

빅데이터 분석 방법론을 활용한 지방자치단체 단위과제 운영 지원도구 개발 연구*

Research on Development of Support Tools for Local Government Business
Transaction Operation Using Big Data Analysis Methodology

김다빈(Kim, Dabeen)** · 이은정(Lee, Eunjung)*** · 류한조(Ryu, Hanjo)****

1. 서론
2. 지방자치단체 단위과제 현황 및 문제점
 - 1) 단위과제의 특성 및 운영현황
 - 2) 단위과제 생성 및 운영상의 문제점
 - 3) 단위과제 운영실정으로 인한 기록관리 문제점
3. 매핑 기반 표준화 도구를 활용한 단위과제 정비 방안
 - 1) 빅데이터 처리 기술의 적용 가능성 검토
 - 2) 표준 기록관리기준표를 활용한 매핑
 - 3) 표준화 도구 개발 및 적용 효과
4. 결론

* 본 연구는 “2021년 행정안전부 국가기록원 기록관리 연구개발사업”의 연구비를 지원받아 수행되었음.

** 한국의국어대학교 일반대학원 정보기록학 석사(제1저자)(ekqls1040@gmail.com).

*** 한성대학교 일반대학원 문헌정보학과 기록관리학전공 석사(공동저자)(su-ji94@hanmail.net).

**** 명지대학교 기록정보과학전문대학원 겸임교수(교신저자)(kalistland@gmail.com).

■ 투고일: 2021년 09월 30일 ■ 최종심사일: 2021년 10월 12일 ■ 최종확정일: 2021년 10월 23일.

■ 기록학연구 70, 85-117, 2021, <https://doi.org/10.20923/kjas.2021.70.085>

〈초록〉

이 연구의 목적은 지방자치단체에서 사용하고 있는 단위과제 현황, 단위과제 운영 및 기록관리 관점의 문제점을 조사 및 분석하여 그 과정에서 도출된 시사점들을 기반으로 텍스트 기반 빅데이터 기술을 활용하여 문제점에 대한 개선방안을 제시하는 것이다.

지방자치단체는 단위과제의 오분류로 인한 보존기간 책정 오류, 과공통사무와 기관공통사무의 유형식별 불가, 단위과제의 과대·과소·중복생성의 오류, 단위과제 명칭의 오류, 참고 가능한 표준의 부재, 통제 가능한 시스템 또는 도구의 부재 등으로 인해 기록관리 운영상 심각한 상태에 놓여 있다. 그러나 단위과제의 수가 약 72만개로 지나치게 많은 수량 때문에 효과적으로 통제할 수 없는 실정이며, 따라서 엄밀하고 통제할 수 있는 도구 및 표준이 필요하다.

본 연구에서는 이와 같은 문제점을 해결하기 위하여 빅데이터 분석 기술 중 텍스트기반 분석 도구인 코퍼스와 토큰화 기술을 적용한 시스템을 개발하고, 이를 기록관리기준표를 구성하고 있는 명칭 및 구성용어에 적용하였다. 이러한 단위과제 운영 지원도구는 통일성 있는 보존기간 책정, 위임사무 기록물 식별, 중복·유사단위과제 생성 통제, 공통과제의 표준적인 운영 등을 지원할 수 있는 도구가 될 수 있어 기록관리 업무에 상당한 기여를 할 수 있을 것으로 예상된다. 따라서 향후 빅데이터 분석 방법론을 활용한 지원도구가 BRM 및 RMS 등과 연계할 수 있다면 기록관리기준표 관리 업무의 품질이 높아질 수 있을 것으로 보인다.

주제어 : 기록관리기준표, 단위과제, 빅데이터, 지방기능분류체계, 지방자치단체

〈Abstract〉

The purpose of this study is to investigate and analyze the current status of unit tasks, unit task operation, and record management problems used by local governments, and to present improvement

measures using text-based big data technology based on the implications derived from the process.

Local governments are in a serious state of record management operation due to errors in preservation period due to misclassification of unit tasks, inability to identify types of overcommon and institutional affairs, errors in unit tasks, errors in name, referenceable standards, and tools. However, the number of unit tasks is about 720,000, which cannot be effectively controlled due to excessive quantities, and thus strict and controllable tools and standards are needed.

In order to solve these problems, this study developed a system that applies text-based analysis tools such as corpus and tokenization technology during big data analysis, and applied them to the names and construction terms constituting the record management standard. These unit task operation support tools are expected to contribute significantly to record management tasks as they can support standard operability such as uniform preservation period, identification of delegated office records, control of duplicate and similar unit task creation, and common tasks. Therefore, if the big data analysis methodology can be linked to BRM and RMS in the future, it is expected that the quality of the record management standard work will increase.

Keywords : Records management standard table, Business Transaction, Bigdata, Local Business Reference Model, Local Government

1. 서론

『공공기록물 관리에 관한 법률』 개정 과정에서 업무에 기반한 기록 관리체계를 구축하기 위해 2008년 도입된 정부기능분류체계(Business Reference Model, 이하 BRM)는 비교적 빠르게 정착되었다. 함께 도입된

기록관리기준표는 단위과제라는 기능분류의 최하위 단위를 기준으로 기능유형을 규정하고, 보존기간을 책정하여 생산되는 기록에 그 속성을 상속하도록 하였다.

중앙행정기관에서부터 도입된 기록관리기준표는 지방자치단체에 도입하면서 약간의 변화가 일어났다. 지방자치단체에 적합한 기능유형을 설정하고, 중앙행정기관과 구분되는 별도의 고유한 기능분류체계를 도입했던 것이다. 이후 지방자치단체에 도입된 기록관리기준표는 현재까지 10년 이상 기록관리의 근간으로 역할을 수행하고 있다.

그러나 지방자치단체의 BRM과 기록관리기준표는 운영기간에 비해 표준화 도구를 쉽게 찾아보기 어려우며, 운영의 엄밀함을 위한 시스템적 도구도 찾기 어렵다. 특히 기록관리의 기본단위인 단위과제의 생성이나 변경을 해당기관의 자율에 부여하여 과소·과대·중복운영을 쉽게 찾아볼 수 있기도 하다. 반면 단위과제별로 운영되는 기록관리기준표의 보존기간의 최종적인 확정권한은 국가기록원에 있어 전국 약 72만 개 이상의 단위과제에 대한 검토책임이 부여되어 있다.

2008년 기록관리기준표 도입 이후 기술의 변화나 행정환경의 변화에 따라 현재의 기록관리기준표의 운영방식을 다시 검토해볼 필요가 있다. 지방자치단체의 업무는 지속적으로 확장되고 있으며, 법령의 변화에 의해 신설되거나 위임되는 업무도 계속해서 변화하고 있다. 또한 대량의 기록을 알고리즘을 이용해 공개여부를 재분류하거나 평가하는 등의 정보통신기술도 속속 등장하고 있다. 기록 뿐만 아니라 기록관리기준표 역시 이러한 흐름에 맞추어 효율적이고 통일성 있는 기록관리체계를 위한 변화를 모색해볼 필요가 있다.

이러한 흐름의 대표적인 연구는 기록관리에 대한 빅데이터의 적용에 관한 연구이다. 빅데이터 기술을 이용하여 기록의 공개재분류 및 분류체계 적용, 기술(Description) 등 대한 개선과제를 도출한 연구가 대표적이며(오효정, 2019), 텍스트·음성·영상인식 등의 기술을 활용하여 기

록 검색의 편의성을 높이기 위한 방안 제시도 의의를 지닌다(김인택, 안대진, 이해영, 2017). 또한 공문서에 대해 텍스트 인식 및 메타데이터 추출을 통해 기록분류와 기술에 대한 적용가능성에 대한 연구도 같은 맥락에서 진행되었다(류한조, 이경남, 황진현, 임진희, 2021).

반면, 현재 기록관리기준표 운영에 대한 문제점을 인식하고 개선하기 위한 연구도 살펴볼 필요가 있다. 국가기록원에서는 운영상의 통일성을 기하기 위해 단위과제의 보존기간 책정이나 기록관리기준표 작성 및 관리에 대한 표준을 제시하였으나(국가기록원a, 2007; 국가기록원b, 2020), 실제 단위과제에 대한 통일성 있는 보존기간의 부여가 일어나지 않는 현실을 지적하고, 과대·과소·중복생성되는 문제를 해결하기 위한 방안에 대한 연구도 진행되었다(강혜원, 2013; 김화경, 김은주, 2014). 또한 지방자치단체의 기록관리기준표에 대한 다양한 오류를 지적하고 이에 대한 해결모형을 제시하는 연구도 찾아볼 수 있다(정상희, 2020).

지방자치단체의 기록관리기준표 운영에 대한 문제는 여러 연구에서 지적하여 개선의 필요성에 대한 공감대가 형성되어 있다고 볼 수 있다. 또한 문서기록이 많은 공기록의 특성 상 텍스트 분석을 토대로 한 빅데이터 기술의 기록관리에 대한 적용에 대한 관심도 확인할 수 있었다. 그러나 기존 연구는 빅데이터 기술에 대한 개괄적인 설명과 타 영역의 사례를 들어 기록관리에 대한 적용가능성을 제시하거나 기록관리기준표 운영 개선을 위해 운영의 엄밀성을 강조하는데 그치고 있다. 범용적으로 적용될 수 있는 구체적 방법론을 제시하기보다 기술의 소개나 문제점을 고발하는데 집중했다고 할 수 있다.

이에 본 연구는 텍스트 분석을 기반으로 한 빅데이터 기술을 구체적으로 활용하여 기록관리기준표 개선에 적용하는 방안을 제시하였다. 파이썬(python) 언어를 기반으로 방대한 텍스트 데이터를 처리하여 용어단위로 인식하는 기술이나 용어간의 품사와 유사도를 토대로 텍스트를 연결하는 기술을 기록관리기준표 운영 개선을 위한 도구개발에 사용하였다.

구체적으로 먼저 지방자치단체에서 일어나고 있는 단위과제의 운영문제점을 확인하기 위해 단위과제 운영에 대한 중복생성, 과대·과소생성에 대한 정량적 분석을 실행하였다. 또한 통일적으로 운영될 필요가 있는 기관공통형태의 단위과제나, 법적 근거로 인해 식별되어야 할 필요가 있는 위임사무가 포함된 단위과제에 대해 정성적인 분석을 진행하였다.

확인된 문제를 개선하기 위한 방안으로 지방자치단체에서 사용하는 단위과제의 언어적 형태를 분석하고, 단위과제를 구성하는 용어와 그 관계어를 토대로 유사단위과제를 식별할 수 있는 표준화 도구를 개발하였다. 이러한 표준화 도구는 유사한 단위과제끼리 비교하여 보존기간을 통일하거나, 위임사무와 같은 특정한 속성을 가져야 하는 단위과제를 식별하는데 활용할 수 있음을 증명하였다.

또한 연구결과가 현장에 적용될 경우 통일성 있는 보존기간 부여, 단위과제의 업무속성(공통업무, 고유업무, 위임사무 등) 구분 등이 가능한지 실제 데이터 적용을 통해 그 타당성을 확인하였다.

2. 지방자치단체 단위과제 현황 및 문제점

1) 단위과제의 특성 및 운영현황

단위과제는 정부기능분류체계의 최하 단위로, 행정업무 수행의 기준이 되는 것으로서 업무의 유사성, 독자성을 고려하여 업무담당자가 소기능을 영역별, 절차별로 세분화한 업무영역을 의미한다(국가기록원, 2020). 최초 도입 시 업무와 서비스, 기술, 성과 등과 연계시켜 행정환경을 개선시킬 것으로 기대되었으나 전면적인 범정부 EA도입이 지연되면서 기록관리와 연계된 용도로 주로 사용되었다. 사실상 기록관리 단위를 위해 운영되고 있는 실정이다.

기록관리기준표는 단위과제를 기준으로 기록관리항목을 추가하여 제시한 표로, 업무설명, 보존기간, 보존기간 책정사유, 비치기록물 해당 여부, 공개여부, 접근권한 등의 기록관리항목이 포함된다. 또한, 단위과제 단위로 보존기간이 책정되며 단위과제 하위 과제관리카드는 단위과제의 보존기간을 상속받아 운영된다. 따라서 기록관리기준표를 통해 단위과제별로 보존기간과 처리과 공개여부값 등이 연결되고 있다.

단위과제는 몇 가지 특성을 찾아볼 수 있는데, 우선 행정언어라는 정확하고 한정된 어휘만을 사용한다는 점을 들 수 있다. 또한, 모든 단위과제는 명사로 이루어져있으며 관형어, 목적어, 서술어의 역할을 하는 각각의 명사로 구분되어 기능한다. 이 중 주로 관형어, 목적어, 서술어로 이루어져 있으나 업무 대상의 명칭만으로 의미가 충분히 전달되는 경우 서술어가 생략되는 경우도 상당수 발견할 수 있다.

이런 특성과 원칙에도 불구하고 단위과제 운영현황은 업무와 기록의 관리를 유기적으로 연계하기 위해 평가, 분류, 편철 등 기록관리의 기준을 BRM의 단위과제로 삼고자 했던 최초의 설계대로 흘러가고 있지 않다. 단위과제 운영현황은 중앙행정기관과 지방자치단체가 다르게 운영되고 있으며, 이에 대한 현황분석결과를 구체적으로 살펴보고자 한다.

단위과제는 수행부서 및 업무의 특성에 따라 기능유형을 크게 처리과 공통업무, 기관 공통업무, 유사기관 공통업무, 위임사무업무, 고유업무 5가지로 구분할 수 있으며, 더 크게 구분한다면 공통업무와 고유업무로 구분할 수 있다. 중앙행정기관의 단위과제 유형은 4가지로 구분되며, 지방자치단체 단위과제 유형은 2가지로 구분된다. 법령에 의해 구분이 필요하다고 여겨진 사항과 중앙행정기관 및 지방자치단체의 단위과제 유형 및 설명을 결합하여 표로 구성하였다. 구분의 상세내역은 <표 1>과 같다.

〈표 1〉 기능분류체계 상 단위과제 유형 및 설명

유형	중앙행정기관	지방자치단체
처리과 공통업무	<ul style="list-style-type: none"> · 행정기관의 모든 과(팀) 단위에서 공통적으로 수행하는 업무 · 상급기구(부처)에서 보내오는 관련 업무에 대하여 처리과의 업무수행 내용보고나 자료제출 등 초안적 성격의 업무 · 처리과 직원에 대한 일상적인 복무관리 등에 관한 업무 · 처리과 서무담당자가 수행하는 업무 · 예시: 서무, 지도감사, 기록물관리, 물품관리, 민원관리, 예산회계, 보안관리 등 	<ul style="list-style-type: none"> · 모든 부서(처리과)에서 공통적으로 수행하는 업무 · 현용 지방자치단체 BRM 내에서 중기능 하위 '과공통일반사무' 분류에 포함되는 업무 · CAMS의 bizType 'B20024' 값에 해당하는 업무 · 예시: 서무, 지도감사, 기록관리, 민원관리, 물품관리, 업무계획보고 등
위임사무	<ul style="list-style-type: none"> · 위임사무는 국가에서 지방자치단체로 위임하는 사무를 지칭하므로 중앙행정기관은 존재하지 않음 	<ul style="list-style-type: none"> · 국가 또는 상급행정기관이 본래 수행하여야 할 사무 중 법률 또는 조례 및 규칙에 의하여 지방자치단체가 처리하도록 위임한 사무 · 지방자치단체 단위과제는 위임사무를 유형으로 구분하는 것이 아니라 속성으로 구분하며, 기초와 광역 값을 추가할 수 있음 · 예시: 여권발급수수료관리, 물가인정화관리 등
기관 공통업무	<ul style="list-style-type: none"> · 독립적으로 운영되는 모든 행정기관을 유지하기 위하여 기관들이 공통적으로 수행하는 업무 · 기관의 고유업무 수행을 위하여 설치된 처리과가 아닌 감사부서, 총무부서, 기획부서, 홍보부서, 예산회계부서 등 지원부서에서 수행하는 업무 	<ul style="list-style-type: none"> · 지방자치단체를 운영·유지하기 위해 수행하는 업무로 개별 부서의 업무 특성에 따라 독자적으로 수행하는 업무 · 지방자치법에서 규정하는 사무를 토대로 확장되는 업무
유사기관 공통업무	<ul style="list-style-type: none"> · 하나의 부처에서 고유기능을 수행하기 위해 업무를 지역별/대상별로 분담하여 기관들이 수행하는 부처 내 공통기능 · 예시: 국세청의 지방국세청 및 세무서 등 	<ul style="list-style-type: none"> · 지방자치단체 특성상 고유사무이지만 타 기관과 중복·유사한 업무가 많음(ex. 복지 등) · 현용 지방자치단체 BRM에서 '자치단체일반' 이라는 분류로 다양한 단위과제 유형이 통합하여 사용되고 있음 · 예시: 교통안전관리, 가맹점관리, 식품안전관리 등
고유업무	<ul style="list-style-type: none"> · 기관 고유기능 · 각 기관의 설립목적을 구현하고 법령 등에 의해 수행하는 각 기관 자체적으로 고유하게 수행하는 업무 · 예시: 기획재정부의 균형발전투자계획총괄 조정 등 	

단위과제 보존기간 책정 조정 지침에 따르면 중앙행정기관의 단위과제의 유형은 크게 처리과 공통업무, 기관 공통업무, 유사기관 공통업무, 고유업무 4가지로 구분된다(국가기록원, 2007)(〈표 1〉 참조).

‘처리과 공통업무’는 주로 처리과 서무업무 담당자가 수행하는 업무를 의미하며, 행정기관의 모든 과(팀)단위에서 공통적으로 수행하는 업무를 의미한다. ‘기관 공통업무’는 독립적으로 운영되는 모든 행정기관을 유지하기 위하여 기관들이 공통적으로 수행하는 업무로, 기관의 고유업무 수행을 위하여 감사부서, 총무부서, 기획부서, 홍보부서, 예산회계부서 등 지원부서에서 수행하는 업무를 의미한다. ‘유사기관 공통업무’는 하나의 부처에서 고유기능을 수행하기 위해 업무를 지역별/대상별로 분담하여 기관들이 수행하는 부처 내 공통기능을 의미한다. ‘고유업무’는 ‘자치사무’를 지칭하며, 각 기관의 설립목적을 구현하기 위해 각 기관에서 자체적으로 고유하게 수행하는 업무를 의미한다.

중앙행정기관의 4가지 구분과 달리 지방자치단체 단위과제 유형은 중앙행정기관의 처리과 공통업무에 해당하는 ‘과공통일반사무’와 그 밖의 모든 사무를 지칭하는 ‘자치사무’ 2가지로 구분되어 있다. 지방자치단체는 지방자치법에 명시되어 있는 업무를 모든 기관에서 수행하므로 공통업무와 구분하는 것에 대한 어려움이 존재한다. 그런 구분이 힘들어서 현재까지 표준이 부재하며 결국은 과공통사무와 자치사무 2가지로 운영되고 있다.

따라서 기관공통업무는 모든 지방자치단체에서 수행하는 업무이며 분명히 존재하나, 별도로 시스템 상에서 구분할 수 있는 도구가 존재하지 않아 통계작성 및 통일적 정책 적용 등을 위해서는 별도의 식별이 필요한 상태이다.

지방자치단체 내에서 단위과제의 정량적인 현황은 연계규격의 문제로 인하여 전체현황을 국가기록원에서 취합하거나 분석할 수 없음을 확인하였다. 추가로 구체적인 현황분석에 대한 결과는 다음과 같다.

첫째, 국가위임사무는 사무유형을 고유사무, 위임사무로 구분하여 입력하면서 자동으로 식별할 수 있도록 설계되어 있으나, CAMS의 연계규격값¹⁾을 기준으로 하는 이에 대한 정보는 확보하기 어려운 상태이다.

또한, CAMS에서 추출한 지방자치단체 BRM 데이터와 한국지역정보개발원에서 추출한 지방자치단체 BRM 데이터는 규격 값이 통일되지 않고, 시스템 상 연계가 미비하여 분산되어 보존되고 있으므로 위임여부 등을 확인할 수 없다.

둘째, CAMS와 지역정보개발원의 BRM 데이터가 연계되고 있으나 CAMS에서 보유하고 있는 정보가 지역정보개발원에서 보유하고 있는 정보보다 적어서 CAMS로 국가 전체의 단위과제에 대한 정보를 포괄할 수 없다. 또한, 이로 인해 단위과제 수량의 차이가 있음을 발견하였다. 해당 사례로, 서울시 A구를 대상으로 분석하였을 때 지역정보개발원의 BRM 데이터에서는 ‘환경개선부담금부과및징수’ 단위과제가 존재하나 CAMS 상에는 관련 단위과제가 존재하지 않음을 확인하였다.

셋째, 기존의 처리과공통 표준과 기관공통 표준은 중앙행정기관 용²⁾으로 정비되어있어 지방자치단체에 적용하기 힘들다. 또한, 위임사무에 관한 표준은 존재하지 않는 실정이다. 따라서 기관 내에서 표준을 참조하여 통일성 있고 정확하게 단위과제를 운영하는데 상대적으로 어려움이 존재한다.

2) 단위과제 생성 및 운영상의 문제점

소기능-단위과제는 업무에서 생산되는 기록을 찾을 때나 업무를 인수인계 할 때, 또는 부서별 기능이관을 하거나 기록을 찾을 경우에 효과적인 업무 수행을 위해 도입되었다. 그러나 실제로 살펴보니 과대·과소생성으로 인하여 엄밀한 인수인계 또는 기능이관이 되지 않고 있어 BRM 도입 당시 설계했던 최초의 취지가 지켜지고 있다고 보기에 어

1) 국가기록원 (2021). NAK 29-4:2021(v1.1) 영구기록관리시스템 데이터연계 기술규격: 제1부 기능분류시스템과의 연계.
2) 국가기록원 (2020). 기록물 관리지침(공통 매뉴얼).

럽다. 따라서 지방자치단체의 단위과제 생성 현황과 이로 인해 발생하는 단위과제 생성 및 운영상의 문제점들을 살펴보고자 한다.

현재 지방자치단체 단위과제 생성 및 운영상 현황을 파악하기 위하여 첫째, CAMS로부터 추출한 기록관리기준표를 토대로 소기능 대비 단위과제 생성비율, 기관별 단위과제 수량, 처리과 공통 및 자치사무의 단위과제 수량을 살펴보았다. 둘째, 소기능 및 단위과제의 과대·과소 생성 현황을 살펴보았다. 셋째, 소기능 및 단위과제의 중복생성 혹은 유사생성 현황을 살펴보았다. 넷째, 지방자치단체의 중요관리대상인 위임사무 운영 및 관리에 관한 문제점과 현황을 살펴보았다.

소기능별 단위과제 생성 비율은 다음 <표 2>와 같다.

<표 2> 광역지자체 소기능, 단위과제, 소기능당 생성비율, 공무원 1인당 단위과제수(일부)

구분	공무원 수	소기능	단위과제	소기능 1개당 단위과제 수	공무원 1인당 단위과제 수	공무원 1인당 소기능 수
광주광역시	2,393	1,029	3,326	3.23	1.39	0.43
대구광역시	3,517	1,572	8,111	5.16	2.31	0.45
대전광역시	2,443	1,142	3,938	3.45	1.61	0.47
부산광역시	4,520	1,452	6,626	4.56	1.47	0.32
서울특별시	10,973	1,565	9,363	5.98	0.85	0.14
세종특별자치시	1,767	1,297	5,093	3.93	2.88	0.73
울산광역시	1,957	1,304	5,707	4.38	2.92	0.67
인천광역시	3,972	1,860	7,965	4.28	2.01	0.47
강원도	2,321	1,287	5,640	4.38	2.43	0.55
경기도	4,341	1,645	7,869	4.78	1.81	0.38
경상남도	2,491	1,374	4,852	3.53	1.95	0.55
경상북도	2,383	1,356	5,728	4.22	2.40	0.57
전라남도	2,288	1,355	6,226	4.59	2.72	0.59
전라북도	1,909	1,199	4,889	4.08	2.56	0.63
제주특별자치도	2,252	849	1,819	2.14	0.81	0.38
충청남도	2,130	1,223	4,216	3.45	1.98	0.57
충청북도	1,830	1,162	4,619	3.98	2.52	0.63

〈표 2〉의 분석결과로 우선 단위과제 운영 시 기관 간 편차가 존재함을 발견하였다. 〈표 2〉에서 소기능 1개당 단위과제 생성수량은 최소 2개, 최대 5개로 약 2.5배 차이가 났으며, 소기능 수, 단위과제 수 등에서 편차가 큼을 확인할 수 있다.

다음으로 불명확한 단위과제 규모와 단위과제수의 적정성 문제이다. 사례로, 현재 B기관의 C처리과제에서는 1개의 단위과제만을 사용하고 있어 지나치게 단위과제를 과대운영하고 있음을 발견하였다. 물론 단위과제의 규모는 기관의 사정과 환경에 따라 적절하게 변동될 수 있다. 그러나 원칙적으로 1개의 소기능을 더 이상 분리될 수 없게 세분화한 개념이 단위과제임을 감안하면 사실상 해당 B기관은 ‘업무의 과정과 결과의 누락 없는 기록화’라는 법의 취지를 구현하지 못하고 있으며, 과제관리카드 운영 현황 또한 오류를 범하고 있을 확률이 매우 높다고 볼 수 있다.

또한, 공무원 1인당 단위과제 수를 분석하였을 때 단위과제의 과소생성이 발생하고 있음을 확인하였다. 특히 가장 적은 수량을 나타내는 제주특별자치도와 서울특별시의 경우 1인당 1개의 단위과제도 담당하지 않는 것으로 보인다. 이는 단위과제의 과소생성의 문제점으로, 소기능이 업무담당자가 수행하는 기능의 업무수준이라는 BRM 운영 원칙에 상당히 위배되는 현황이라고 볼 수 있다.

단위과제 생성 시 문제점 중 수량이나 기관 간 편차 외에도 명명의 문제가 심각함을 발견하였다. 국가기록원의 기록물관리 지침 등을 살펴보면 단위과제 명명 시 사용하지 말아야 할 규칙들을 제시하고 있으나 제대로 지켜지지 않고 있음을 확인하였다. 사례로, ‘myjobidea추진’, ‘(용달/택배/개별)화물자동차운송사업관리’, ‘경로당’ 등이 있다. 특수문자사용 금지, 영어 사용금지, 업무를 나타낼 수 있는 표현 사용 등의 규칙을 지키지 않은 채로 단위과제를 생성하고 있음을 발견하였다.

현재 단위과제 생성 시 문제점은 기관 간 편차가 심하다는 것과 1인

당 단위과제 및 소기능이 과소 생성되고 있다는 것이다. 동일한 지방자치법을 근거로 한 자치사무에 대해 담당수준의 소기능을 세분화한 것이 지방자치단체의 단위과제이므로, 단위과제 수량의 편차가 큰 현상은 기관마다 단위과제를 생성하는 문화나 기준이 다르기 때문으로 추정된다. 또한, 단위과제를 생성하기 위해서는 분류체계 등 입력하여야 할 정보가 많을뿐더러 기관 BRM 담당자의 승인을 받아야 하는 등의 불편함과 복잡함이 있다. 이에 비해 과제관리카드는 단위과제에 비해 비교적 간단하고 쉽게 만들 수 있으므로 과제관리카드가 과대 생성되는 문제점이 발생한다. 그밖에도 또 다른 문제로 중복생성 혹은 유사생성 등의 문제가 있음을 발견하였다.

광역지방자치단체 중 가장 많은 단위과제를 생성한 D광역시의 소기능 별 단위과제 생성 및 활용 현황에 대해 분석 한 결과 소기능 별 단위과제의 중복생성 혹은 유사생성 등의 문제점을 발견하였으며 일부 예시는 <표 3>과 같다.

〈표 3〉 D광역자치단체 동일처리과 소기능 별 단위과제의 중복생성 사례(일부)

동일처리과 동일 소기능 아래 중복단위과제 생성	처리과명	소기능	단위과제
	시민소통기획관시민봉사담당관	민원제도개선	120다산콜센터상담원교육
	시민소통기획관시민봉사담당관	민원제도개선	120다산콜센터상담원교육
	시민소통기획관시민봉사담당관	민원제도개선	120다산콜센터상담자료관리
	시민소통기획관시민봉사담당관	민원제도개선	120다산콜센터상담자료관리
동일처리과 타 소기능 아래 중복단위과제 생성	처리과명	소기능	단위과제
	소방서소방행정과	지도감사사무	감시수감및결과조치
	소방서소방행정과	지도감사(사무)	감시수감및결과조치
	소방서예방과	지도감사사무	감시수감및결과조치
	소방서예방과	지도감사(사무)	감시수감및결과조치
	소방서재난관리과	지도감사사무	감시수감및결과조치
소방서재난관리과	지도감사(사무)	감시수감및결과조치	

	처리과명	소기능	단위과제
동일처리과 유사동일 소기능 아래 유사 단위과제 생성	푸른도시국조경과	녹화사업추진	가로녹지량확충
	푸른도시국조경과	녹화사업수행	가로녹지량확충및관리
	푸른도시국조경과	가로변녹화	가로수관리
	푸른도시국조경과	도시숲조성및가로녹화	가로수조성및관리
	푸른도시국조경과	도시숲조성및가로녹화	가로수조성사업

D광역자치단체의 경우 동일 처리과에서 사용하고 있는 같은 이름의 소기능 임에도 불구하고 ‘120다산콜센터상담원교육’, ‘120다산콜센터상담자료관리’와 같이 각 2건씩 유사하거나 동일한 단위과제명이 중복으로 생성되어 활용되고 있음을 확인하였다. 또한, D광역자치단체는 ‘감사수감및결과조치’ 라는 동일한 단위과제를 중복으로 생성하여 활용하고 있었으며, 소기능도 ‘지도감사서무’와 ‘지도감사(서무)’ 가 명명만 다를 뿐 기능은 같으므로 소기능 또한 중복으로 생성되어 활용되고 있음을 확인하였다. 뿐만 아니라 ‘녹화사업추진’-‘녹화사업수행’과 같이 유사하거나 동일한 업무를 포괄하는 소기능 아래 ‘가로녹지량확충’-‘가로녹지량확충관리’와 같은 유사한 단위과제를 중복 생성하여 사용하고 있음을 발견하였다.

위임사무의 경우 「공공기록물 관리에 관한 법률」 제11조 7항에 따르면 지방기록물관리기관의 장은 중앙기록물관리기관의 장이 국가위임사무에 관한 기록물의 원본 또는 사본의 이관을 요청할 시 이에 관하여 협조하여야 한다. 즉, 중앙기록물관리기관은 국가위임사무에 관한 기록물의 원본 또는 사본 이관을 요청하기 위해서는 사전에 국가위임사무와 이에 해당하는 기록물이 무엇인지 식별하고 관리하여야 한다.

그러나 위임사무 운영에 관하여 다음과 같은 문제점을 발견하였다. 첫째, 동일한 근거법령이 존재함에도 조례에 의해 기관 간 위임사무 운영에 관한 편차가 존재한다. 둘째, 법령으로 식별된 위임사무가 단위과제로 생성되는 원칙과 절차를 발견할 수 없다. 셋째, 포괄적으로 통제

하는 조직이나 담당이 없어, 기관 내 BRM을 총괄하는 부서의 주관성이 반영되고 있다.

지역정보개발원에서 추출한 광역자치단체 기록관리기준표의 위임사무 속성값을 부여받은 단위과제 현황을 분석하였다. 광역별로 전체 단위과제별 위임사무 단위과제 생성비율을 분석하였으며 그 결과는 <표 4>와 같다.

<표 4> 광역별 위임사무 현황분석 결과(일부)

광역	위임사무 단위과제수(개)	생성비율
서울	3	0.004%
부산	1058	2.053%
광주	703	3.617%
대전	1302	7.982%
경상북도	1581	3.262%
전라북도	882	3.904%
제주도	539	4.592%

전반적인 현황 분석 결과 광역별로 편차가 심하다는 것을 확인할 수 있다. 위임사무에 해당하는 단위과제가 가장 많은 곳은 경상북도이며 가장 적은 곳은 서울특별시이다. 서울특별시는 가장 많은 단위과제 수량을 보유하고 있음에도 위임사무가 가장 적은 곳으로 파악되었다. 위임사무 조사 결과 법령 내용을 통해 위임을 규정하고 있는 법률은 288개, 시행령은 267개로 각 법령 내용에 의해 위임된 사무는 289개 이상임을 확인할 수 있었다. 그러나 서울시의 경우 위임사무로 지정하여 사용하고 있는 단위과제의 개수가 3개뿐으로, 위임사무를 정확히 인식하지 못하며 위임사무를 단위과제명으로 반영하지 못하고 있음을 알 수 있다.

또한, 위임사무에 대한 정확한 식별이 어려워 과공통사무에 위임사무 속성을 부여한 사례도 확인할 수 있다. 해당 사례로, 단위과제 ‘일반서무업무’는 일반공공행정>일반행정>일반행정지원>과공통일반사무>일반서무>일반서무업무의 분류체계를 가지고 있으며 기능유형 또한 과공통일반으로 분류되어있어 명백하게 과공통일반사무이나, 위임사무 속성값이 부여되어 위임사무로 오분류되고 있다. 뿐만 아니라, 동일 위임사무에 대해 단위과제가 중복으로 생성된 사례도 발견하였다. E구청의 경우, 「건설산업기본법」에 의거하여 위임사무 ‘건설업 등록접수 및 등록’에 대해 ‘전문건설업등록관리’, ‘전문건설업신고및등록관리’ 두 가지의 단위과제가 생성되어있음을 발견하였다.

위임사무는 국가의 사무를 법률 등에 따라 지방자치단체가 처리하는 것으로, 위임사무를 수행하는 기록물은 반드시 관리하여야 한다. 따라서 사전에 국가위임사무와 이에 해당하는 기록물이 무엇인지 식별하여야 한다. 즉, 법으로 정해진 위임사무를 수행하는 단위과제에 정확한 위임여부 속성값이 부여되어야 하나 지켜지지 않고 있음을 확인하였다. 해당 사례로, F구청의 경우 「공유수면 관리 매립에 관한 법률」에 의거하여 위임사무 ‘지방관리구역항 내 공유수면의 점용 실시계획 승인 사항에 대한 준공검사’에 대한 단위과제가 반드시 생성되어야 하나 생성되지 않음을 확인하였다. 또한, G시청의 경우 「악취방지법」에 의거하여 위임사무 ‘생활악취를 위한 대책 수립 및 시행’에 대해 마찬가지로 단위과제가 생성되지 않음을 확인하였다.

지역정보개발원의 위임사무 단위과제 현황분석 결과 발견한 문제점은 위임사무의 개념이 명확하게 정의되어 있지 않아 이로 인해 해석의 여지가 증대될 수 있다는 점이다. 이는 단위과제 등록 단계 시 위임사무를 통제할 수 있는 절차나 시스템의 부재, 참고 가능한 표준이나 지침 등의 부재, 법적의무를 수행할 수 있는 제반요건의 부재 등으로 나타날 수 있는 결과이다.

종합적으로 단위과제 운영 시 가장 두드러진 문제점은 과대·과소·중복 생성의 오류 등이며 이로 인하여 효율적이고 효과적인 기록 분류 및 보존기간 책정이 일어나지 않고 있다. 또한, 과공통일반사무 뿐만 아니라 기관공통적 성격의 단위과제의 생성 및 운영에서 중앙행정기관에 비해 통제되지 않고 다양한 형태를 보이고 있다. 그밖에도 불분명한 단위과제의 규모, 단위과제 명명의 오류 등으로 인하여 엄밀하고 통일성 있게 단위과제가 운영되고 있지 않았다.

3) 단위과제 운영실정으로 인한 기록관리 문제점

단위과제는 기록관리의 근간이 되는 최소 단위로, 단위과제를 기준으로 보존기간, 공개여부, 비밀여부 및 접근권한 등을 분류하여 기록관리가 이루어져야 한다. 그러나 앞서 언급하였듯이 현재의 기록관리는 통일성 있게 생성·운영되지 못하는 공통과제 등으로 인하여 과제관리카드의 생성도 효과적으로 일어나지 않고 있는 심각한 상황이다.

이로 인한 기록물 관리 및 활용의 문제점은 다음과 같다. 첫째, 보존기간의 잘못된 책정으로 인한 문제가 있다. 둘째, 단위과제별 공개여부 책정 시, 하나의 단위과제 아래 비공개와 공개기록물이 혼재될 수 있다. 셋째, 처리과공통/기관공통사무의 구분이 어려워 오분류가 상당수 발생하고 있다. 넷째, 위임사무 기록물은 법에 근거하여 법에 명시된 바, 반드시 위임사무 기록물을 관리하여야 하나 현재는 식별하지 못하여 법적 요구에 대응하기 어려운 실정이다.

단위과제 생성 시 분류계층을 오분류하여 잘못 지정하는 사례가 흔히 있으며, 특히 과공통업무와 기관공통업무에 대한 혼동이 많다. 이러한 오분류는 단기간 가치가 있는 기록이 장기보존 되거나 반대로 장기보존 되어야 할 중요문서가 단기간 보존될 수 있는 심각한 문제로 이어지게 된다(정상희, 2016). 또한, 단위과제 하위 과제관리카드는 단위과

제의 보존기간 및 공개여부 등의 정보를 상속받으므로 과제관리카드의 보존기간 문제로까지 이어질 수 있다. 또한, 보존기간 책정은 우선 처리과에서 1차로 책정하고, 이를 기록관에서 검토한 후 국가기록원과 협의 및 확정하는 절차를 거쳐야 한다. 기록관과 국가기록원은 적게는 수천 개에서부터 많게는 만 개가 넘는 기관 전체의 단위과제 보존기간을 책정하여야 하다 보니 업무에 대한 면밀한 분석과 충분한 검토 없이 보존기간을 확정한다(정상희, 2016). 이렇듯 기관 내에서도 보존기간 단위과제 생성-검토-확정 과정에서 지나치게 업무 부담이 많다.

기록관리기준표의 관리항목 중 공개여부는 기록물 관리의 현실을 고려하여 근본적인 재검토가 요구된다(정상희, 2016). 또한, 공개여부는 건별로 생산당시 결정되므로 실무적인 입장에서도 공개여부나 접근제한은 건별로 파악할 수밖에 없는 것이 현실이다(설문원, 2006). 단위과제와 과제관리카드를 잘못 생성하게 된다면 개인정보와 같은 보호받아야 할 기록물이 부분공개 또는 공개가 되는 심각한 문제가 발생하게 된다.

처리과공통업무의 경우 현용 지방자치단체 BRM 내에서는 분류명을 잘못 표기하여 사용하는 경우, 처리과공통업무 임에도 불구하고 과공통일반사무에 분류되지 않는 등의 단위과제 운영실정으로 인한 기록관리의 문제점을 발견하였다. 또한, 국가기록원에서 제시한 처리과공통업무 표준 기준표(국가기록원, 2020)에 명시된 소기능을 중심으로 해당 소기능에 종속된 단위과제를 처리과 공통업무로 식별하도록 되어 있으나 소기능의 오지정 등 오류가 발견되었으며, 뿐만 아니라 ‘보안관리’ 소기능을 ‘보안’, ‘보안일반’, ‘보안업무일반’ 등으로 변경하여 사용하는 경우도 발견되었다.

기관공통업무의 경우 기관공통업무와 유사기관공통업무로 구분될 수 있으나, 지방자치법에 규정된 다양한 업무를 공통적으로 수행하는 지방자치단체의 특성상 유사기관공통업무를 실질적으로 구분하기 어

려워 속성 식별에 문제점이 있음을 확인하였다.

국가위임사무 단위과제의 값의 정확성 여부를 검토하는 절차 및 확인활동이 별도로 존재하지 않으므로 정확성을 담보하기 어렵다. 뿐만 아니라 앞서 분석한 결과처럼 명확하게 처리과공통사무임에도 불구하고 위임사무의 속성값이 부여되어 공통사무와 위임사무가 서로 혼재되어 있음을 확인하였다.

앞서 분석한 결과를 토대로 도출한 종합적인 문제점은 다음과 같다. 첫째, 단위과제 속성값이 통일되지 않고 있다. 둘째, 지방자치단체 공통업무 단위과제에 대한 표준화가 구축되지 않았다. 셋째, 위임사무에 대하여 단위과제 유형이 별도로 생성되지 않았으며, 속성값이 일치되게 부여되어 있지 않다. 넷째, 처리과 공통표준이 존재하나 상당수 지켜지고 있지 않다. 다섯째, 지방자치단체에도 기관공통 성격의 업무가 분명하게 존재하나 중앙행정기관용 업무 표준을 그대로 준수하기 어려운 상황이며, 지방자치단체에 적용 가능하거나 참고 가능한 표준이 없다. 여섯째, '지방자치법에 의거한 업무'와 '기관공통업무'는 모두 전 지방자치단체에서 공통적으로 수행한다는 특성으로 인하여 기관공통과 처리과공통에 대한 구분이 매우 모호하다. 일곱째, 개별법령에서 보존기간이 존재할 경우 이를 준수하기 위한 별도의 조치(단위과제 보존기간 책정 등)가 취해지지 않고 있다. 여덟째, 유사·공통업무를 수행하는 경우에도 동일업무의 단위과제에 대해서 단위과제명, 보존기간 등이 상이하여 통일성 있는 기록관리가 이루어지지 못하고 있다.

이를 종합해보면 효율적이고 효과적인 기록 분류 및 보존기간 책정을 위해서 도입된 BRM 체계가 잘 작동되고 있지 않음을 알 수 있다. 또한, 현재 지방자치단체 기록관리시스템에서 관리되는 단위과제는 최소 72만 건 이상으로(17개 광역의 경우, 단위과제 평균 6~7천 건이며 228개 시군구의 경우 약 4~5천 건) 기록관리시스템의 엑셀 함수 기능만으로는 체계적인 관리의 한계가 있다.

3. 매핑 기반 표준화 도구를 활용한 단위과제 정비 방안

1) 빅데이터 처리 기술의 적용 가능성 검토

단위과제 생성 및 운영상의 문제, 단위과제 운영실정으로 인한 기록 관리 문제점으로 인해 BRM 체계가 잘 작동하고 있지 않아 효율적인 기록 분류와 보존기간 책정이 수행되지 않음을 확인하였다. 그러나 현재 까지 단위과제의 과소·과대·중복 생성 및 오분류 등에 대한 개선방안은 찾기 어려운 실정이다. 이에 대한 개선을 위해 각 지방자치단체에서 사용하고 있는 기록관리기준표를 대상으로 실제로 빅데이터 기술이 적용 가능한지를 검토하여 가능성을 확인해보고, 어떠한 기술의 적용이 타당할지에 대해 검토해보고자 한다.

BRM 구조에서는 단위과제를 기준으로 보존기간, 공개여부, 업무설명 등 기본적인 관리항목을 부여하게 되어있다. 따라서 본 연구에서는 각 지방자치단체에서 기록관리기준표를 운영하며 사용하고 있는 단위과제명을 연구의 분석 대상이자 분석의 기준으로 삼고자 한다. 단위과제 명칭을 통한 분석은 제한된 언어 사용으로 인해 분석이 상대적으로 수월하다는 장점을 가진다.

다양한 빅데이터 분석기법이 존재하나, 그중에서 비정형 데이터인 텍스트를 대상으로 하는 빅데이터 분석기법을 적용하고자 한다. 단위과제명의 분석을 위해 적용 가능한 기술로는 먼저 텍스트 마이닝³⁾ 기법과 NLP(Natural Language Processing)⁴⁾ 전처리(Preprocessing) 도구의 활

3) 텍스트 마이닝(Text Mining) 기법은 빅데이터 분석 또는 데이터 과학의 대표적인 분야로, 비정형 텍스트에서 자연어 처리와 형태소 분석 기술을 두어 유용한 단어를 추출해 빈도수를 분석하고 맥락 수준의 의미를 찾아내는 분석방법을 의미한다.

4) 자연어 처리(NLP)는 컴퓨터를 이용하여 사람 언어의 이해, 생성 및 분석을 다루는 인공지능 기술로(박대민, 2016) 테스트의 수준에 따라 문자 또는 문자열, 형태소,

용이 있다. 전처리 과정을 통해 단위과제명을 구성하는 텍스트를 정제하는 과정을 거쳐 텍스트로부터 정보를 추출하기 위한 사전 작업이 요구되기 때문이다. 텍스트로부터 정보 추출 시 다의성과 비정형성, 모호성 등 고유한 특성(김수현 외, 2019)으로 인해 원하는 정보만 추출하기가 어렵고 인지능력만으로 텍스트의 자료를 처리하는 데 한계가 있다. 자연어는 동일한 의미를 가지는 단어가 다양한 형태로 표현되거나 단어의 조합에 따라 새로운 의미가 만들어지는 경우가 많아 텍스트가 가지는 정보가 정확하게 인식되지 않을 수 있다. 마찬가지로 단위과제명은 언어로 구성된 텍스트 데이터로서 텍스트가 가지는 특징을 상속한다. 이 때문에 단위과제명 텍스트가 가진 정보의 정확도를 높여주는 역할을 할 수 있는 텍스트 전처리 과정이 필요하다.

연구를 위한 분석 대상으로 지방자치단체에서 사용하는 기록관리기준표 단위과제 총 724,005개를 수집하여 살펴보았다. 단위과제를 구성하는 행정어는 자연어보다 범위가 작고 엄밀하다는 특성이 있으며 자연어에 비해 용어가 통일되어 있다. 보편적으로 ‘허가, 인가, 확인, 통지, 수리’ 등 행정행위로 끝나는 형태들을 보이며 행위의 유형을 토대로 행정업무의 기능을 파악할 수 있음을 발견하였다. 또한, 빅데이터 기술을 통해서 단위과제 명칭을 구성하는 관형어, 목적어, 서술어 등으로의 구분을 시도해볼 수 있다. 행정어를 활용한 구성과 관형어, 목적어, 서술어로 분해가 가능하다는 점은 일반적인 자연어 처리와 비교하여 도드라지는 차이점으로, 텍스트 마이닝 기술 적용에 있어 자연어에 비해 더 수월한 분석이 가능할 것으로 보인다.

구체적으로는 토큰화(Tokenization) 개념을 적용할 수 있다. 텍스트로부터 정보를 추출하기 위해서는 의미 단위의 분해가 필요한데, 이때 토큰화 개념을 적용하여 단어를 식별하고 이를 하나의 유용한 의미 단위

단어, 구, 절, 단문과 복문 등의 문장, 문단, 문서, 복수 문서 등의 다양한 수준에서 수행될 수 있다.

로 분해할 수 있다. 단위과제명의 특성상 형태소, 조사, 띄어쓰기, 어미 등의 변화는 제한적으로 일어나기 때문에 비교적 간단한 전처리 과정을 통해서 텍스트 토큰화가 가능할 것으로 보인다.

오히려 본 연구에서는 단위과제명을 구성하고 있는 구조 혹은 용어에 대한 이해가 빅데이터 처리에 더 중요한 요소로 작용할 수 있다. 또한, 토큰화 된 모든 용어를 분석에 사용하지 않으며 연구 과정 안에서 의미파악에 활용되지 않는 부사 등의 용어나 특수문자는 분석에서 제외시키는 과정이 필요하다. 단위과제명을 대상으로 텍스트 토큰화를 실행하여 단·복수형, 축약어 및 띄어쓰기, 오타 등을 정제하는 과정을 거친 후 코엔엘파이(KoNLPy)의 코퍼스(Corpus) 등 자연어 전처리 기술을 파이썬에서 구현하는 방식을 통해 단위과제를 구성하는 용어들을 추출하여 분석이 가능하다.

단위과제에 대해 직접적으로 품사 구분의 개념 적용은 쉽지 않다. 우리말인 명사는 목적어와 서술어가 직접적으로 지칭되지 않으므로 동명사의 개념에도 다양한 논쟁이 있다. 따라서 토큰화를 통해 언어를 분해한다 하더라도 행정 환경에서 사용하는 목적어와 서술어 등을 별도의 라이브러리를 구성하여 구분 할 수 있는 추가적인 도구가 필요할 수 있다.

용어들의 문법적인 기능에 따라 관형어, 목적어, 서술어의 품사를 부여하면 용어의 품사별 집합 관리가 가능하다. 단, 72만개 가량의 단위과제명을 살펴보았을 때 목적어와 서술어 두 가지의 품사로서 기능하는 단어들이 존재함을 알 수 있었다. 서술어 뒤에 '~하다'가 생략된 형태로 표현된 경우(ex. 처분, 등록)와 그 자체가 행위의 대상인 경우(ex. 처분요구, 등록대상)가 발견되었다. 이와 같은 예외적인 사항의 처리를 위해서는 관형어, 목적어, 서술어 외에도 목적어/서술어 품사의 별도 개설 등이 필요할 수 있다. 또, 품사 구분을 했을 때 업무의 명칭만 존재하고 '관리' 등의 서술어가 생략된 경우도 발견할 수 있었다. 빅데이터 처리를 위해서는 이와 같은 예외사항을 별도로 반영할 필요가 있다.

용어의 품사 구분은 빅데이터 처리를 활용한 단위과제 정비방안에 중요한 작용을 할 수 있다.

텍스트 마이닝과 토큰화를 통해 시스템 안에서 텍스트의 의미 단위 인식이 가능해짐에 따라 용어 간의 의미 관계 분석이 가능해진다. 이는 텍스트에서 추출된 용어가 순서대로 배열되는 현상이 다른 용어와 일치할 경우 이를 하나의 의미 단위로 취급하는 알고리즘을 기반으로 수행된다. 이를 기준으로 유사한 텍스트를 군집화한 후 시스템 알고리즘 안에서 같은 의미를 가진 유사한 용어로 인식하게끔 할 수 있다. 뿐만 아니라 군집화를 통해 용어들을 유사한 의미별로 가치 있게 구조화하고 활용할 수 있으며 품사별로 구분된 용어들의 집합적 관리를 수행할 수 있다. 즉, 빅데이터 분석기술인 텍스트 마이닝과 자연어 전처리, 토큰화, 품사 인식 및 군집화 등의 기술을 약간의 보완 및 적용으로 단위과제에 대한 분석 및 검토가 가능하다.

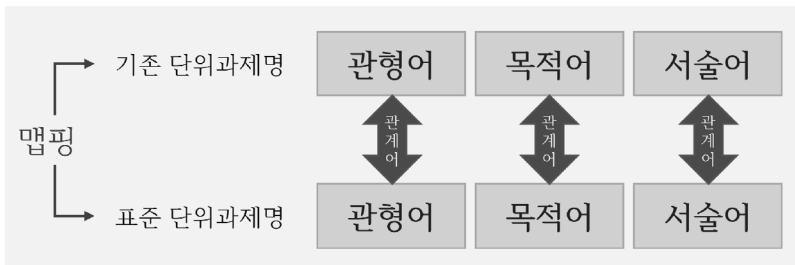
2) 표준 기록관리기준표를 활용한 매핑

기록관리기준표 개선에 기여할 수 있는 구체적인 방안으로 표준 기록관리기준표를 활용한 매핑을 제시하였다. ‘매핑’이란 표준 단위과제가 존재한다는 전제하에 품사별 매핑 알고리즘 형태가 시스템에 적용된 방식으로, 현장에서 사용하고 있는 단위과제명의 사례들을 표준 단위과제와 연결 짓는 방식을 의미한다. 이 과정은 앞서 언급한 빅데이터 분석기술의 적용에 기반하여 수행될 수 있는데, ①현장 사례와 표준의 단위과제명의 구조가 같다는 점과 ②텍스트 분석을 통해 시스템 내에서 용어의 의미관계 인식 및 용어의 군집화가 가능하다는 조건을 통해 작동한다.

단위과제명은 관형어+목적어+서술어의 형태를 보이는 동시에 일반적인 용어에서 시작하여 구체적인 용어로 끝나는 경향성을 가진다. 제

일 뒤에 위치하는 서술어를 통해 업무처리 방식을 명확히 하고, 관형어와 목적어를 통해서는 업무의 범위와 대상을 정의한다. 이러한 명명규칙에 따라 현장에서 사용하고 있는 단위과제명과 표준 단위과제명이 모두 같은 조건을 가지고 있으므로 이를 기반으로 매핑 개념의 도입이 가능하다. 텍스트 매핑 개념을 간단하게 도식화 해보면 다음 그림과 같이 표현할 수 있다.

〈그림 1〉 텍스트 매핑 개념 도식화



기존 단위과제명과 표준 단위과제명의 매핑은 단위과제를 구성하는 용어 간의 유사성에 의해 발생한다. 표준 단위과제명을 구성하는 목적어와 현장에서 사용하고 있는 단위과제명(기존 단위과제명)의 목적어가 같은 품사로서 기능하며 같은 의미 관계 안에 포함되는 관계어(또는 유사어) 이고, 마찬가지로 관형어와 서술어 모두 같은 조건을 만족할 시 기존 단위과제명과 표준 단위과제명의 매핑이 일어난다.

매핑과정에 적용되는 원칙을 정리하면 다음과 같다.

1. 대상과제의 목적어와 서술어를 구성하는 용어는 표준의 용어와 관계어로 구성되어야 한다.
2. 관형어가 존재하는 경우 관형어, 목적어, 서술어 모두 관계어로 구성되어야 한다.
3. 서술어와 목적어로 동시에 사용될 수 있는 용어는 서술어를 우선하여 적용한다.

예를 들어 표준 단위과제 ‘교육훈련(목적어)관리(서술어)’는 기존 단위과제 ‘직무교육관리’, ‘공직자교육훈련관리’와 매핑될 수 있다. ‘교육훈련’이 목적어로 기능하며 관계어인 ‘직무교육’, ‘공직자교육훈련’과 연결되고, 서술어인 관리가 같은 조건을 만족하므로 매핑이 발생한다. 이렇듯 매핑 기술의 구체적인 적용은 표준과 현장 사례를 일대 다 방식으로 매핑하여 통제할 수 있는 방법으로 발전이 가능하다. 이를 통해 단위과제의 다양한 이형을 식별하고 이에 대한 표준 단위과제를 제시함으로써 표준화를 유도하는 방식으로 기록관리기준표 개선에 기여할 수 있을 것으로 보인다.

위와 같은 알고리즘에 따라 매핑이 시스템적으로 처리되기 위해서는 준용할 수 있는 표준과 용어들 간의 관계가 정의되어 있는 용어사전 구축이 전제되어야 한다. 또한, 매핑이 시스템 등에 의해 통일성 있고 적절하게 적용되기 위한 방식으로 몇 가지 원칙이 필요하다. 먼저 표준 개발 과정에서 자치단체에서 사용하고 있는 다양한 기관의 단위과제 현황과 용례를 함께 고려하여 융통성 있는 설정 규모를 정하고 이를 원칙으로 도출할 필요가 있다. 단위과제는 업무 간 유사성과 독자성을 고려하여 소기능을 세분한 업무 영역으로 각 업무 분야별 담당자 수, 절차적 단계, 근거법령 등에 따라 과제의 설정 규모가 달라질 수 있기 때문이다. 또, 표준 개발 시 효율적이고 효과적인 매핑 지원을 위해 표준 단위과제의 명칭은 구체적인 철명이나 업무절차를 나열하기보다 정확한 대상과 수행 업무를 표현할 수 있어야 한다.

원활한 매핑을 위해서는 용어사전에 대한 엄밀한 원칙이 요구된다. 이때 각 용어는 동일한 품사 내에서만 관계어를 설정하며 타 품사와는 중복하여 관계어를 설정하지 않아야 원활한 매핑이 이뤄질 수 있다. 또, 관계어를 설정하는 단계에서 상위어 및 하위어는 같은 의미로 묶지 않는다. 1대 다 방식으로 매핑 되는 방식의 특성상 상위어/하위어를 사용한 표준에 너무 과대·과소한 단위과제들이 매핑될 가능성이 있기 때문이다.

마지막으로 매핑 알고리즘에서 단위과제명을 인식할 때 맨 뒤에서부터 앞의 순서로 인식하도록 만드는 원칙이 필요하다. 이 원칙은 수행하는 업무의 핵심 행위를 맨 뒤에 위치하는 서술어로 구분하는 단위과제명의 특성으로 인해 요구된다. 즉, 단위과제명의 의미를 해석하고자 할 때에는 과제명의 마지막 용어에서부터 앞의 순서로 인식하여 업무 행위를 서술어를 기준으로 구분하여 해석하는 것이 효율적이다.

매핑 알고리즘과 원칙의 적용을 통해 처리과공통사무 표준 ‘교육훈련관리’를 활용하여 기존 지자체에서 과공통일반사무로 분류하여 사용하고 있는 단위과제에 대한 식별을 시도해보았다. 그 결과의 예시 일부는 다음 <표 5>와 같다.

<표 5> 매핑 알고리즘과 원칙을 적용한 예시(일부)

표준	처리 내역(기존 단위과제명)
교육훈련관리	공무원교육일반
	공무원교육훈련관리
	공무원상시학습
	교육관련
	교육상시학습
	상시학습실적관리
	직원교육관리
	공무원교육훈련상시학습
	환경관련공무원교육
	정보화교육운영
	직무교육운영
	교육일반
	공직자교육일반
	교육일반상시학습
	교육훈련관련

1개의 처리과 공통단위과제와 유사하게 사용되는 과공통단위과제의 개수는 총 174개로 다양한 이형으로 분산되어있는 단위과제들이 하나의 표준과 유사/동일단위과제로 매핑됨을 확인할 수 있다.

3) 표준화 도구 개발 및 적용 효과

단위과제의 표준화된 운영을 위해 기존 단위과제명과 표준 단위과제명의 매핑을 토대로 한 기록관리기준표 통제 도구를 개발해 볼 수 있다. 매핑 기반 분석 도구는 표준 기록관리기준표 도출과 용어사전 구축 등 다양한 제반 작업을 종합하여 적용한 결과물로, 매핑 알고리즘이 탑재된 시스템 형태로 개발이 가능하다. 분석 도구는 전처리 도구와 용어사전, 표준이 유기적으로 맞물려 돌아가는 구조이며 기록관리기준표에 대한 표준화를 유도하는 표준화 도구로서 기능한다. 시스템 사용자가 개선이나 정비를 원하는 기록관리기준표를 업로드하고 정비를 실행하면 해당 기록관리기준표 단위과제명과 매핑되는 표준을 제안하는 방식의 시스템으로 구현될 수 있다. 시스템 형태의 표준화 도구는 향후 AMS, RMS 및 BRM 관리 시스템과의 연계도 고려해볼 수 있다. 연계를 통해 RMS의 검토 및 CAMS에서 '확정' 하는 단위과제의 보존기간 책정에 도움을 줄 수 있고, BRM과 연계하여 사전에 중복 생성 여부를 파악할 수 있다. 이러한 연계가 가능해진다면 나아가 단위과제 정비 사업 추진이나 국가기록원의 보존기간 검토 책임에 대한 업무 부하를 줄일 수 있을 것으로 보인다.

표준화 도구는 기록관리기준표 운영 과정에서 발견한 문제들에 대한 개선을 지원할 수 있다. 표준화 도구를 통한 여섯 가지 문제점 개선 방안을 제시하였다. 첫째 '국가위임사무 단위과제의 식별'과 두 번째 '기관공통업무 단위과제 식별'은 상대적으로 빠르게 실현이 가능하다. 그러나 나머지 네 개의 개선방안의 경우 표준화 도구의 개발과 보완이 충

분히 완료된 후 단계적으로 적용이 가능해질 것으로 예상된다.

첫째, 국가위임사무 기록이 생산되는 단위과제를 식별하고 대표할 수 있는 단위과제를 표준으로 규정하여 지속적으로 위임사무를 관리할 수 있는 체계를 마련할 수 있다. 현재 자치단체에서는 국가위임사무 성격의 단위과제 식별이 이루어지지 않아 기록관리의 통일성 부여 및 국가기록물에 대한 사전통제 등의 업무에 어려움을 겪고 있다. 위임사무의 식별을 위해서는 국가위임사무가 포함되어 있는 조례와 관련 사무를 정의한 후 이와 매핑되는 단위과제를 빅데이터 분석 도구를 활용해 식별하는 방식이 적용 가능하다.

둘째, 지방자치단체 기관공통업무의 식별이 가능하다. 기관공통업무의 경우 모든 자치단체에서 수행하는 업무이나 시스템적으로 구분할 수 있는 방안이 존재하지 않아 별도의 식별이 요구되었다. 지방자치단체 기관공통업무의 식별을 위해서는 국가기록원에서 공개한 기관공통업무에 해당하는 소기능 및 단위과제명을 기반으로 기관공통업무를 정의한 후 이와 유사한 단위과제를 빅데이터 분석 도구를 활용해 식별하는 방식이 적용 가능하다. 기관공통업무 식별 후 지속적으로 이를 관리할 수 있는 체계를 마련할 수 있다.

셋째, 단위과제의 다양한 이형을 식별하고 이에 대한 일원화된 보존기간 등의 속성을 부여하여 속성값의 통일성 있는 관리와 통제가 가능하다. 표준 개발 과정에서 보존기간 등의 속성값을 작성하여 시스템에 탑재하는 방식으로 기존 단위과제에 해당하는 올바른 속성값을 제안할 수 있다. 이 경우 표준과 매핑되는 단위과제들을 그룹화하여 속성값을 손쉽게 통일시킬 수 있다. 유사·동일 단위과제에 대해 각기 다르게 책정되어 있던 속성값의 통일적인 관리로 더 양질의 기록관리기준표의 운영이 가능해질 것으로 보인다.

넷째, 단위과제의 중복·과소·과대 생성 여부를 파악하여 통제가 가능하다. 많은 지방자치단체가 동일·유사 단위과제를 중복 생성하여

분류체계의 장점을 활용하지 못하고 있는 실정이다. 빅데이터 분석도구를 이용할 경우 표준과 일대 多로 매핑되는 유사·동일 단위과제의 파악 및 통제가 가능해질 수 있다. 또한, BRM과 연계하여 단위과제의 신설 과정에서 표준화 도구를 통해 기존에 생성된 유사·중복 단위과제가 존재하는지 점검이 가능해져 단위과제 생성 이전에 타당성을 검증할 수 있다.

다섯째, 타당한 소기능의 선택에 도움을 주어 기록관리기준표 분류체계 오지정 문제를 개선할 수 있다. 표준 기록관리기준표 개발 과정에서 소기능을 정확히 지정하여 적합한 소기능 하위 단위과제를 파악할 수 있게 한다. BRM과 연계하여 단위과제의 신설 과정에서 표준화 도구를 통해 표준 기록관리기준표를 기반으로 해당 단위과제가 속하는 소기능을 안내하거나 제안하도록 시도해볼 수 있다. 이를 통해 필요한 기록을 쉽게 찾도록 하는 분류체계 지원의 효과를 얻을 수 있을 것으로 보인다.

여섯째, 개별법령에서 보존기간을 제시하고 있는 단위과제들을 식별하고 법령에서 책정하고 있는 정확한 보존기간을 준수하도록 도울 수 있다. 지방자치단체가 준수하여야 개별법령을 정의한 후 이와 매핑되는 단위과제를 빅데이터 분석 도구를 활용해 식별하는 방식이 적용 가능하다.

기존 단위과제명과 표준 단위과제명의 매핑을 통한 단위과제의 유형별 식별을 위해서는 무엇보다 엄밀한 표준의 구축이 전제되어야 한다. 표준의 정합성을 높이기 위해 속성값 작성 시 적합한 근거를 마련하고, 단위과제 명명규칙 준수 및 과제 규모 설정의 타당성 등을 중요하게 고려해야 할 것으로 보인다.

2014년 국가기록원에서 배포한 지방자치단체 기록관리기준표 표준에서는 전자제어를 통해 유사 단위과제를 표기하여 단위과제의 의미별 유사성 파악을 시도한 바 있다. 그러나 유사하거나 동일한 업무임에도

단위과제명에 대한 이형이 다수 발견되었을 뿐만 아니라 많은 이형들을 모두 전거제어 형식으로 표준에 수록하지 못했다는 한계를 지니고 있다.

본 연구에서 제안하는 표준화 도구는 매핑 방식을 통해 이형을 식별하는데 효과적으로 작용한다. 단위과제명의 파악과 통제를 활용할 수 있어 각각 지방자치단체에 보급되어 단위과제 생성에 어려움을 겪거나 기존 단위과제의 정비를 계획하는 업무담당자에게 지원 도구가 될 수 있다. 또한, 현재 기관들이 사용하는 기록관리기준표 생성 문화를 적절히 반영하여 데이터를 입력할수록 효용성이 높아 질 것으로 예상된다.

4. 결론

본 연구는 빅데이터 분석기술을 적용한 기존 단위과제명과 표준 단위과제명의 매핑을 통해 기록관리기준표 운영 과정 중 단위과제의 생성 전·후에 모두 적용 가능한 보조 도구를 제시하였다는 데에 의의가 있다.

현황분석을 위해 지방자치단체의 단위과제 총 724,005개를 수집하여 살펴본 결과 많은 기관에서 단위과제 생성·운영상의 문제 및 기록관리 관점의 문제점이 발견되었다. 단위과제의 과소, 과대, 중복 생성으로 효과적인 기록 분류 및 보존기간 책정이 일어나지 않고 있으며, 단위과제 생성 시 분류계층과 유형을 오분류하는 사례가 다수 발견되었다. 이는 기관 공통적 성격의 단위과제와 위임사무 단위과제의 식별이 명확하게 이루어지지 않아 발생하는 문제이다. 이를 통제할 수 있는 절차나 시스템이 부재함과 동시에 기록관리의 기본단위인 단위과제의 생성과 변경을 해당 기관의 자율에 부여하였다는 점이 원인으로 작용했다고 볼 수 있다. 그러나 현재까지 단위과제의 과소·과대·중복 생성

및 오분류 지정 등에 대한 개선방안은 찾기 어려운 실정이다.

다양한 빅데이터 기법이 존재하나 본 연구에서는 그중에서 비정형 데이터인 텍스트를 대상으로 하는 빅데이터 분석기법을 적용하여 단위과제명의 분석을 시도하였다. 단위과제는 행정언어라는 정확하고 한정된 어휘만을 사용하고, 모든 단위과제는 명사로 이루어져 있으며 각각 관형어, 목적어, 서술어로서 기능한다. 이러한 특성으로 단위과제명칭을 통한 분석은 제한된 언어 사용으로 인해 분석이 상대적으로 수월하다는 장점을 가진다. 기술의 적용 가능성을 검토해본 결과 빅데이터 분석기술인 텍스트 마이닝과 자연어 전처리, 토큰화, 품사 인식 및 군집화 등의 기술을 약간의 보완을 거쳐 적용하면 단위과제에 대한 분석 및 검토가 가능함을 확인하였다.

빅데이터 기술을 적용하여 기록관리기준표 개선에 기여할 수 있는 구체적인 방안으로 표준 기록관리기준표를 활용한 매핑을 제시하였다. 기존 단위과제명과 표준 단위과제명의 매핑이 발생하는 작동 논리와 몇 가지 원칙을 설명하였다. 또한, 매핑 알고리즘과 원칙의 적용을 통해 표준을 활용하여 현재 지방자치단체에서 사용하고 있는 단위과제에 대한 식별을 시도해 본 사례를 일부 제시하였다. 매핑 알고리즘이 탑재된 시스템 형태로 기록관리기준표 통제가 가능한 표준화 도구의 개발 가능성을 연구하였다. 시스템 형태의 표준화 도구는 매핑 방식을 통해 이형을 식별하는 데 효과적으로 작용한다. 단위과제명의 파악과 통제를 활용할 수 있어 각각 지방자치단체에 보급되어 단위과제 생성에 어려움을 겪거나 기존 단위과제의 정비를 계획하는 업무담당자에게 지원 도구가 될 수 있다.

그러나 표준화 도구는 실제 단위과제 관리카드의 명칭은 반영하지 못하였기 때문에 단위과제 관리카드의 생성 및 운영에 대한 개선은 다루지 못한다는 한계가 있다. 또한, 유사하거나 동일한 업무가 기관별로 다르게 적용되는 예외사항에 대해서는 대응하기 어렵다. 따라서 보조

도구로서의 의미를 지니지만, 전국 72만 개 이상의 단위과제 대용량 처리에는 효과적으로 작용할 수 있다. 앞으로도 빅데이터 분석기술을 활용한 표준화 도구에 대한 한계점 극복 방안이나 기능요건에 대해 더 많은 추가적인 연구가 필요할 것으로 여겨진다.

〈참고문헌〉

- 국가기록원 (2007). NAK-P-2009-08 단위과제 보존기간 책정조정 지침.
- 국가기록원 (2014). 지방자치단체 기록관리기준표 표준.
- 국가기록원 (2018). NAK 34:2018(v1.0) 교육(행정)기관 공통 단위과제 분류기준 및 보존기간 책정기준.
- 국가기록원 (2020). 중앙행정기관 단위과제 정비 매뉴얼.
- 국가기록원 (2020). 기록물 관리지침(공통 매뉴얼).
- 국가기록원 (2021). NAK 4:2021(v2.2) 기록관리기준표 작성 및 관리절차.
- 국가기록원 (2021). NAK 29-4:2021(v1.1) 영구기록관리시스템 데이터연계 기술규격: 제1부 기능분류시스템과의 연계.
- 김수현, 이영준, 신진영, 박기영 (2019). 경제 분석을 위한 텍스트 마이닝. 한국은행.
- 김인택, 안대진, 이해영 (2017). 인공지능을 활용한 지능형 기록관리 방안. 한국기록관리학회지, 17(4), 225-250.
- 김화경, 김은주 (2014). BRM 운영을 위한 단위과제 정비방안. 한국기록관리학회지, 14(4), 199-219.
- 류한조, 이경남, 황진현, 임진희 (2021). 지능형 문서처리 도입과 기록관리 변화에 관한 연구. 기록학연구, 68, 41-72.
- 박대민 (2016). 뉴스 기사의 자연어처리. 커뮤니케이션 이론, 12(1), 4-52.
- 설문원 (2013a). 단위과제 기반 공공기록물 평가제도의 문제점과 개선방안. 한국기록관리학회지, 13(3), 231-254.
- 설문원 (2013b). 기록분류를 위한 정부기능분류체계의 적용 구조 및 운용 분석. 한국비블리아학회지, 24(4), 23-51.
- 오효정 (2019). QFD 기반 공공기관 기록관 서비스 품질 비교 및 개선방안. 기록학연구, 61, 119-156.
- 정상희 (2016). 광역지방자치단체 단위과제 운영에 관한 연구. 기록학연구, 49, 327-359.

정상희 (2020). 지방자치단체 기록분류체계 개선방안 연구 : 지방 BRM 정책영역 '지역맞도시'를 중심으로. 한국외국어대학교 대학원 정보기록학과 박사학위논문.