘텐츠 파일을 공유할 수 있는 체계에서는 논리적 이관의 방식을 채택하는 것이 합리적인 선택이다. 그런데, 온나라2.0 체계에서는 여전히 물리적 이관 방식을 고수하고 있다. 중앙행정기관에 온나라2.0의 보급이 완료됨에 따라 이에 조응하는 cRMS의 보급을 서둘러야 하다보니 기존의 표준기록관리시스

템 기능을 그대로 클라우드 서비스화한 것으로 이해된다. 하지만, 전자기록의 이관을 여전히 물리적 이동을 통해 하기로 한 것은 디지털 변환(Transformation)을 도모하지 못한 것으로 아쉬운 결과이다. 시간과 비용이 들더라도 논리적 이관의 방식으로 전환한 cRMS를 완성하고 배포했다면 공공기록의 전자기록 관리가 한 발 더 진전했을 것이다.

3. 기록물 저장공간의 할당 전략

3.1 클라우드 저장소의 특성

우리는 알게 모르게 이미 민간 클라우드 저장소를 다양한 용처에서 유무료로 이용하고 있다. 휴대폰에 넘쳐나는 사진을 저장하기 위해 특정 회사의 클라우드 저장소 계정을 만들었던 경험이나 프로젝트팀원끼리 문서를 공유하기 위해 클라우드 저장소 서비스를 이용한 경험이 있을 것이다. 민간 클라우드에 업로드한 사진이 최종적으로 어떤 종류의 저장매체에 어떤 방식으로 저장되게 되는지 우리는 구체적으로 확인하기 어렵다. 다만 인터넷이 되는 곳이라면 업로드한 사진이나 파일을 원할 때 다운로드받아 열어볼 수 있다는 점을 알고 있을 뿐이다.

클라우드 저장소 서비스를 제공하는 측에서는 거대한 저장 공간을 효율적으로 운영하기 위해 여러 가지 전략을 사용하게 된다. 클라우드 저장소는 다양한 크기와 다양한 유형의 저장매체들을 상호 연결하여 가상화하여 구성된다. 이 때, 저장소에 접근하는 어플리케이션들이 필요로 하는 저장 공간의 양과 품질 기준은

다를 수 있으므로 저장소를 서비스 수준에 따라 논리적인 볼륨으로 나누어 관리할 수 있다. 이용자들의 요구 수준이나 사용 방식에 따라 차별화하여 저장소 볼륨을 배분하고 이에 맞춰 과금할 수 있다. 콘텐츠 파일 하나는 클라우드 저장소에 사본 하나만 저장될 수도 있고 여러 개가 분산되어 저장될 수도 있다. 업로드한 뒤로 거의 찾아보지 않는 콘텐츠 파일들은 일정 시간이 지난 후 온라인(online) 저장매체 중심의 볼륨에서 니어라인(nearline) 중심의 저장매체 볼륨으로 옮겨 둘 수도 있다. 온라인 스토리지와 니어라인 스토리지의 가격에 차이가 나기 때문에 대용량 저장 공간을 구비해야 하는 클라우드 저장소 서비스 제공자 입장에서는 비용 효율을 높이기 위해 이용자군별 저장 전략을 잘 세워나가야 한다.

서울시의 경우 표준기록관리시스템을 구축하여 운영 중이며 스토리지는 외산 NAS로 160 테라 용량을 갖추고 있다. 표준기록관리시스템의 백업 장비는 니어라인 스토리지를 활용하고 있으나 운영 장비에는 온라인 스토리지만 사용 중이다. 최근에 스토리지 부족으로 추가 구매를 하게 되었는데 그 이유는 장기보존포맷 변환을 하면서 스토리지 사용량이 급증했기 때문이다. 저장소가 부족하니 신규 스토리지를 구매하긴 해야 하는데, 장기보존포맷 변환이라는 작업이 일년 내내 지속적으로 이루어지는 작업이 아니고 이관 작업 기간 중의 일시적인 작업이라면 나머지 기간 동안 해당 스토리지는 유휴자원이 되므로 자원의 낭비를 감수할 수 밖에 없다. 단독 서버로 구축된 기록관리시스템은 저장소의 수요가 최대치에 달하는 수준까지를 고려하여 스토리지를 구비해야만 작업에 지

장이 없게 된다. 일시적으로 자원이 추가로 필요할 때, 클라우드 저장소를 사용하면 효율성이 극적으로 높아진다. 클라우드 저장소는 저장 공간의 여유분을 형성해 두고 스토리지를 필요로 하는 서비스에 필요한 만큼 할당을 해주었다가 사용 후 불필요해진 스토리지를 반납 받는 방식으로 운영할 수 있다. 이것은 저장소를 공유하는 서비스들이 최대치의 스토리지를 필요로 하는 시점이 어느 정도 분산되는 것으로 가정한 시나리오이다. 클라우드 저장소를 여러 서비스가 공용으로 사용하는 환경에서 가능한 운용방식이며, 모든 서비스가 동시에 최대의 저장 용량을 필요로 하는 경우가 발생하면 클라우드 저장소도 한계에 도달하여 도움이 되지 않을 수도 있다.

3.2 기록의 저장소 요건

기록관의 업무에 필요한 저장소의 요건과 기록의 특징을 살펴보면 다음과 같다. 첫째, 기관 전체의 기록이 한 군데로 모이므로 대용량의 저장소가 필요하고, 둘째, 공개재분류나 보존포맷 변환과 같은 작업을 할 때는 일시적으로 대량의 기록에 접근해야 하고, 셋째, 포맷 변환 등의 작업을 위해서는 임시저장소가 필요하며, 넷째, 보존기간이 만료되면 복구 불가능한 상태로 삭제해야 하고, 다섯째, 기록관에 신규로 이관되어온 기록에 비해 보유한 시간이 오래된 기록의 활용성은 줄어드는 것이 일반적이다. 첫째, 둘째, 셋째 요건을 고려했을 때, 기록관에는 클라우드 기반의 저장소가 일반 저장소보다 유리하다. 클라우드는 좀 더 저렴한 디스크들을 다량 구입하여 큰 용량의 스토리지를 가

화하여 구성할 수 있으며, 일시적으로 필요한 저장소를 자동 스케일 아웃하여 할당했다가 불필요해지면 자동 스케일 인하여 다른 곳에 활용할 수 있는 체계이기 때문이다. 넷째의 요건은 기록의 생애주기별로 저장 공간을 할당함으로써 삭제의 효과성과 효율성을 높이기 위한 전략이 필요한 대목이다. 다섯째의 특징은 온라인, 니어라인, 오프라인 저장매체를 적절하게 구비하여 저장소를 꾸릴 필요성을 제기한다.

위의 요건을 확인하기 위해서는 우선 기록관에서 보유한 기록의 유형별 통계와 이용 통계가 필요하다. 서울시 기록관의 경우 표준기록관리시스템을 사용하고 있는데 시스템에 기본 이용 통계를 보는 기능이 없었던 터라 유지보수사업팀에 이용 통계를 산출하도록 요청하였다. 2019년에야 웹로그 분석 도구를 설치하고 검색엔진과 문서보안 솔루션 등에서 확보된 데이터를 근거로 통계를 산출하기 시작했다. 2019년 6월의 통계를 보면 총 방문자 수는 2,268명이며, 57,882번 기록관리시스템 기능이 사용됐고, 그 중 11,727번은 기록물 건의 내용 정보를 들여다 본 횟수였다. 전체 보유 기록에 대비하여 활용한 기록이 1~2%에 불과하므로 이 정도의 활용도는 확연히 현용 단계를 지난 준현용 혹은 비현용 단계의 활용 수준이라 할 수 있다. 만약 저장 비용을 절감하고자 한다면, 지속적으로 빈도수가 높은 검색어를 분석하고 자주 살펴본 기록물 철 메타데이터와 기록물 건 메타데이터를 분석하여 활용도가 높을 가능성이 있는 콘텐츠들을 온라인 스토리지에 저장하고 나머지는 니어라인 스토리지에 배치하는 전략을 고려할 수 있을 것이다.

기록관리시스템에는 동일 내용을 담고 있으

나 용도에 따라 포맷이 다른 콘텐츠 파일이 복수로 존재할 수 있다. 예를 들면, 디지털 객체의 썸네일 이미지, 저해상도의 디지털 사본, 고해상도의 원본 등이다. 이 경우 검색의 용이성을 위해 썸네일 이미지는 메타데이터와 함께 필수적으로 제공하고, 저해상도의 디지털 사본은 온라인 스토리지에 저장하여 검색 결과로 바로 조회할 수 있도록 제공하며, 고해상도 원본 파일은 니어라인이나 오프라인 스토리지에 저장하여 특별히 원본을 요청하는 경우에 제공하도록 서비스 정책을 구상할 수 있다. 저해상도의 사본만으로도 이용자의 정보 요구가 충족될 수 있고, 고해상도의 원본에 대한 요구는 추가 비용을 청구하는 정책을 취할 수도 있을 것이다. 특히, 시청각 기록물의 경우 디지털화했을 때 고해상도의 원본은 대용량 파일인 경우가 많다. 하나의 기록 내용에 대해 복수 개의 사본이 존재할 때, 각 사본 별로 용도를 정하고 그에 맞는 저장 매체 종류를 선택하도록 하는 것을 고려해야 한다.

서울시의 경우 시청각기록물은 메타데이터만 기록관리시스템에 등록하고 내용물은 DVD, 테이프 등 원래의 저장매체 상태로 보관하고 있다. 종이문서의 경우 30년 이상 보존기간을 가진 기록은 해마다 디지털화 작업을 하고 있는데, 디지털화 결과물인 pdf파일들은 기록관리시스템 온라인 스토리지에 저장하고 jpg 파일들을 DVD에 수록하여 서고에 보관하고 있다. 현재까지는 온라인 스토리지의 용량이 기록관에서 보유한 기록정보의 양을 충분히 담을 수 있다. 하지만, 시청각기록물의 디지털화를 하게 된다면 결과물을 온라인 스토리지에 모두 담기는 어려울 것으로 예상된다. 결국 활용도가 높

을 것으로 예상되는 콘텐츠들을 선정하여 일부만 온라인 스토리지에 올려야 하는데, 어떤 기준으로 스토리지를 배정할 것인가에 대해 원칙을 수립하는 것이 중요한 과제가 될 것이다.

3.3 생애주기 기반의 저장 전략

기록관에서 보유한 기록이 대량이고, 폐기든 이관이든 결국 100%의 기록정보가 처분대상이라는 점을 감안한다면 기록의 생애주기를 고려한 저장 전략을 세울 필요가 있다. 나아가 기록의 생애주기와 정보시스템의 생애주기를 연관하여 관리하는 것이 효과적이다. 기록관리 시스템의 저장소 구성 시에 최소한 장기보존 대상 기록과 유한기간 보존 대상 기록을 저장하는 공간을 구분하는 것을 고려할 수 있다. 유한기간 보존 대상 디지털 객체들을 저장하는 저장소에는 주기적으로 콘텐츠 파일들이 신규 유입되고 또한 폐기 처분되는 과정을 거치게 된다. 따라서 잦은 파일 입출력의 성능과 정확성을 보장하고, 잦은 삭제에 따라 발생하는 디스크 조각(Fragmentation)을 잘 관리할 수 있는 저장 매체를 선정하는 것이 필요하다. 장기보존 대상 디지털 객체들은 가능하면 그 저장소에서 저장매체가 노후화되어 교체할 때까지 이동없이 보존되는 것이 여러모로 유리하다. 콘텐츠 파일이 다른 저장 공간으로 이동되는 시점에 무결성 훼손의 위험도가 높아지기 때문이다. 기록관에서 장기보존 대상의 콘텐츠 파일들만 별도의 저장 매체에 따로 저장해 두었다가 영구기록물관리기관에 매체 단위로 이관하는 것도 고려해볼만 하다.

단일한 하나의 저장소를 사용하는 체계에서

는 의미가 없지만 다수의 디스크를 엮어서 논리적 개념의 저장소를 만드는 클라우드 체계에서는 또 다른 접근도 가능해진다. 클라우드 저장소의 경우 하이브리드한 디스크들을 엮어서 논리적인 하나의 저장소를 제공하게 된다. 이때, 논리적인 저장공간을 볼륨으로 나누어 정할 수 있는데, 단기간 보존하는 기록은 상대적으로 저렴한 디스크 영역에 보관하고 장기보존 대상 기록을 위해서는 장기간 보존 안정성을 제공해주는 고비용 스토리지 영역을 할당하는 방식으로 저장소를 운영할 수도 있다. 이 경우, 장기보존 기록의 저장매체 리프레쉬 횟수를 줄일 수 있어 그만큼 물리적 이관에 따르는 위험성을 최소화할 수 있을 것이다. 단기간 보존하는 기록의 경우, 보존기한이 만료되는 연도가 같은 기록을 같은 저장소 영역에 저장하도록 한다면 폐기 결정 후 데이터를 삭제가 훨씬 용이해지고 저장소 재활용도 용이해질 것이다.

기록관의 저장소를 구성할 때 디지털 저장매체에 여러 종류가 존재한다는 점을 고려해야 한다. 자기방식과 광학방식을 구분하는 것, 온라인/니어라인/오프라인을 구분하는 것, NAS와 SAN을 구분하는 것, 클라우드 저장소 여부 등 다양한 선택지가 존재한다. 그러나, 대부분의 기록관리자들은 현재까지의 기록관리시스템을 통해 경험한 저장매체의 종류가 한정적이다 보니 상상력을 발휘하여 저장 전략 구상하기가 어려운듯하다. 특히, 기록의 저장매체에는 온라인 저장소만이 가능한 것처럼 사고하는 경향이 있는데, 현실에서는 모든 기록의 콘텐츠 파일을 온라인 저장소에 보관하는 것은 비용 측면에서도 관리 측면에서도 바람직하지 않다. 특히 영구기록물관리기관에서 보관하는 상당

량의 기록들은 니어라인이나 오프라인 저장소를 활용하는 것이 유용하다. 기록관에서도 앞에서 언급한 바대로 대용량 고화질의 시청각 기록물의 디지털 원본들은 니어라인이나 오프라인 저장소에 두는 것이 비용 측면에서 지속 가능성이 있다. 온라인 저장소에는 이용자들이 자주 접근하여 활용하는 기록물을 중심으로 활용용 사본 콘텐츠를 저장하도록 배치하는게 유리하다. 기록관이 서비스 활동 초기에는 기록의 목록과 썸네일 수준의 콘텐츠만 온라인 저장소에 두고 운용하다가, 이용자의 요구가 있는 콘텐츠 파일을 순서대로 온라인 저장소에 올려두고, 온라인 저장소가 부족해질 때는 가장 활용한 지 오래된 콘텐츠 파일부터 오프라인 저장소로 내려 공간을 확보하는 방식으로 운용할 수도 있다.

기록관리시스템을 설계할 때, 위에서 언급한 요건들에 따라 저장매체들을 논리적인 저장소 영역으로 구분하고, 콘텐츠 파일들을 보존기간이나 공개여부, 활용도 등에 따라 적절한 위치에 구분하여 저장하고, 콘텐츠 파일들 간의 원본/사본 관계를 관리하고, 접근 요청이 들어왔을 때 특정 콘텐츠 파일을 찾아주고, 주기적으로/수시로 무결성 검증을 실행하는 저장소 서비스를 만들 필요가 있다. 현재 표준기록관리시스템에서 저장매체로 WORM스토리지를 사용하도록 하는 것은 장기보존 대상 기록물에 대해서는 좋은 선택이지만 유한기간 보존하다가 폐기할 기록물의 저장매체로는 불리하게 작용하는 점도 있다. 또한, 장기보존 기록물의 경우 NEO 장기보존포맷으로 변환을 하고, 재변환을 하는 등의 절차를 거쳐야 하는 경우 파일이 고정되어 불변하는 것이 아니므로 WORM

스토리지를 잘 활용하는 것이 용이하지 않다. 적합한 용처에 WORM스토리지를 활용하도록 하되 그 밖의 다양한 저장매체들도 용도에 맞게 배치하는 융통성있는 저장 전략이 요구된다.

4. 파일 중복을 최소화하는 관리

4.1 콘텐츠 중복 현황

기록관에서 전자기록관리를 할 때 저장소 운영의 핵심은 콘텐츠 파일 관리이다. 서울시의 경우 현재 사용 중인 업무관리시스템은 온나라 1.0을 개작한 버전인데, 스토리지에 저장된 본문파일과 첨부파일 중 서로 중복되는 콘텐츠 파일이 전체의 4% 가량된다. 다중 부서에서 접수 받는 기록의 경우 첨부파일을 공유하는 구조로 설계되어 있어서 중복이 4% 수준인 것이다. 업무관리시스템에서는 다중 부서에 접수된 문서의 경우 콘텐츠 파일들이 한 부만 서버에 저장되고 동일 문서를 접수한 여러 부서에서는 메타데이터만 생성되도록 관리된다. 그러나, 이 접수 문서가 기록관리시스템으로 이관되면서 콘텐츠 파일들을 건 단위로 각각 포함하는 방식으로 처리되어 기록관리시스템의 저장소에는 중복 콘텐츠 파일이 급격히 증가하는 구조이다. 앞에서 언급한 바와 같이 서울시 기록관리시스템에는 약 70%의 파일이 중복 콘텐츠이다.

서울시의 경우를 관찰해 보니 업무담당자들이 업무관리시스템에서 문서관리카드를 만들 때 본문은 기안기를 사용하여 새로이 작성하게 되지만 첨부파일에는 다른 결재문서 본문이나 첨부파일을 PC에 다운로드해 두었다가 업로드

하는 경우가 많았다. 업무관리시스템에 있는 결재문서 본문을 참조할 수도 있는데 굳이 PC에 다운로드 받아 사용하는 것은 관인을 생략하거나 결재자의 승인 정보를 없앤 문서를 첨부하고 싶어서인 경우도 있다. 서울시는 정보소통광장(opengov.seoul.go.kr)을 통해 과장급 이상 결재문서 원문을 시민에게 제공하고 있는데 이 시스템 스토리지에 저장된 파일 중 중복되는 콘텐츠 파일은 전체의 약 30%가 넘는다. 실제로는 훨씬 많은 콘텐츠 파일들이 중복되고 있으나 관인을 생략하거나 결재정보를 감춘 버전으로 만든 문서의 경우 내용은 동일해도 해시값이 달라져서 중복으로 계산되지 않고 있는 것이다.

콘텐츠 파일이 다량 중복되어 여러 측면의 비효율이 발생하고 있기에 온나라2.0과 cRMS가 설계될 때 중복 제거가 적용될 것으로 기대하였다. 그러나 아쉽게도 온나라2.0이나 cRMS에서도 콘텐츠 파일의 중복은 그대로 유지되고 있다. 행정안전부로부터 전해들은 바로는 초기에 온나라 문서2.0 설계 시 중복 콘텐츠 파일을 제거하고 하나의 콘텐츠 파일을 다중으로 참조하도록 만들었다고 한다. 하지만, 국가기록원 협의 과정에서 콘텐츠의 중복을 요청하여 설계를 변경했다고 한다. 국가기록원이 콘텐츠의 중복을 요청한 근거는 소위 기록 건(item)들은 서로 명확히 구분되어야 한다는 원칙을 시스템 상에서 관철하기 위해서라고 한다. 동일한 콘텐츠 파일을 참조하는 컴포넌트 개수만큼 저장소에 파일을 중복하여 저장해야만 건별로 구분이 명확해진다는 것이다. 기록 건들이 서로 간에 명확히 구분되어야 한다는 것은 개념적으로도 논리적으로도 옳다. 그러나, 개념과 논리를

구현하는 기술적이고 물리적인 방식은 다양한 선택이 가능하다. 하나의 컴퓨터 파일이 서로 다른 기록 건에 컴포넌트로 포함된다고 해도 기록 건이 서로 겹치지 않고 명확히 구분되어 활용되고 보존되도록 할 수 있기 때문이다. 지금 클라우드 서비스로 구현된 업무관리시스템은 하드웨어를 여러 기관이 공유하여 사용하는 구조이다. 유사한 논리로 클라우드 업무관리시스템 배포 초기에 기관별로 서버가 구분되어야 하는데 여러 기관의 서비스를 하나의 클라우드 플랫폼에 얹는 것이 불법 아니냐는 논의가 있었다고 한다. 여러 논쟁 끝에 클라우드 플랫폼 안에 데이터베이스 인스턴스를 기관별로 구별하여 동작하는 방식으로 기관 간 구분을 명시화하고, 클라우드 저장소는 논리적인 경계 구분을 하는 것으로 마무리되었다고 한다. 클라우드 서비스가 공공영역에 도입되면서 기관별로 서버나 스토리지를 독립적으로 구축해야만 한다는 원칙은 이미 폐기된 것이다. 즉, 물리적인 구분보다는 논리적인 구분이 필요한 환경이 되었다는 것이다. 이처럼 새로운 기술적 환경에서는 개념적, 논리적 원칙이 물리적으로 구현될 때 새로운 방식으로 적용될 수 있어야 한다.

국가기록원의 또 다른 주장은 같은 콘텐츠 파일이라도 결재문서 마다 역할이 다르기 때문에 따로 저장되어야 한다는 것이다. 동일한 콘텐츠 파일이 건 별로 역할이 다를 수 있는 것은 당연한 상황이다. 콘텐츠의 파일이 해당 결재문서에서 본문의 역할인지, 첨부1의 역할인지, 첨부2의 역할인지는 해당 컴포넌트의 메타데이터 값으로 구분되어 저장되면 된다. 콘텐츠 파일은 기록의 내용을 품고 있는 것이건에서의 역할을 의미하는 것은 아니므로 콘텐츠 파

일을 중복적으로 저장해야만 하는 것은 아니다. 콘텐츠 파일과 컴포넌트의 관계는 실제로 일대다(1:多)인데 국가기록원은 일대일(1:1) 관계로 만들어 관리하도록 시스템을 설계한 것이다. 앞에서 클라우드 환경이 도입되면서 기관별 서버 독립의 원칙이 달리 적용되었듯이 기록 건의 구성도 물리적으로 구분되는 방식이 아니라 논리적으로 구분되는 방식으로 이해한다면 설계는 변경될 수 있다.

4.2 중복 제거를 위한 방법

업무관리 및 기록관리와 관련하여 콘텐츠 파일의 관리에 몇 가지 문제점이 있다. 중복 콘텐츠 파일이 다량으로 생산된다는 점과 콘텐츠 파일 단위의 식별과 통제가 미비하며 진본성 확인체계가 불충분하다는 점이다. 이에 콘텐츠 파일 관리체계에 대해 몇 가지 제안을 하고자 한다.

동일한 콘텐츠 파일이 여러 문서, 여러 기록 건에 중복적으로 활용될 수 있다. 그런데, 콘텐츠 파일이 한 본만 저장되어도 여러 기록 건에서 참조하여 활용되는데 문제가 없음에도 이렇게 중복적으로 파일을 만들어 관리하게 된 것은 어떤 연유일까? 현재까지 추적 인터뷰를 한 바에 따르면, 전자문서시스템에서도 업무관리 시스템에서도 접수문서의 경우 당연히 중복이 많이 발생하므로 여러 접수 기록에서 동일한 콘텐츠 파일에 대해 링크 정보를 보유하는 방식으로 콘텐츠 파일의 중복을 없앴다고 한다. 하지만, 생산 문서의 경우에는 기안자가 첨부

문서를 직접 업로드하기 때문에 중복 여부를 알 수 없다고 한다. 기안자가 업로드하는 문서가 이미 서버에 존재하는 콘텐츠 파일과 동일본일 때는 굳이 중복하여 콘텐츠를 저장하지 않고 기존 콘텐츠 파일의 링크 정보를 연결시켜주면 되는데, 이에 콘텐츠 파일 중복 여부에 대한 확인절차를 갖지 않는다는 점이 문제인 것이다. 콘텐츠 파일의 중복 여부를 신속하고 정확하게 확인하기 위해서 모든 콘텐츠 파일의 해쉬값을 메타데이터로 관리하고 있어야 한다. 새로 업로드되는 콘텐츠 파일의 해쉬값을 이용하여 기존 것들과 중복 여부를 확인하면 된다. 저장소의 모든 콘텐츠 파일에 대해서는 UUID(Universally Unique Identifier)⁵⁾를 부여하여 관리하도록 한다. UUID체계와 해쉬값을 관리함으로써 콘텐츠 파일별 식별과 진본성 확인이 용이해진다.

서울시도 업무관리시스템을 클라우드 기반으로 고도화할 계획을 수립 중인데, 서울시 뿐만 아니라 25개의 자치구 간에도 콘텐츠를 파일을 공유할 수 있도록 클라우드 저장소를 운영할 예정이다. 서울시와 25개 자치구 간에 문서 수발신이 자주 이루어지고 있어 여기서 발생하는 중복 콘텐츠 파일의 양이 상당한데 통합 클라우드 저장소를 공동활용하게 된다면 중복을 최소화하게 될 것이다.

최근 국가기록원은 전자기록의 장기보존과 관련한 정책의 수정을 검토 중이다. 국가기록원이 전자기록 장기보존 정책과 전자기록 보존포맷 정책에 관해 밝힌 바에 따르면, 법령에서 사용 중인 문서보존포맷, 장기보존포맷이라는

5) UUID는 개념적으로 MoReq2010에서 제시하는 내용을 준용하는 것으로 제안한다. 최소한 기록 건, 컴포넌트, 콘텐츠 등에 UUID를 부여하여 관리하는 것이 기록관리 체계에 도움이 될 것이라 판단된다.

용어를 각각 '보존포맷', '장기보존 패키지'라는 용어로 전환하겠다고 한다. 그간 두 개의 포맷이 의미하는 것이 무엇인지를 기록학계 내에서 설명하는 것이 어려웠는데 OAIS참조모형에 근거하여 용어를 바꾸는 것은 학계 내부의 이해도를 높이는 좋은 선택이라 생각한다. 다만, 그간 국가기록원이나 학계에서 장기보존포맷, 즉 패키지에 관해 OAIS참조모형을 오독하여 적용한 부분이 있어 이에 관해 논의를 할 필요가 있다⁶⁾. 요약하자면, 패키지는 물리적인 단일 파일로 만드는 것을 의미하는게 아니라 논리적, 개념적인 컨테이너라는 점이다. 물리적 패키징이란 건이나 철 단위의 기록정보 완전성을 기하기 위해 메타데이터와 모든 콘텐츠 파일들을 하나의 컨테이너(예를 들면, xml파일이나 zip파일)에 묶는 과정이다. 논리적 패키징이란, 데이터베이스와 파일서버에 메타데이터와 콘텐츠 파일들을 저장해둔 채 하나로 묶어야 할 대상들을 식별해두고 관리하는 방식이다. 건 단위로 물리적 패키징을 할 때는 콘텐츠 파일이 중복적으로 저장될 수 밖에 없지만, 논리적 패키징을 할 때는 콘텐츠 파일을 중복 저장할 필요가 없는 것이다. 패키지의 의미를 물리적 패키지로만 이해하게 되면 콘텐츠 파일의 수많은 중복을 감수해야 한다는 논리로 발전하게 된다. 물론, 영구기록물관리기관에서 중요기록물의 이중보존을 위해 별도 저장매체에 수록을 해둘 때에는 영구기록관리시스템이 없이도 기록을 식별할 수 있도록 기록물 건과 철 단위

로 모든 정보를 담아내고 자기 기술 정보를 충분히 넣어 물리적으로 패키징해야한다. 하지만, 수시로 콘텐츠 파일에 접근해야 하는 검색 및 활용 시스템에서는 논리적 패키징이 여러모로 유리하니 시스템을 그렇게 설계하면 좋겠다는 것이다. 이관 시에도 물리적 패키지를 주고 받게 되면 디지털 객체의 용량이 훨씬 커지므로 불리해진다. OAIS참조모형에서도 패키지의 의미를 개념적, 논리적 패키지라는 점을 강조하고 있는 만큼 패키지를 구현할 때 용도와 상황에 맞게 논리적 패키징, 물리적 패키징 방법 중에 잘 선택할 필요가 있다.

4.3 생산 및 관리 과정의 변화

클라우드 저장소에 저장되는 콘텐츠 파일의 중복을 최소화하는 방식으로 관리하고 기록생산시스템과 기록관리시스템이 콘텐츠 파일들을 공유하게 되면, 그에 맞춰 기록의 생산 및 관리 과정이 변화해야 한다. 기록생산시스템에서도 기록관리시스템에서도 저장소에 새로운 콘텐츠 파일들이 CRUD(Create, Read, Update, Delete)⁷⁾될 때의 처리 방식이 공통적으로 변경되어야 한다. 즉 콘텐츠들이 처음 생성되거나 업로드될 때, 콘텐츠들에 접근하여 내용을 열어볼 때, 콘텐츠들의 비트스트림이 변경될 때, 콘텐츠를 삭제할 때마다 공통적인 처리가 요구되는 것이다. 콘텐츠들의 중복을 최소화하면서 기관 내의 핵심 콘텐츠를 정보자산으로 효과적

6) 2019년 3월 27일 국가기록원 연구세미나에 참석하여 토론문에서 밝힌 내용이다.

7) CRUD는 대부분의 컴퓨터 소프트웨어가 가지는 기본적인 데이터 처리 기능인 Create(생성), Read(읽기), Update(갱신), Delete(삭제)를 묶어서 일컫는 용어이다. 사용자 인터페이스가 갖추어야 할 기능(정보의 참조/검색/갱신)을 가리키는 용어로서도 사용된다. <https://ko.wikipedia.org/wiki/CRUD> 참조.

으로 관리하기 위해 다음과 같은 작업 방식을 제안한다. 먼저, 콘텐츠 파일들이 저장소에 새로 만들어질 때는 고유 지문이라 할 수 있는 해시값을 만들고 이를 포함하여 UUID를 만들어 저장소 콘텐츠 파일 목록에 등재한다. 해시값이 동일한 콘텐츠 파일이 등장할 때는 저장소에 저장하지 않고 기존의 콘텐츠 파일을 참조하도록 처리한다. UUID는 콘텐츠 파일들이 저장소 내 뿐만 아니라 외부 어디에 유통되더라도 진본여부와 출처를 확인하는 기본체계로 작동하도록 설계한다. 하나의 콘텐츠 파일은 하나 이상의 기록 건 컴포넌트와 일대다(1:多)의 관계로 연결한다. 해시값이 동일한 콘텐츠 파일이 파일명이 다르게 입수될 수도 있다. 컴포넌트 별로 콘텐츠 파일의 역할 정보와 파일명, 포맷정보, 해시값을 메타데이터로 등록하도록 한다.

해시값이 동일한 콘텐츠 파일의 중복을 없애면 검색 서비스가 향상될 수 있다. 온나라 문서 2.0의 경우 콘텐츠 파일의 텍스트 검색기능을 제공하고 있는데, 동일한 콘텐츠 파일이 여러 사본 존재하면 검색의 결과도 중복될 것이므로 이용자 입장에서 비효율적이게 된다. 검색 결과로 유일한 콘텐츠 파일이 조회되고, 필요시 해당 콘텐츠 파일을 참조하고 있는 다중의 컴포넌트 목록을 연결하여 조회할 수 있도록 하는 것이 좀 더 고도화된 서비스라 할 것이다. 서울시의 경우 정보소통광장 저장소의 콘텐츠 파일에 대해 해시값을 생성하여 30% 가량의 중복이 존재한다는 것을 확인한 바 있다. 서울시는 이 결과를 이용하여 정보소통광장의 이용자의 서비스를 보완하는데 활용하고 있다. 이용자가 특정 결재문서 원문을 조회한 경우, 동일한 콘텐츠 파일이 포함된 타 결재문서를 연

관문서로 보여주도록 한 것이다. 동일 콘텐츠 파일을 공유하는 결재문서에 대해서 이용자가 관심을 가질 가능성이 높기 때문이다. 나아가 동일 콘텐츠 파일이 참조된 기록을 추적해보면 조직 간 업무적 연관도를 분석할 수 있으며, 조직 내에서 협업이 필요한 지점을 확인하여 이후 업무개선의 대상으로 설정할 수도 있을 것이다. 서울시의 경우 업무관리시스템, 기록관리시스템, 정보소통광장에서 통합 클라우드 저장소를 함께 사용하도록 한다면 콘텐츠 파일의 중복 저장을 획기적으로 줄이고, 콘텐츠 파일의 식별체계를 단일화할 수 있을 것이다.

동일한 콘텐츠 파일이 여러 기록 건에 중복적으로 활용되고 있다는 것을 파악하고 나면 기록관리 및 통제에 일관성을 기하기가 용이해진다. 현재 정보공개법령에 따라 결재문서 기록은 공개, 부분공개, 비공개 등으로 설정을 하여 시민에게 제공하도록 하고 있다. 공개여부를 설정하는 단위는 본문파일이나 첨부파일 등 각 콘텐츠 파일이 기본단위가 된다. 시민에게 서비스할 때에도 콘텐츠 파일 별로 원문공개를 하고 있다. 그런데, 동일한 콘텐츠 파일이 서로 다른 결재문서에 참조되면서 공개여부 값이 각기 다르게 설정된다면 문제가 아닐 수 없다. 현재는 업무담당자들이 결재문서 건별로 공개여부 값을 설정하고 있으므로 중복 콘텐츠 파일 간에 일관성을 확보하기 어려운 상태이다. 비공개 콘텐츠 파일의 경우 조직 내에서 어느 범위까지 볼 수 있도록 허용할 것이냐는 문제도 마찬가지로 일관성을 확보하기 어려운 상태이다. 중복을 허용하지 않는 체계에서는 업무담당자들이 특정 콘텐츠 파일에 대해 공개여부 값과 접근권한 범위를 하나의 값으로 설정하여

일관되게 통제할 수 있다.

저장소에 새로운 콘텐츠 파일들이 만들어지는 경우는 기록생산시스템에서 완전히 새로운 내용의 기록이 생성될 때와 이미 저장소에 있는 콘텐츠 파일을 발췌 가공하여 만들 때이다. 시민에게 제공하기 위한 사본 생성, 기록 보존을 위한 포맷 변환, 원문공개용 개인정보를 가린 발췌사본, LOD 등 개방형 접근을 위한 XML/JSON/RDF/RIC 등을 발행할 때 등이 포함된다. 내용은 동일해도 포맷이 변하면 해시값이 바뀌므로 고도화된 기록관리를 위해서는 내용적 동일성 여부를 확인하기 위한 체계까지 갖추는 것이 요구된다. 예를 들어, 해시값이 다르지만 내용은 동일한 콘텐츠 파일들, 예를 들어 ODT 포맷의 결재문서 본문과 PDF/A-1로 변환한 콘텐츠 파일의 경우 해시값은 다르지만 내용적으로 동일한 버전임을 표기해두어야 한다. 가능하다면 내용적으로 같은 버전의 콘텐츠 파일들이 왜 생성되게 되었는지 사유와 용도를 함께 메타데이터에 기재해 두면 유용할 것이다.

대부분의 기록관에서는 기록관리시스템 안에 이미 많은 중복 콘텐츠 파일들을 보유하고 있을 것이다. 그러나 기록관리시스템에 위와 같이 포맷만 다르고 내용이 동일한 콘텐츠 파일을 추적할 수 있는 메타데이터가 없을 것이다. 이런 상황에서는 인공지능 딥러닝 알고리즘을 이용하여 콘텐츠 파일의 공통 속성(feature)을 분석하여 콘텐츠 간의 동일성을 추정해볼 수 있다. 또한, 원천 콘텐츠 파일과 이로부터 파생된 가공 콘텐츠도 추적해 볼 수 있다. 이런 분석이 가능해지면 조직 내에서 원천 지식 콘텐츠를 만들어 낸 최초의 공헌자를 파악하는 등 지식관리에 새로운 접근이 가능해질 것이다.

하나의 콘텐츠가 필요에 따라 여러 사본 파일로 제작된 경우, 기록관리 관점에서 본다면 보존기간이 지나 폐기처분하거나 영구기록물관리기관에 이관하는 시점에 사본들을 모두 함께 처분해야 한다. 콘텐츠 파일의 목록의 정보가 있어야 사본들의 위치를 추적하여 완전한 처분을 실행할 수 있게 된다. 기록의 처분과정에서는 처분 대상 기록의 컴포넌트별로 관련 콘텐츠 파일을 삭제할 것인지 확인하는 절차를 추가해야 한다. 클라우드 저장소에 보관 중인 전자기록을 영구기록물관리기관에 이관하거나 보존기간 만료로 폐기하기로 결정했다면 메타데이터에 처분이력을 추가하고 존재를 증명할 잔존 메타데이터만 남기고 나머지는 삭제해야 한다. 또한 컴포넌트를 구성하는 콘텐츠 파일들을 복구불가능한 상태로 저장소에서 삭제해야 한다. 그러나, 새로이 제시하는 저장소 운영 방식에서는 콘텐츠 파일 삭제 부분의 절차를 다소 변경해야한다. 콘텐츠 파일과 기록 컴포넌트 관계가 일대다의 관계이므로, 기록 건 하나가 폐기되거나 이관된다고 해서 관련 콘텐츠 파일을 바로 삭제해서는 안되고, 해당 콘텐츠 파일을 참조하는 다른 기록 건들이 남아있는지 확인해야 하기 때문이다.

5. 맺음말

클라우드 컴퓨팅 환경이 갖는 특징 중 하나가 플랫폼 자체가 대부분 개방형 코드(Open Source)로 이루어져 있기에 기존의 상용솔루션과는 개발 언어, 개발자군, 마인드 등에서 다른 점이 많다는 것이다. 최신의 전자정부프레

임워크에도 개방형 소프트웨어가 상당수 포함되어 있고, G-클라우드에 메타데이터의 확장이 용이한 NoSQL 데이터베이스관리시스템도 탑재되어 있다. 향후 클라우드 기반의 컨테이너 배포 등 여러 진보적인 기술 개념들을 기록관리시스템의 개발과 배포, 운영 및 유지 분야에서 적용해 가야할 것이다.

클라우드의 장점 중 기록관리와 연관되는 점은 클라우드 저장소를 활용하여 기록객체가 여러 어플리케이션에서 공유될 수 있다는 점이다. 그간 전자기록을 종이기록처럼 물리적으로 이관해온 관행을 극복할 수 있는 확실한 기술적 기반이 마련된 것이다. 물리적 이관은 그 자체로 비용이 많이 들고 무결성이 훼손될 위험에 노출되는 과정이므로 기술적 기반이 따라준다면 물리적 이관을 최소화하여 관리하는 것이 유리하다.

중앙정부는 G-클라우드를 구축하여 활용하고 있으며 민간클라우드를 활용하여 공공서비스를 할 수 있도록 지침을 개정하는 등 클라우드 컴퓨팅을 활성화하는 방향으로 나아가고 있다. 서울시와 같은 지방정부도 클라우드 컴퓨팅 환경을 점차 확대하여 적용하고 있으나 아직은 가상머신 수준에 머물고 있으며, 중앙정부 G-클라우드 수준의 광역형 클라우드 기반을 구축하기 위해서는 예산과 조직 확보라는 난관을 해결해야 하는 상황이다.

서울시 은평구처럼 지방자치단체 중에서도 클라우드 온나라2.0을 독립형으로 도입하는 곳이 생김에 따라 국가기록원이 지자체용 cRMS 보급에 대해 검토 중이며, 2019년 중에 가이드라인을 마련한다는 소식도 들린다. 그런데, 독립형 온나라는 클라우드의 장점을 크게 살리기

어려운 구조이다. 앞에서 살펴본 바와 같이 cRMS 또한 클라우드의 장점을 살리지 못하는 구조이다. 지자체들은 온나라와 cRMS 도입 시 이러한 상황을 정확히 인지할 필요가 있다. 서울시는 2019년 하반기에 차세대 업무관리시스템을 위한 BPR/ISP를 수행하고 있고, 2020년에 클라우드 기반의 업무관리시스템을 구축할 예정이다. 서울시는 앞에서 살펴본 여러 문제들을 바로잡아 새로운 설계를 하고자 한다. 클라우드 저장소에서는 콘텐츠 파일을 여러 어플리케이션에서 URL로 용이하게 공유할 수 있고, 이 장점을 살려 처리과-기록관 단계로의 이관은 메타데이터만 복사하고 콘텐츠 파일에 대한 관리권한을 이양받는 방식으로 설계할 예정이다. 또한, 클라우드 저장소를 장기와 단기 저장용으로 논리적 볼륨을 나누어 콘텐츠 파일을 생애주기 관리에 용이하게 저장할 예정이다. 콘텐츠 파일마다 UUID를 부여하여 관리하고, 해쉬값을 필수 메타데이터로 관리하며, 콘텐츠 파일을 참조하고 있는 기록 건과의 관계 정보를 관리할 것이다. 업무관리시스템에서 만들어진 결재문서는 서울시 정보소통광장을 통해 다음 날 원문공개가 되고 있는데 개인정보와 민감정보를 가린 문서를 만들어 공개하고 있다. 콘텐츠 파일의 가공된 버전들도 UUID를 통해 추적할 수 있도록 체계를 만들어갈 예정이다.

디지털 시대의 기록관리를 논의할 때 비평의 기준은 기록관리 도구들이 디지털 속성에 친화적이냐 여부로 볼 수 있다. 아날로그적 원리를 그대로 답습하면서 표면적으로만 디지털화하는 것은 근본적인 한계를 노정하게 된다는 점에 주목해야 한다. 또한, 향후 미래 기술로 떠오르는 기술들과의 호환성, 연계성, 확장성을 보장하는

쪽이 유리하다는 점에 주목해야 한다. 따라서, 디지털 환경에 맞춰 전자기록을 관리하고자 한다면, 아날로그의 디지털화 수준이 아니고 디지털의 본성에 맞춰 사고할 필요가 있다. 예를 들어, 클라우드 컴퓨팅 기술이 보편화되면 그 기술

을 기반으로 기록이 생산되고 관리되는 과정을 면밀히 살펴보고, 자연스럽게 전자기록이 흐르면서 이용되도록 설계하고, 기록관리에서 특별히 신경써야 할 진본성과 장기보존이라는 목표에 유리하도록 기술을 활용하는 것이 필요하다.

참 고 문 헌

국가기록원 (2016). 클라우드 기반의 RMS 전환 추진현황. 대전: 국가기록원.

국가기록원 (2017). 차세대 기록관리모델 재설계 연구 개발 보고서. 대전: 국가기록원.

국가기록원 (2019). 정부지식 공유활용기반 고도화 4차(중앙부처 클라우드 기록관리 확산) 완료보고회. 대전: 국가기록원.

김주영, 김순희 (2019). 클라우드 저장소를 활용하여 기록생산시스템에서 기록관리시스템으로 전자기록물을 이관하는 방안에 관한 연구. 한국기록관리학회지, 19(2), 1-24.
<http://dx.doi.org/10.14404/JKSARM.2019.19.2.001>

대통령비서실 (2004). 기록관리체계 재설계 방안 연구.

서울특별시 (2016). 서울기록원 정보화전략계획(ISP) 수립 용역.

안대진, 임진희 (2017). 제4차 산업혁명 기술의 기록관리 적용 방안. 기록학연구, 54, 211-248.

이생동 (2014). 클라우드컴퓨팅에 기반한 기록관리시스템 체계 개선방안 연구. 석사학위논문. 명지대학교 기록정보과학전문대학원, 기록관리학과.

정부혁신지방분권위원회 (2005). 기록관리혁신 로드맵.

[관련법령]

- 「공공기록물 관리에 관한 법률」.
- 「전자정부법」.
- 「정부 클라우드저장소(G드라이브) 이용지침」.
- 「클라우드컴퓨팅 발전 및 이용자 보호에 관한 법률」.

• 국문 참고자료의 영어 표기
 (English translation / romanization of references originally written in Korean)

An, Dae-Jin & Yim, Jin-Hee (2017). Application of 4th Industrial Revolution Technology to

- Records Management. The Korean Journal of Archival Studies, 54, 211-248.
- Kim, Ju-Young & Kim, Soon-Hee (2019). A Study on Transferring Electronic Records from Record Production System to Record Management System Using Cloud Storage. Journal of Korean Society of Archives and Records Management, 19(2), 1-24.
<http://dx.doi.org/10.14404/JKSARM.2019.19.2.001>
- Lee, Saeng-Dong (2014). A Study on The Improvement of Cloud Computing-based Records Management System. Unpublished master's thesis, Myongji University, Seoul, Korea.
- National Archives of Korea (2016). Report on Cloud-based RMS Implementation. Daejeon: National Archives.
- National Archives of Korea (2017). A Study on Conceptual Redesign of The Next Generation Record Management. Daejeon: National Archives.
- National Archives of Korea (2019). Advanced Knowledge Base on Utilization of Government Knowledge Sharing 4th(Diffusion of Cloud Record Management at Central Government) Completed report. Daejeon: National Archives.
- Presidential Committee on Government Innovation and Decentralization (2005). Records Management Innovation Roadmap.
- Presidential Office (2004). A Study on Redesign of Recording Management System.
- Seoul Metropolitan City (2016). The Information Strategy Planning for Seoul Metropolitan City Archives.

[Laws]

- Act on The Development of Cloud Computing and Protection of Its Users.
Electronic Government Act
Government Cloud Storage Usage Guidelines.
Public Records Management Act.