

지능형 아카이브 솔루션을 활용한 공개재분류 연구: 2020년 국가기록원 공개재분류 사업을 중심으로

A Study on Access Re-Review Using Intelligent Archive Solutions: Focusing on the Access Re-Review Project of the National Archives of Korea in 2020

송주형(Zoo Hyung Song)

E-mail: sanin@nate.com

(주)디앤디기술 책임연구원, 한국외국어대학교 정보기록학과 박사수로



논문접수 2021-10-19
최초심사 2021-10-21
게재확정 2021-11-05

ORCID

Zoo Hyung Song
<https://orcid.org/0000-0001-7505-5776>

초 록

기록물 공개의 전(前)단계인 공개재분류 업무는 가치있고 중요한 업무다. 그러나 그 의미와 중요성과 별개로 일선 기록연구사에게는 귀찮고 부담스러운 업무이기도 하다. 이에 기록연구사의 업무 부담을 경감해주는 공개재분류 자동화가 제안되기도 했다. 이런 상황에서 국가기록원의 ‘2020년 공개재분류 결과 건별 입력 위탁사업’ 수행 간 공개재분류 솔루션을 활용하였고, 이를 사람이 한 작업과 비교·분석해 보았다. 물론 해당 사업은 공개재분류 솔루션에 대한 분석과 연구가 중심인 연구용역 사업이 아니었고, 상용프로그램을 실험적으로 사용해 본 제한적 결과라는 한계는 존재한다. 그럼에도 지능형 아카이브의 공개재분류에 대한 거시적이고 피상적인 논의만 있는 현 상황에서 기록연구사가 활용가능한 공개재분류 솔루션을 실제 사업에서 활용한 결과를 살펴보는 것은 여러모로 의미가 있으리라 생각된다. 본 논문은 공개재분류 솔루션의 활용 사례에 대한 분석을 통해 공개재분류 업무를 경감시킬 수 있는 실질적이고 미시적인 차원의 논의이다.

ABSTRACT

Access re-review is a valuable and important task, but it is burdensome for archivists. Thus, an access re-review automation was proposed to address this. In this situation, the National Archives of Korea actually utilized the access re-review solution in the performance of the “2020 Access Re-Review Project” and compared and analyzed it with human work. The project was, however, not a research project centered on analysis on access re-review solutions, and it has a limited result in terms of experimental use of commercial programs. Nevertheless, in the current situation where there are only macro and superficial discussions on access re-review of intelligent archives, it would be meaningful to apply the access re-review solution to archivists in real businesses and examine the results. This paper seeks to discuss the practicality that can mitigate the task of access re-review through an analysis of use cases of access re-review solutions.

Keywords: 공개재분류, 기계학습, 지능형 아카이브 솔루션, 전자기록물, 국가기록원
access re-review, machine learning, intelligent archive solution, electronic records,
National Archives of Korea

© 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.

<https://jksarm.koar.kr>

www.kci.go.kr

1. 들어가며

“잡지는 기록물에 날개를 달자”

국가기록원이 2009년에 펴낸 ‘기록물공개재분류 실용매뉴얼’의 제목이다. 기록관 서고 속에 단지 보존만 되어 있는 비공개기록물을, 공개 또는 부분공개로 전환해 국민들이 열람·활용할 수 있게 하는 기록물공개재분류 업무를 잘 표현한 문장이다. 기록물 공개는 국민의 알권리를 증진하고 공공기관이 수행한 업무에 대한 설명책임성을 확보하는 중요한 요소로, 기록 관리의 궁극적 목표와도 맞닿아있다.

이런 기록물 공개의 전(前)단계인 공개재분류 업무는 공공기관 입장에서 무척이나 가치있고 중요한 업무다. 그러나 그 의미와 중요성과는 별개로 일선 기록연구사에게는 귀찮고 부담스러운 업무이기도 하다. 대다수의 기관이 기록관리전문요원 1명을 배치해 운영하는 현실에서 방대한 양의 비공개기록물을 매년 재분류하기에는 인력과 시간이 턱없이 부족하다. 또한 공개재분류 업무는 수집·이관·정리, 평가·폐기 등 다른 업무와 달리 당장 표시가 나는 업무도 아니기에 업무 우선 순위에서도 뒤로 밀리는 경우가 많다. 실제 중앙행정부처 기록연구사를 대상으로 한 설문(송혜진, 김유승, 2012)에서도 공개재분류 업무는 44개의 기록전문직 업무요소 중 중요도 순위 30위에 불과했고, 수행도는 21위로 평균 이상으로 조사됐다. 설문대상을 교육청 기록연구사로 바꾼 다른 연구에서도 공개재분류 업무는 24개 업무영역 중 중요도는 22번째로 낮은 반면, 난이도는 4번째로 높은 업무라고 답변한 조사결과(이을지, 이호신, 2016)도 있다. 기록연구사들은 공개재분류 업무를 중요성은 낮으면서, 수행하기는 까다롭고, 품이 많이 드는 업무로 인식하고 있는 셈이다.

이렇다보니 기관에서 공개재분류 업무는 형식적으로 이뤄지는 경우가 많고, 심지어 아예 공개재분류 업무를 수행하지 않는 기관도 있을 정도다. 그나마 외부 위탁사업을 통해 진행하는 국가기록원, 국회 등 일부 기관의 경우는 형편이 나은 상황이다. 물론 기관의 성격과 보유기록물의 차이, 현실적 어려움 등을 고려해볼 때 일정 부분 이해가 되는 부분이 없진 않지만, 공공기록물 관리에 관한 법률(이하 기록물관리법)의 취지나 기록연구직의 존재 이유 등을 고려할 때 마냥 손을 놓고만 있을 수도 없는 노릇이다.

공개재분류를 다룬 대부분의 선행연구에서 공통적으로 지적한 문제점은 방대한 비공개기록물의 양에 비해 이를 처리할 수 있는 기록연구사가 부족하다는 점이다(신성혜, 2009; 이현자, 2010; 임희연, 2016; 장혜진, 2016). 이에 대한 해결책으로 인력 확충을 비롯한 다양한 법적·제도적 개선방안을 제안하고 있으나, 모두 당위론적이고, 이상적인 제안들이라 당장 적용하기는 어려운 것이 사실이다. 최근 개정된 기록물관리법¹⁾에서는 5년 주기 재분류의 유예가 적용되는 비공개기록물을 명시함으로써 대상 기록물을 획기적으로 줄이는 제도적 보완을 이뤄냈지만, 그럼에도 기록연구사 1인이 해결할 수 있는 양은 아니기에 현장에서의 어려움은 여전하다. 최근 인공지능 활용 지능형 아카이브에 대한 연구(류한조 외, 2021; 방재현, 2018)에서는 기록연구사의 업무 부담을 경감해주는 공개재분류 자동화가 제안되기도 했다. 이런 연구를 바탕으로 한 공개재분류 솔루션도 몇 년 전 이미 개발되어 시판되고 있으나, 아직까지 많이 알려지지 않았다. 경기도, 강릉시 등 극히 일부 지자체에서 활용된 사례가 있었으나 어느 정도 실효성이 있었는지까지는 확인이 어렵다.

이런 상황에서 국가기록원의 ‘2020년 공개재분류 결과 건별 입력 위탁사업’ 수행 간 공개재분류 솔루션을 실제 활용해볼 기회가 생겼고, 그 결과를 사람이 수행한 작업과 비교·분석해보았다. 물론 해당 사업은 공개재분류 솔루션에 대한 분석과 연구가 중심인 연구용역 사업이 아니었고, 영구기록물관리 기관이라는 특성에 맞춰 커스터마이징 되지 않은 상용프로그램을 실험적으로 사용해 본 제한적인 결과라는 한계는 존재한다. 그러다보니 적용 대상도 다양화해볼 수 없었고, 결과분석도 단순 일치도 확인에 그치는 수준이었다. 그럼에도 공개재분

1) 2019. 12. 3. 개정된 공공기록물 관리에 관한 법률 제35조 제1항에 “다만 기록물관리기관이 공공기관의 정보공개에 관한 법률 제9조 제1항 제6호에 해당하여 비공개로 재분류한 기록물에 대해서는 생산연도 종료 후 30년까지 공개여부 재분류를 실시하지 아니할 수 있다.”를 추가했다.

류와 지능형 아카이브에 대한 거시적이고 피상적인 논의만 있는 현 상황에서 기록연구사가 직접 활용가능한 공개재분류 솔루션을 실제 사업에서 적용하고, 이에 대한 활용 결과와 개선점 등을 살펴보는 것은 여러모로 의미가 있으리라 생각된다. 부족하지만 본 논문은 공개재분류 솔루션의 활용 사례에 대한 소개와 결과 분석을 통해 기록연구사의 공개재분류 업무를 경감시킬 수 있는 실질적이고 미시적인 차원의 논의를 시작해보고자 한다.

2. 국가기록원 기록물 공개재분류

2.1 국가기록원 공개재분류 업무 현황과 절차

국가기록원은 국민의 알권리 제고 및 기록서비스 추진 기반 마련을 목표로 매년 공개재분류를 실시하고 있다. 2007년 4월 기록물관리법의 전면 개정에 따라 소장기록물의 공개재분류가 법적 의무사항이 된 까닭이다. 기록물관리법 35조에 따라 비공개로 재분류된 기록물은 매 5년마다 공개여부를 재분류하고 있으며, 생산된 지 30년이 경과한 비공개기록물을 ‘공개원칙’을 적용해 공개재분류를 실시해오고 있다. 또한 동법 시행령 제 72조에 의거 소장기록물은 공개여부를 구분하여 관리하도록 규정하고 있는 바, 미분류 기록물에 대한 공개재분류 역시 필수적으로 이뤄진다. 2007년부터 2020년까지 13년간 국가기록원은 약 310만 권, 8,607만 건의 공개재분류를 완료했다. 5년 주기 재분류에 따른 중복건수가 포함되어 있긴 하지만 매년 방대한 양의 공개재분류가 이뤄졌다. 다만 이러한 양도 법정 대상량에 비하면 턱없이 부족한 수치이긴 하다. 2021년 사업대상은 145,847권(2449,33건)이며 2020년과 2021년 추진한 공개재분류 분량은 법정 대상량의 각각 7.9%, 10.3%에 불과하다. 그간 추진된 국가기록원의 기록물 공개재분류 세부 내역을 살펴보면 <표 1>과 같다(국가기록원, 2021).

초기에는 ‘비공개기록물 재분류 3개년 기본계획’에 의거 국가기록원 자체적으로 공개재분류를 진행하였고, 2010년부터 2021년 현재까지는 5억 5천만 원 안팎의 예산으로 외부업체에 위탁·진행하고 있다. 제안 요청서 등을 통해 확인한 사업 위탁의 이유로는 입력 작업량이 방대하고, 시스템 반영을 위한 파일변환 기술 등이 필요하다는 것을 꼽고 있다. 아울러 범정부적으로 추진 중인 일자리 창출 및 경제 활성화에 기여한다는 점도 이유 중 하나이다.

2010년부터 2021년까지 조달청 나라장터(www.g2b.go.kr)에 올라온 공개재분류 사업 입찰공고문을 분석해보니 공개재분류를 위탁사업 형태로 진행한 기관은 국가기록원 외에도, 대통령기록관(2009-2017), 서울시(2015-2017), 강원도(2015), 양산시(2015-2016), 충청남도(2016-2019), 인천시 동구(2016), 새마을운동중앙회(2020) 등이 있었다. 2021년 10월 현재는 국회(2012-2019, 2021)와 서울시교육청(2012-2015, 2021), 경상남도기록원(2020-2021) 등 세 곳을 확인할 수 있었다. 대통령기록관은 2017년을 끝으로 자체사업을 진행하고 있으며, 국회는 2020년에 공개재분류가 아닌 ‘분류점검사업’을 발주하였다. 서울시교육청은 공개재분류 5개년 계획 중 4년차까지 이뤄진 후 사업을 중단했으나, 2021년 재개되었다. 경남기록원은 2020년 공공데이터 기업매칭 지원사업으로 공개재분류 사업을 시작한 이래 2021년부터 위탁사업을 진행 중이다. 이밖에도 ‘중요기록물 전자화(DB구축) 및 공개재분류 사업’으로 디지털화와 공개재분류를 함께 진행한 다수의 기관들도 있으나, 이는 전자화 대상 종이기록물의 공개값을 단순 책정하는 경우가 대부분이라 공개재분류라고 보긴 어려운 형태다(임희연, 2016).

공개재분류사업을 추진한 대부분의 기관은 기록물 공개여부 검토서, 기록물유형별 검토서의 작성부터 입력, 업로드 등 공개재분류 업무 전반을 업체에 위탁하지만, 국가기록원은 기관에서 작성한 공개재분류 기준서를 바탕으로 건별 결과값 입력과 업로드만을 위탁한다. 때문에 국가기록원은 공개재분류 사업이 아닌 공개재분류 결과 ‘건별 입력 위탁사업’이라는 명칭을 사용하는 것으로 보인다.

<표 1> 국가기록원 공개재분류 추진실적

구분	합계(건)	공개			비공개
		소계	전부공개	부분공개	
계	86,074,468	57,111,854	19,062,660	38,049,194	28,962,514
	100%	66.4%	22.1%	44.2%	33.6%
'07년~'09년	38,814,631	24,548,617	15,614,526	8,934,091	14,266,014
	100%	63.2%	63.6%	36.4%	36.8%
'10년	2,707,843	1,298,854	169,928	1,128,926	1,408,989
	100%	48.0%	13.1%	86.9%	52.0%
'11년	1,231,191	971,319	339,063	632,256	259,872
	100%	78.9%	34.9%	65.1%	21.1%
'12년	3,914,276	3,121,668	206,720	2,914,948	792,608
	100%	79.8%	6.6%	93.4%	20.2%
'13년	17,223,708	11,719,823	245,812	11,474,011	5,503,885
	100%	68.0%	2.1%	97.9%	32.0%
'14년	4,046,982	3,360,998	403,642	2,957,356	685,984
	100%	83.0%	12.0%	88.0%	17.0%
'15년	6,891,460	6,261,102	269,599	5,991,503	630,358
	100%	90.9%	4.3%	95.7%	9.1%
'16년	2,841,253	718,664	170,910	547,754	2,122,589
	100%	25.3%	23.8%	76.2%	74.7%
'17년	1,345,903	1,113,612	134,587	979,025	232,291
	100%	82.7%	12.1%	87.9%	17.3%
'18년	2,241,681	1,182,572	332,086	850,486	1,059,109
	100%	52.8%	28.1%	71.9%	47.2%
'19년	2,309,025	1,549,110	590,479	958,631	759,915
	100%	67.1%	38.1%	61.9%	32.9%
'20년	2,506,515	1,265,515	585,308	680,207	1,241,100
	100%	50.5%	46.3%	53.7%	49.5%

※ 국가기록원, 『21년 비공개기록물 공개재분류 추진계획』에서 발췌 및 가공

국가기록원의 업무 중 공개재분류 절차와 그중 위탁사업이 이뤄지는 부분은 <그림 1>과 같다. 단계별로 살펴보면, 먼저 공개재분류 대상을 선정하는데, 중앙연구기록물관리시스템(Central Archives Management System, 이하 'CAMS')에서 재분류 대상기록물 목록(철, 건)을 추출한다. CAMS에는 생산 당시 기관명으로 등록되어 있어, 이를 현재 기관명 및 담당 기관 등을 고려해 생산기관별, 기록물 유형별로 분류한다. 이렇게 분류된 기록물은 유형별로 기록물 원본을 실사하고, 기존 심의이력 등을 검토하여 '기록물 공개재분류 기준서'(이하 기준서)를 작성한다.

유형별로 작성된 기준서는 생산이력, 기록물 세부내용, 공개이력, 공개여부, 비공개사유, 비공개 대상정보 등의 항목으로 이루어져 있다. 기록물 공개의 일관성과 편의성을 위해 작성된 기준서를 바탕으로 위탁업체는 건건이 공개값 및 비공개 대상 정보 등을 확인해 건 별로 결과 값을 입력하게 된다. 이런 작업이 <그림 1>에서는 생산기관 의견조회와 심의회 및 위원회 심의 다음으로 위치해 있지만, 240일간의 사업기간을 고려하여 해당 과정들은 중첩되어 진행된다. 일단 기준서대로 입력하고 생산기관 의견조회나 기록물공개심의회 및 국가기록관리위원회 심의 과정에서 변경되는 결과 값이 있을 경우 이는 추후에 반영하게 된다. 이렇게 수정까지 마친 최종 결과값 등은 워드문 형태로 작성되어 CAMS에 일괄 업로드 된다. 최종 공개재분류 현황과 공개기록물 목록은 통상 다음해 2월 경 국가기록원 홈페이지의 공개재분류 기록물 서비스 메뉴(<https://theme.archives.go.kr/next/closed/page/viewMain.do>)와 관보에 게시된다(권미현, 2019; 국가기록원, 2020c).



〈그림 1〉 국가기록원의 비공개기록물 공개재분류 절차
 ※ 20년 비공개기록물 공개재분류 결과 권별 입력 위탁사업 착수보고자료

2.2 비공개기록물 공개재분류 결과 권별 입력 위탁사업

국가기록원은 ‘비공개기록물 공개재분류 결과 권별 입력 위탁사업’이라는 명칭으로 10년 이상 공개재분류 사업을 진행해왔다. 공개재분류 대상이 되는 기록물은 시기별로 도래하고, 분산 배치되어있기 때문에, 작업은 국가기록원 본원(정부 대전청사 내), 나라기록관(성남), 역사기록관(부산), 행정기록관(대전)에서 4곳에서 각각 이뤄진다. 사업의 주요 과업은 ‘사전작업, 기록물 반출 및 반입, 공개재분류 결과 권별입력, 기록물 정리 및 검수, CAMS 업로드 규격 준수하여 업로드용 데이터 작성’ 등으로 역시 거의 변화가 없었다.

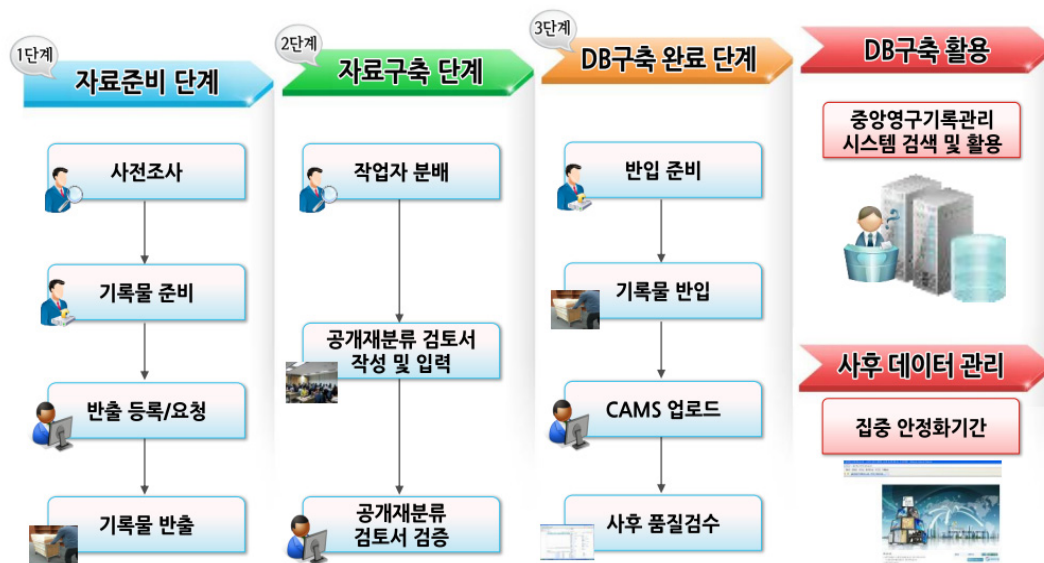
권별 입력 항목 16개는 <표 2>와 같다. 대부분 날짜, 심의회 여부 등과 같이 명확한 사항들로, 입력 항목 대부분을 기관에서 받아서 정확하게 입력하는 게 사업의 관건이다. 사업을 수행하는 위탁업체는 철 단위 기준서를 건단위로 적용하는 부분이나, 세부 비공개대상 정보를 찾아내는 부분에서 판단의 여지가 일부 있으나 주어진 기준서에 따라 단순 입력한다는 점에서 다른 공개재분류 사업에 비해 업무난이도는 높지 않은 편이라 할 수 있다. 공개재분류의 핵심인 기준서를 기관이 직접 작성하기 때문이다.

대상 기록물을 반출해 기준서에 따라 입력하고, 다시 반입하는 <그림 2>와 같은 위탁사업의 절차는 지난 10여년 간 큰 변동없이 이뤄졌다. 물론 운용 방식의 변화는 국가기록원의 조직 개편과 정책 등에 따라 여러 차례 있었다. 그간의 사업을 분석해보면 2014년 이전 초기에는 대전 본원에서 사업을 총괄하는 ‘중앙 집중식’으로 운영되었다. 담당 주무관이 성남, 대전, 부산을 이동하며 기준서를 작성하다보니 업무 과중 및 지연이 발생하였다. 2015년에는 권역별 통합기록관리 체계로 조직이 개편됨에 따라 ‘소장 기록관별’로 각각의 주무관이 기준서를 작성하였다. 이에 따라 업무 분담은 되었으나, 소장된 기록물을 대상으로 기록관별로 기준서를 작성하다보니 동일 생산기관, 동일 유형의 문서일지라도 작성자에 따라 기준서가 다른 경우가 발생하기도 했다. 2016년부터 2018년까지는 ‘소관 기록별’로 기준서를 작성하였다. 즉, 나라기록관에서는 서울, 경기, 강원, 세종시 소재 기관이 생산한 기록물에 대한 기준서를 작성하고, 역사기록관은 영남지역 소재 기관, 대전 본원 및 행정기록관은 호남, 충청, 제주지역 소재 기관이 생산한 기록물에 대해 각각 기준서를 작성하였다. 2019년부터 다시 ‘중앙집중식 수행’ 및 주관부서가 대전 본원에서 나라기록관으로 변경되었으며, 처음으로 기관별 기준서가 아닌 공통 업무 기준서가 도입되었다.

〈표 2〉 공개재분류 입력항목

입력항목	내용 및 예시
재분류일자	예) 2020-12-28
공개구분	공개, 부분공개, 비공개
비공개유형	「공공기관의 정보공개에 관한 법률」 제9조 제1항 제1호~제8호 중 해당 항목 입력
비공개대상정보	부분공개일 경우 비공개대상정보(이름, 주소, 주민등록번호 등)를 구체적으로 입력
검토의견	작업지시서·공개재분류검토서의 검토의견을 반영하여 입력
생산기관 의견조회 여부	Y or N 입력
생산기관 의견조회 회신 일자	예) 2020-12-28
생산기관 의견 회신문서번호	환경부 운영지원과-4867호
기록물공개심의회 여부	Y or N 입력
기록물공개심의회 일자	예) 2020-12-28
국가기록관리위원회 여부	Y or N 입력
국가기록관리위원회 일자	예) 2020-12-28
서비스용 첩 제목 서비스 여부	Y or N 입력
서비스용 첩 제목	개인정보 등은 ○○○ 처리
서비스용 건 제목 서비스 여부	Y or N 입력
서비스용 건 제목	개인정보 등은 ○○○ 처리

※ 비공개기록물 공개재분류 결과 건별입력 위탁사업 제안요청서



〈그림 2〉 국가기록원의 비공개기록물 공개재분류 건별입력 위탁사업 절차
 ※ 20년 비공개기록물 공개재분류 결과 건별 입력 위탁사업 착수보고자료

이런 작은 변화와 달리 2020년 사업에서 전자기록물이 최초로 추가된 큰 변화가 있었다. 2019년까지의 국가 기록원 공개재분류는 각급 기관에서 이관되어 온 비전자기록물만을 대상으로 했다. 그러나 2020년 사업은 2004년, 2005년 생산되어 2014년, 2015년 국가기록원으로 이관된 전자기록물의 공개재분류 시기가 도래해 이에 대한 공개재분류를 해야만 했다. 처음 시작이라 양이 많지는 않았지만 공개재분류 대상이 되는 전자기록물이 기하급수적으로 증가하는 앞으로의 추세를 생각하면 전자기록물에 대한 공개재분류의 첫 단추를 잘 끼워야 하는 상황이었다.

<표 3> 2020년 비공개 기록물공개재분류 결과 건별입력 위탁사업 대상량

경과기간	구분		수량	
	위치	권	건	
30년 경과 (2,507권/267,205건)	나라기록관	13,435		166,361
	대전 본원	3,385		26,626
	행정기록관	6,499		46,194
	역사기록관	1,888		28,024
30년 미경과 (128,502권, 2,161,193건)	나라기록관	63,946		853,710
	전자기록물	24,268		430,082
	대전 본원	2,336		26,686
	행정기록관	20,761		134,950
	역사기록관	17,191		715,765
합계		153,709		2,428,398

※ 제안요청서 상의 최초 목록에서 중복제거 등을 통해 산출한 착수보고 당시 대상량임.

2020년 사업 대상량은 <표 3>과 같이 153,709권, 2,428,398건으로 이중 전자기록물이 2.4만권 포함되어 있었다. 전자기록물의 경우 공개 여부와 상관없이 전체 대상(5년 주기)이 선정됐으며, 이는 총 사업 대상량의 약 15%에 해당하는 수치였다. 당시 제안요청서는 전자기록물 공개재분류에 있어 ‘내용식별 솔루션을 사업수행에 활용’할 것을 요구했다. 특정 솔루션을 지정한 것은 아니고 시판되고 있는 상용 소프트웨어 중에 ‘전자문서 내 개인정보 추출(개인정보보호법에 의한 개인정보 내역, 파일 사이즈, 특정 키워드 추출), 특정 키워드를 중심으로 양쪽 문장 추출, hwp, doc, excel, zip파일 내 전자문서 등 다양한 파일 지원, 추출 결과를 excel 또는 특정한 양식으로 정리’를 주요 기능으로 하는 전자기록내용식별 솔루션이라면 어떤 솔루션도 관계는 없었다. 이에 위탁 사업자는 2015년 개발되어 판매 중인 S사의 공개재분류 솔루션을 선정해 사업을 진행하였다(비공개 기록물 공개재분류 결과 건별 입력 위탁사업 제안요청서, 2016~2021, 2020년 비공개 기록물 공개재분류 추진계획, 사업제안서 등).

3. 공개재분류 솔루션을 활용한 공개재분류 결과 분석

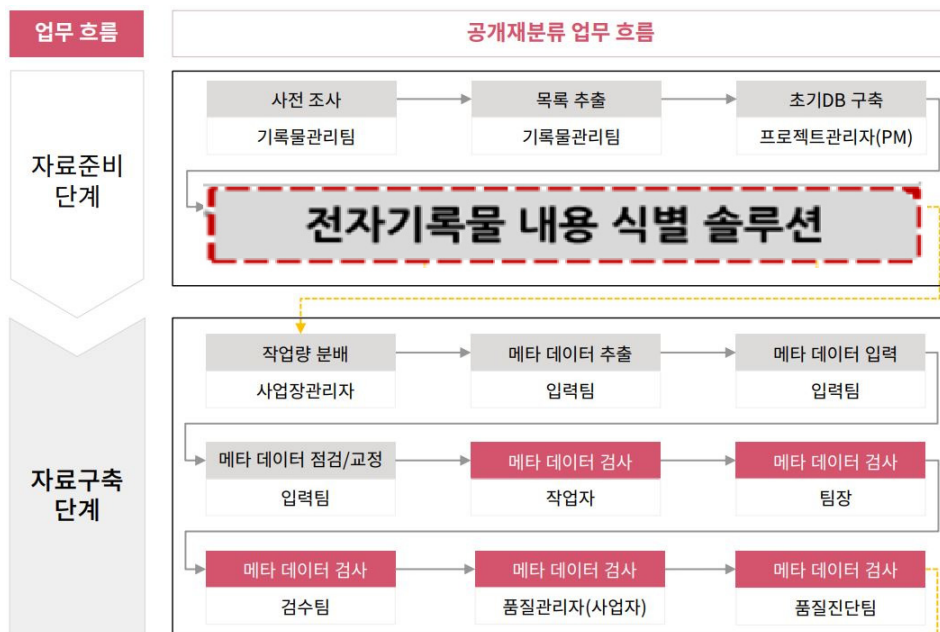
3.1 공개재분류 솔루션 개요 및 적용

지능형 업무처리는 대량 반복 업무 처리에 강점을 가진다고 평가받는데, 공개재분류는 기록관리 업무 중에서 이에 해당되는 분야로 첫 손에 꼽힌다. 인공지능을 기반으로 한 지능형 업무처리를 공개재분류 업무에 적용한다면 자연어를 처리하여 유사문서를 구분하고 동일한 기준을 적용할 수 있게 된다. 즉 지능형 문서처리를 기반으로 한 공개 재분류가 실행된다면 업무수행의 품질, 일관성, 속도 등을 개선하여 생산성을 높일 수 있을 것으로 기대(류한조 외, 2021)한 것이다. 공개재분류 사업에서 기관이 요구했던 솔루션의 기능은 단순 개인정보 추출이었으나, 그 기능과 함께 공개값(공개, 부분공개, 비공개), 비공개·부분공개의 사유(공공기관의 정보공개에 관한 법률 제9조 1항 1호~8호)를 추천하는 기능까지 포함된 솔루션을 선정해 활용하였다.

유사한 기능을 지닌 몇몇 공개재분류 솔루션²⁾이 있었으나, 그중에서 가장 먼저 개발되었고, 개발자의 개발이력을 확인할 수 있는 S사의 A프로그램을 선정하였다. 이 제품은 ‘기계학습 기반 지능형 기록물 공개 관리시스템’이

2) 본 사업에서 활용한 S사의 솔루션 외에도 또 다른 S사가 개발한 A솔루션, D사의 A솔루션 등 공개재분류 솔루션이 출시되어 있다.

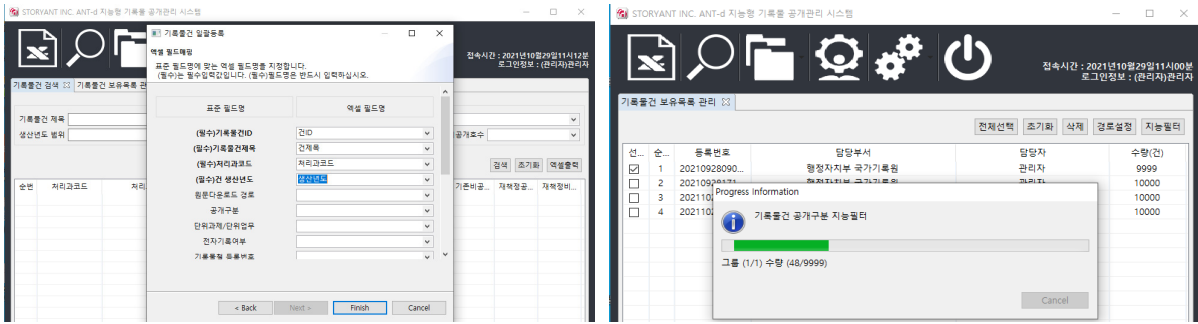
라는 명칭으로 특허가 등록(특허 제10-1627550호)되어 있었고, 학습데이터가 공개되어 있어 신뢰성을 확보했다고 판단했다. 해당 솔루션이 학습한 데이터는 2015년 4월 30일부터 2016년 5월 30일까지 13개월 동안 검색사이트를 통해 확보한 전국 지방자치단체 공개재분류결과 목록 4,475,202건과 정보공개포털을 통해 수집된 31,691건 등 총 4,506,893건의 기록물 건 정보(제목, 공개여부 항목)로, 이를 3개의 클래스(공개/부분공개/비공개)로 구분해 학습모델을 생성(방재현, 2018)하였다고 한다.³⁾ 개인정보의 추출의 경우 행정안전부 개인정보 보호과에서 2012년 펴낸 『홈페이지 개인정보 노출방지 가이드라인』에서 예시하고 있는 주민등록번호, 여권번호, 운전면허번호, 휴대폰번호, 외국인등록번호, 신용카드번호, 건강보험번호, 계좌번호 등 8개의 패턴화된 정보를 검색·필터링하는 기능을 보유하고 있었다.



<그림 3> 전자기록물 공개재분류 절차도
 ※ 20년 비공개기록물 공개재분류 결과 건별 입력 위탁사업 착수보고자료

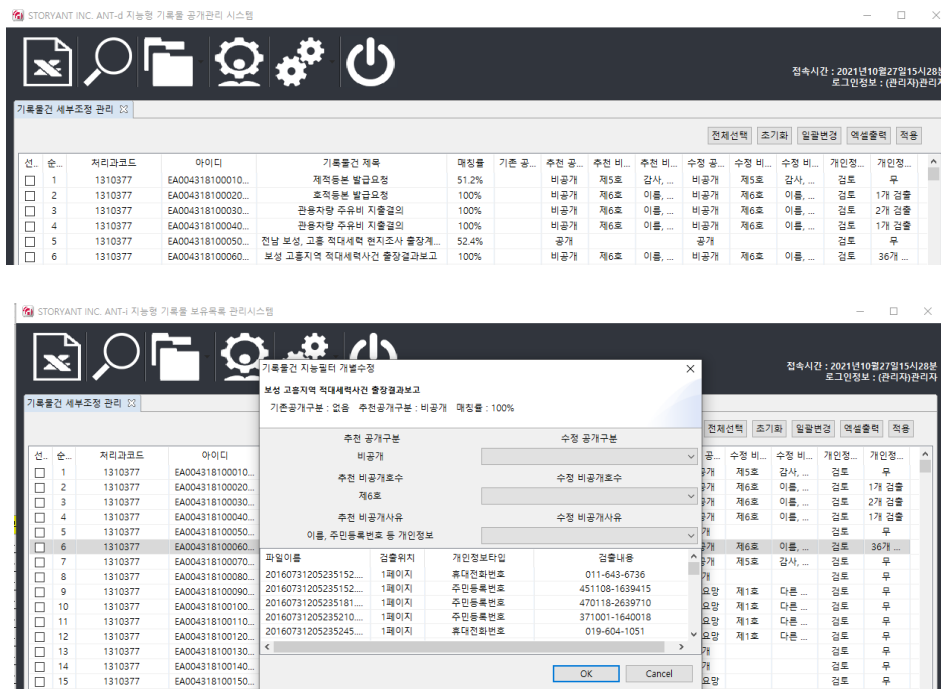
이런 솔루션을 활용한 전자기록물 공개재분류는 <그림 3>의 흐름으로 이뤄졌다. 앞선 <그림 2> 위탁사업 절차의 자료 준비단계와 자료 구축단계를 비교할 때 기록물 반입·반출 과정이 빠지고 그 대신 전자기록물 내용식별 솔루션이 추가되었다. 전자기록물 공개재분류의 경우 시스템 상에서 직접 수행하기도 하지만 기록물 뷰어의 로딩 속도, 다수의 작업자 접속에 따른 부하 등을 고려해 전자기록물 원문을 다운로드받아 작업하기로 했다. 원문 다운과 목록 추출은 CAMS 유지보수업체인 C사에서 담당했으며, 이렇게 추출된 공개재분류 대상 전자기록물의 목록과 원문은 외장하드에 제공받았다. 이를 솔루션이 설치된 관리자 PC에 옮겼으며, 솔루션이 요구하는 방식으로 전자기록물 목록을 수정하였다. 처리과 코드를 국가기록원으로 통일하고, CAMS의 관리번호와 BSID, DSID⁴⁾를 조합해 각각의 건 ID를 생성하였다. 또한 CAMS에서 원문이 추출된 폴더 구조에 맞춰 건에 대한 이미지 경로를 지정해주었다.

3) 솔루션에 사용된 기술의 이론적 배경이나 개념은 본 논문의 주제 범위를 넘어서는 부분으로, 솔루션 개발자이기도 한 방재현(2018)의 논문에서 자세히 다루고 있다.
 4) CAMS에서 기록물이 등록되고 관리되는 번호로 관리번호는 철단위로 부여되는 총 9자리 숫자로 알파벳 2자리+숫자 7자리로 구성된다. BSID는 철에, DSID는 건에 각각 부여되는 12자리 숫자이다.



<그림 4> 공개재분류 솔루션에 등록 및 구동

이렇게 수정 작업을 거친 목록을 <그림 4>의 좌측처럼 견제목, 처리과코드, 생산년도 등을 입력해 솔루션에 등록하고, 지능필터를 클릭하면 <그림 4> 우측과 같이 구동된다. 구동시 프로세싱 인포메이션 팝업창이 뜨면서 그룹, 수량의 진행 정도를 표현해준다. 기록물 건의 페이지 수나 OCR 변환으로 인한 원문의 차이 등으로 개인정보 탐색 시간이 차이를 나타냈으며, 결과 값은 <그림 5> 상단의 형태로 구현됐다. 순번, 처리과코드, 아이디, 기록물건 제목, 매칭률, 기존 공개구분, 추천 공개구분, 추천 비공개호수, 추천 비공개사유, 수정 공개구분, 수정 비공개호수, 개인정보검토, 개인정보 유무 등 총 13개의 컬럼이 표현되었고, 세부 건을 클릭하면 <그림 5> 하단처럼 팝업이 떠 파일이름, 검출된 위치, 개인정보 타입, 검출내용을 세부적으로 표시해준다. 이렇게 생성된 결과값은 표시된 형태와 동일하게 엑셀로 추출할 수 있다.



<그림 5> 공개재분류 솔루션 구동 결과 값 표현 화면

솔루션은 2020년 6월 29일 국가기록원 나라기록관 공개재분류 사업장 관리자 PC 1대에만 설치했으며, 1만 건 단위로 끊어 일과시간에만 구동하였다. 대상량은 모두 30년 미경과 기록물로 2014, 2015년 이관된 403,657건,

비상대비 6,011건 등 총 409,668건이었으나, 이중 중복파일, 원문없음 등을 제외한 24,237건 407,894건이 최종 대상이었다. 모든 대상을 솔루션에 구동 완료한 시점은 시작한지 3개월이 조금 안되는 9월 24일이었다.

3.2 공개재분류 솔루션 활용 결과 비교 분석

관리자가 구동했던 솔루션과는 별개로 작업자들은 비전자기록물과 동일하게 기준서를 바탕으로 공개값을 입력하고 비공개 대상 정보를 찾는 ‘전자기록물 공개재분류’ 작업을 실시하였다. 실물 기록물을 보고 공정프로그램에 입력하는 비전자기록물과 방식은 동일하였으나, 전자기록물을 모니터를 통해 봐야하기 때문에 듀얼 모니터를 설치했다. 한 모니터에서는 전자기록물 원문을 보고, 다른 모니터에서는 공정프로그램을 띄워 공개재분류 결과 및 비공개대상정보를 입력하는 형태였다. 실물 기록물을 서고에서 반출시키는 작업이 없는 대신 전자기록물의 각각의 철, 건, 첨부문서 폴더들을 작업자가 일일이 열어서 확인해야 했다. 건의 페이지가 많은 경우 로딩 시간도 소요됐고, 모니터를 보는 계속 봐 눈의 피로가 심해진 작업자의 휴식시간도 필요했다. 이렇게 작업된 전자기록물 공개재분류 입력된 현황은 <표 4>와 같다.

<표 4> 전자기록물 공개재분류 입력현황

입력 현황		총 투입 인력(명)	일 평균입력/ 1인		비고
철	건		철	건	
24,237	429,930	375	64.6	1,146.5	비공개 포함
19,891	288,350		53.0	768.9	비공개 미포함

기준서에서 비공개로 주어진 기록물에 대해서는 실제 작업을 하지 않았지만 이를 포함해 입력현황을 분석할 경우 1인당 하루 평균 64.6철, 1,146.5건을 입력했다. 비공개를 포함시키지 않은 실제 작업량은 53철, 768.9건의 입력실적을 기록했다. 물론 대상 기록물의 유형이나 작업환경, 작업자의 숙련도 등 다양한 변수가 있지만 약 2만건의 기록물을 공개재분류하는데 3명이 5개월은 걸린다는 이야기이다. 기준서가 있는 상태임에도 그러니 작업 속도에서만큼은 공개재분류 솔루션은 사람과의 비교를 불허한다 하겠다. 솔루션은 사람과 달리 24시간도 가동은 물론 PC 여러 대의 활용도 가능한만큼 적어도 속도에 있어서만큼은 사람은 절대 따라올 수 없는 뛰어난 효율성을 보인다.

문제는 정확도라 할 수 있는데, 이에 대한 정확한 확인은 어려웠다. 다만 이를 간접적으로 확인하기 위해 작업자에 의해 입력된 공개재분류 결과 값을 솔루션의 추천 결과 값과 비교해보았다. 사람이 한 작업도 완벽하다고 볼 수는 없지만 검수를 여러 번 거치기 때문에 이를 기준으로 솔루션 결과의 경향은 확인할 수 있으리라 생각했다. 솔루션을 적용시킨 407,893건의 기록물 중 솔루션은 6,296건에 대해 ‘확인요망’이라는 결과 값을 내놓았는데, 대부분 파일이 깨지거나 오류가 있는 경우 등이었다. 이러한 전자기록물의 단순오류 조차도 사람은 일일이 파일을 열어보고 확인해야 하지만 솔루션은 즉각적인 확인이 가능했다. 작업 전을 솔루션을 통해 이렇게 문제있는 파일만 추려서 확인해도 사람의 업무는 많이 경감될 수 있을 것이다.

이런 오류가 발생한 기록물을 제외하고 남은 401,597건의 기록물을 대상으로 사람이 한 작업과 솔루션이 추천한 결과 값의 일치도를 분석한 결과는 <표 5>와 같다. 공개, 부분공개, 비공개를 판단하는 공개여부는 솔루션과 사람의 일치율은 48.61%였다. 일치도가 절반도 채 되지 않는 이유는 솔루션의 경우 개인정보가 검출된 기록물 11,932건을 무조건 비공개로 추천하였고, 사람은 개인정보가 있다하더라도 이를 경우에 따라 공개(1,353건)하거나, 부분공개(10,579건)로 재분류하는 경우가 많았기 때문이다. 특정 정보만을 제외하고 공개하는 부분공개의 판단은 가치판단이 이뤄져야 해 작업자들도 어려워하는 만큼 현재의 솔루션 수준에서 당장 해결하기는 힘든 부분으로 보여졌다.

<표 5> 솔루션 추천 값과 사람 작업 결과의 일치도

구분	공개여부		추천호수(단수)		추천호수(복수)	
	건	비율	건	비율	건	비율
일치	195,233	48.61%	2,639	3.54%	21,004	28.20%
불일치	206,364	51.39%	71,843	94.46%	53,478	71.80%
합계	401,597	100%	74,482	100%	74,482	100%

또 솔루션을 학습시킨 데이터가 일관된 기준이 적용된 정제된 데이터가 아니라는 한계도 있었다. 지방자치단체 기록관을 타겟으로 한 상용 프로그램이다 보니 영구기록물관리기관 보다 기록관 단계의 지방자치단체의 공개재분류 결과값이 학습 데이터로 많이 포함되어 보수적으로 공개값을 추천하는 경향을 보였다.

공개를 제외한 부분공개와 비공개에의 경우 공공기관의 정보공개에 관한 법률 제9조 제1항의 제1호부터 제8호까지의 비공개 사유를 넣도록 되어있다. 솔루션과 작업자의 공개값이 일치하는 195,233건 중 호수가 필요한 74,482건의 호수를 비교해보니 단 3.54%만 일치하였다. 이는 사람들의 작업에서는 여러 호수를 함께 넣은 복수 호수가 많았으나, 솔루션은 모두 단수 호수만을 추천했기 때문이다. 이는 솔루션 설계에서 공공기록물 관리에 관한 법률 시행규칙 별지 제13호의 2 서식에 따라 ‘동일 건에 비공개 사유가 2개 이상인 경우에는 주요 비공개 사유를 표기’하는 방식을 적용했기 때문으로 보인다. 사람이 복수로 호수를 선정한 기록물과 솔루션의 단수 추천 호수와의 부분 일치하는 건을 확인해보면 21,004건으로 일치도가 28.20%로 다소 상승하는 결과를 확인할 수 있는 만큼 설계상 복수 호수를 적용하면 개선될 수 있는 부분으로 판단된다.

기관에서 요구한 솔루션의 가장 중요한 기능은 개인정보 추출이었다. 선정된 솔루션은 주민등록번호 등 숫자로 정형화된 개인정보 8가지만을 검출할 수 있는데, <표 6>에서 보듯 ‘원문 없음’으로 검토를 하지 못한 381건을 제외하고, 대상의 5.6%인 22,861건에서 개인정보를 검출하였다. 개인정보가 검출된 22,861건 중 작업자가 기록물을 확인한 공개, 부분공개 기록물은 11,932건이었다. 이중 작업자가 개인정보를 비롯한 비공개 대상정보를 찾아낸 기록물은 10,579건, 못찾아낸 것은 1,353건이었다. 이는 솔루션과 작업자간의 일치도가 88.66%에 달하는 높은 수치이다. 불일치하는 1,353건의 경우 개인정보가 아닌 관리번호 등과 같은 형태가 주민등록번호와 유사한 일련번호인 경우도 있었고, 솔루션은 발견했으나 작업자가 실수로 미처 발견하지 못한 개인정보도 다수 있었다. 작업자는 1,090,223건에서 비공개대상 정보를 찾아냈고, 솔루션은 이중 90% 이상의 비공개 대상정보는 검출하지 못하였다(<표 7> 참조).

<표 6> 솔루션 추출 개인정보

구분	건	비율
개인정보 검출	22,861	5.60%
개인정보 미검출	378,355	94.30%
미검토(원문없음 등)	381	0.09%
합계	401,597	100%

<표 7> 개인정보 검출의 일치도

구분	개인정보(솔루션 검출)		비공개대상 정보(작업자 검출)	
	건	비율	건	비율
일치	10,579	88.66%	10,579	9.70%
불일치	1,353	11.34%	98,444	90.30%
합계	11,932	100%	109,023	100%

작업자는 개인정보 외에도 다양한 비공개 대상정보를 찾아내지만 솔루션은 개인정보만을, 그중에서도 정형화된 것만을 찾을 수 있다보니 나온 당연한 결과이다. 이는 향후 이름, 주소 등 비정형 개인정보는 검색키워드 지정 등의 기능을 솔루션에 추가하는 형태로 보완할 수 있을 것이다.

이밖에도 솔루션 활용의 효과를 제대로 보지 못한 데에는 PDF파일에서의 OCR 인식문제, 인터넷접속이 제한된 환경에서의 클라우드 서버의 미활용, 낮은 PC 사양으로 인한 간이버전 솔루션, 검색 키워드 미지정 등 다양한 이유를 꼽을 수 있었다. 이는 오류 수정 후 재작업 등을 커스터마이징되지 않은 상용프로그램에서는 할 수 없었기 때문에 실제 솔루션을 도입할 경우 개발자에게 보완을 요구할 수 있는 부분이다.

3.3 공개재분류 솔루션 활용의 보완점 및 시사점

공개재분류 솔루션의 활용 결과를 사람이 작업한 내용과 비교·분석해보니 정확도가 그리 높다고 할 순 없었다. 하지만 단순히 보여주는 수치만으로 이를 폄하해버리기에 활용 가능성이 많이 엿보였고, 보완할 수 있는 지점도 쉽게 확인할 수 있었다. 앞서 언급했듯 본 솔루션이 주로 학습한 데이터가 영구기록물관리기관이 아닌 지방자치단체 기록관의 공개재분류 결과다 보니 기준 차이에 따른 값의 왜곡이 나타났다. 이는 기관별로 정제된 데이터를 학습시키면 쉽게 해결이 될 것으로 보였다. 현재도 기존의 기준서와 심의이력 등을 검토해 유사하게 진행하는 만큼 누적된 국가기록원의 8,607만 건의 공개재분류 결과를 학습시켜 공개 값과 호수를 추천해준다면 업무의 부담을 감소시키는 것은 물론 일관성과 정확성 확보에도 많은 도움이 될 것 같다. 부분 공개 확대, 복수 호수 기능 추가 등과 기관 고유 데이터의 학습, 8가지 개인정보 외에 추가 키워드를 설정해 비공개 대상정보를 추출하는 등 분석과정에서 언급한 내용을 보완한다면 정확도가 비약적으로 향상될 것이다.

현 수준에서는 학습데이터와 분석데이터의 정합도를 표현한 매칭율을 활용해 매칭율 8~90% 이상의 단순한 기록물은 솔루션이 추천한 그대로 재분류하고, ‘확인요망’이나 낮은 매칭율의 기록물, 가치판단이 필요한 복잡한 기록물만 추려 별도로 공개재분류 작업을 한다면 서두에서 이야기한 공개재분류 대상의 방대함의 문제점도 어느 정도 극복할 수 있을 것이다. 또 딥러닝 솔루션의 특성상 정제된 데이터가 많아지면 정확도가 높아지는 만큼 결과값의 피드백을 통해 계속 업그레이드해 나간다면 별도의 작업은 점점 줄어들 것이다.

분석 대상 사업에서는 솔루션 활용 목적 자체가 상용 프로그램을 테스트 형식으로 사용해보는 것이었기에 처음부터 큰 기대를 하지는 않았다. 사업 PM 입장에서는 지금까지의 사업과 달리 솔루션 구입비용과 사용 공수가 추가될 뿐인 과업이기 때문에 더욱 부정적이었다. 하지만 솔루션을 사업간 활용해보면서 기계학습 기반의 지능형 아카이브 구현 가능성을 훨씬 더 높게 보는 계기됐다. 현재 개발된 공개재분류 솔루션 자체가 애초에 완벽하게 공개재분류를 수행하기 위한 것이 아닌 기록연구사의 업무를 효율적으로 수행하고 경감할 수 있도록 설계된 것 인만큼 보조적인 수단으로는 충분히 역할을 할 수 있으리라 판단되었다.

하지만 마냥 장밋빛 미래가 펼쳐질 거 같진 않다. 이미 판매된 지 5년 이상 된 솔루션이지만 많이 알려지지도 않았고 사용하는 기관은 거의 없다. 그러다보니 솔루션의 업데이트나 추가 데이터 학습, 미비점 보완 등도 활발하게 이루지지 않고 있다. 또한 기계학습 기반의 솔루션이다보니 학습데이터 확보가 무엇보다 중요한데 대부분의 기관이 공개재분류 후에도 공개 목록 정도만을 게시하는 현실에서 학습 데이터를 확보하는 것은 거의 불가능해 보인다. 공공데이터 개방 등의 움직임이 없는 것은 아니지만 애초 공개재분류는 비공개기록물을 대상으로 하다보니 기관이 전향적으로 변하지 않은 한 요원한 일이다. 실제 2020년 5월 13일 온라인 화상회의로 진행됐던 공개재분류 위탁사업 착수보고회에서도 대통령기록관의 한 담당자는 솔루션 활용에 대해 “민간기업의 솔루션을 세금 써서 테스트해주는 거 아니냐”고 지적하기도 했다. 그 자리에선 “우리나라의 기록관리 발전을 위해 누군가, 어디 선가는 해야 할 일이고, 그게 국가기록원이 되어야 하지 않겠느냐”고 답변을 했지만 현재 기록연구사의 인식과 현실은 단적으로 보여주는 사례이다.

4. 나오며

기록물 공개재분류는 그 의미와 중요성에 비해 평가절하 되어 있다. 수고에 비해 태가 나지 않고, 활용성도 떨어지다 보니 인정받지 못하는 부분이 없진 않다. 그럼에도 끊임없이 재분류를 하고, 또 해 기록물 공개비용을 과격적으로 높여야 한다고 생각한다. 분명 이렇게 생각하는 기록연구사들이 많이 있을 터인데 공개재분류 업무의 압박감은 그들을 소극적으로 만든다. 이러한 현실을 타계하는 방법은 생산 단계에서부터 비공개기록물을 줄이는 문화를 만들어가고 공개재분류에 대한 법적, 제도적 개선을 추구해야 할 것이다. 이런 장기적으로 진행되어야 하는 당위적이고 피상적인 해결책 외에 그나마 단기적이고 현실적인 방법 중 하나는 지능형 아카이브 솔루션을 공개재분류에 활용하는 것인 듯하다. 활용하기에 따라 기록연구사의 업무를 줄여줄 수 있는 여지가 있기 때문이다. 다행스럽게 이미 몇몇 솔루션이 나와있고 이에 대한 관심도 조금씩 커지고 있는 상황이다. 실제 솔루션을 이용해 공개재분류를 수행한 소수의 기관도 있는 것으로 알려졌지만 그 결과를 확인할 길은 없었다.

때마침 2020년부터 국가기록원 공개재분류 대상에 전자기록물이 포함되면서 사업에서 솔루션을 활용하는 과업이 추가되었고, 사업은 무사히 마무리되었다. 그 활용 결과를 부족하지만 본 논문을 통해 공유하고자 했다. 물론 기관의 특성에 맞춰 커스터마이징 되지 않은 상용프로그램을 실험적으로 사용해 본 제한적인 결과라는 한계는 분명하지만 직접적으로 기록연구사가 활용가능한 공개재분류 솔루션의 적용 경험과 과정 등을 소개하고, 그 결과를 최초로 공개했다는 데에서 의미를 찾을 수 있을 것이다.

하지만 국가기록원의 공개재분류 위탁사업 자체가 비전자기록물을 위주로 이뤄지는 사업이고, 솔루션에 대한 분석이나 개발 등의 연구가 중심인 연구용역 사업이 아니었던 만큼 결과값의 분석이나 개선점 제안에 있어 정치하지 않고 단순 소개에 그쳤다는 지적은 피할 수 없을 것이다. 솔루션 위주의 과업이었다면 분석대상 건수를 줄여서라도 정확한 대상과의 결과값 비교를 통해 정확도(precision)나 재현율(recall)을 계산할 수 있었을 것이다. 또한 오류내역을 세분화, 구체화하는 작업 등을 병행했다면 더욱 가치있는 자료가 됐을 거란 아쉬움도 남는다. 또 해외 사례 등에 대한 조사가 부족한 점도 인정한다. 이런 아쉬움은 국가기록원 등의 다양한 연구용역과 사업이 진행 중인만큼 후속 연구에서 보완되길 기대해본다.

참고문헌

- 국가기록원 (2009). 잠자는 기록물에 날개를 달자: 기록물 공개재분류 실용매뉴얼: 기록관 편.
- 국가기록원 (2020a). 국가기록관리혁신 자료집.
- 국가기록원 (2020b). 2020년 비공개 기록물 공개재분류 추진계획.
- 국가기록원 (2020c). 2020년 비공개기록물 공개재분류 결과 건별입력 위탁사업 제안요청서.
- 국가기록원 (2021). 국가기록원 소장기록물 공개제도 개선계획.
- 국가기록원 홈페이지 (2021.08.20.). 공개재분류 기록물 서비스 출처: <https://theme.archives.go.kr//next/closed/page/viewMain.do>
- 권미현 (2019). 비공개기록물 공개재분류 업무절차 개선 방안. 제8차 기록관리연구세미나 발표자료.
- 나라장터 국가종합전자조달 (2021.07.21.). 입찰정보 검색. 출처: <http://www.g2b.go.kr/index.jsp>
- 류한조, 이경남, 황진현, 임진희 (2021). 지능형 문서처리 도입과 기록관리 변화에 관한 연구. 기록학연구, 68, 41-72.
<https://doi.org/10.20923/kjas.2021.68.041>
- 방재현 (2018). 지능형 아카이브 시스템을 위한 기계학습 기술 적용 방안 연구. 심층신경망 적용을 중심으로. 박사학위논문, 한국외국어대학교 대학원 정보·기록학과.
- 송재혁 (2010). 비공개기록물의 공개재분류 활성화 방안에 관한 연구. 석사학위논문, 중앙대학교 대학원 기록관리학과.
- 송혜진, 김유승 (2012). IPA를 활용한 중앙행정부처 기록전문직 직무인식에 관한 연구. 한국기록관리학회지, 12(3), 117-136.

<https://doi.org/10.14404/JKSARM.2012.12.3.117>

신성혜 (2009). 기록물의 공개 재분류 효율화 방안. 석사학위논문, 원광대학교 대학원 사학과.

안희진 (2006). 비공개기록물의 재분류제도 연구. 석사학위논문, 충남대학교 대학원 기록보존학협동과정.

이을지, 이호신 (2016). 교육청 소속 기록연구사의 직무분석에 관한연구. 한국기록관리학회지, 16(3), 131-156.

<https://doi.org/10.14404/JKSARM.2016.16.3.131>

이현자 (2010). 기록관의 공개 재분류제도에 관한 연구. 석사학위논문, 충남대학교 대학원 기록보존학협동과정.

임희연 (2016). 기록관에서의 공개재분류 제도 개선 방안. 서울특별시교육청 사례 중심. 기록학연구, 49, 277-297.

<https://doi.org/10.20923/kjas.2016.49.277>

장혜진 (2016). 기록물 공개 재분류 제도 개선방안 연구. 석사학위논문, 한남대학교 대학원 기록관리학과.

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

An, Hui-Jin (2006). Study about the release system of classified record. Master's thesis, The Graduate School of Chungnam National University The Department of Archival Science.

Bang, Jae Hyun (2018). A Study on Application of Machine, Learning for Intelligent Archive System, -Focusing on Application of Deep Neural Network. Doctor's thesis, The Graduate School of Hankuk University of Foreign Studies The Department of Information and Archival Studies.

Jang, Hye-jin (2016). A Study on Improvement Measures for the Access Re-Review System of Records. Master's thesis, The Graduate School Hannam University The Department of Archives Management.

Korea ON-line e-Procurement System (2021.07.21.). Bid information search. available: <http://www.g2b.go.kr/index.jsp>

Kwon, Mihyun (2019). Measures to improve the business procedure for public reclassification of non-disclosed records. the 8th Seminar on Record Management.

Lee, Eul Ji & Lee, Hosin (2016). A Study on the Analysis of the Tasks of Record Managers in Offices of Education. Journal of Korean Society of Archives and Records Management, 16(3), 131-156.

<https://doi.org/10.14404/JKSARM.2016.16.3.131>

Lee, Hyun Ja (2010). (A) Study on Access Re-review System at the Records Center in Korea. Master's thesis, The Graduate School of Chungnam National University The Department of Archival Science.

Lim, Heeyeon (2016). Improvement of access re-review in archives: The Seoul Metropolitan Office of Cases. The Korean Journal of Archival Studies, 49, 277-297. <https://doi.org/10.20923/kjas.2016.49.277>

National Archives of Records (2020a). National Archives of Records Management Innovation Data.

National Archives (2020b). Plan to promote the public re-classification of non-disclosure records in 2020.

National Archives (2020c). Request for proposal for an input consignment project for each case as a result of the 2020 non-disclosure record disclosure reclassification.

National Archives (2021). plan to improve the disclosure system of archives held by the National Archives.

National Archives of Korea (2009). Let's attach wings to sleeping records: Re-classification of records Practical Manual: Archives.

National Archives of Korea homepage (2021.07.21.). Public reclassification record service. available:

<https://theme.archives.go.kr//next/closed/page/viewMain.do>

Ryu, Hanjo, Lee, Kyungnam, Hwang, Jinhyun, & Yim, Jinhee (2021). A Study on the Introduction of Intelligent Document Processing and Change of Record Management. The Korean Journal of Archival Studies, 68, 41-72.

<https://doi.org/10.20923/kjas.2021.68.041>

Shin, Sung-hye (2009). An Efficient Method of improving Access re-review of Records. Master's thesis, The Graduate School

of Wonkwang University The Department of Archival Science.

Song, Hye Jin & Kim, You-Seung (2012). A Study of Job Recognition for Record Manger using the IPA. Journal of Korean Society of Archives and Records Management, 12(3), 117-136. <https://doi.org/10.14404/JKSARM.2012.12.3.117>

Song, Jea Hyunk (2010). (A)study on the reclassification of non-disclosed records. Master's thesis, The Graduate School of Chung-ang University Department of Record & Archive Management Science.

