

AtoM 시스템의 구축과 커스터마이징 방법에 관한 연구*

안 대 진** · 김 익 한***

1. 머리말
2. AtoM 도입 시 고려사항
 - 1) OSS의 개념 및 특징
 - 2) AtoM 구조 및 기능
3. AtoM 개발사례 분석
 - 1) AtoM 개발 사례 및 특징
 - 2) 시사점
4. AtoM 시스템의 구축과 커스터마이징
 - 1) 요구분석
 - 2) 적합성 평가
 - 2) 마이그레이션
 - 3) 커스터마이징
5. 맺음말

* 본 연구는 명지대학교 기록정보과학전문대학원 석사학위논문(2014) 「AtoM 시스템의 구축과 커스터마이징 방법에 관한 연구」를 보완·수정한 것임.

** 사단법인 한국국가기록연구원 선임연구원(주저자).

*** 명지대학교 기록정보과학전문대학원 교수(교신저자).

▪투고일 : 2015년 6월 18일 ▪최초심사일 : 2015년 6월 26일 ▪게재확정일 : 2015년 7월 1일.

[국문초록]

ICA는 소규모 기록관을 지원하기 위해 2010년 웹 기반의 기록물 기술 소프트웨어인 AtoM 공개버전을 무료로 배포했다. AtoM 프로젝트의 목적은 단일기관, 혹은 복수의 기관이 웹을 통해 ICA 기술표준에 따라 기록물을 기술하고 온라인을 통해 접근을 제공하는 것이다. 기술적 전문성을 공유함으로써 실무경험을 촉진하고 사용자와 개발자들의 활발한 커뮤니티를 조성하는 것이 AtoM 프로젝트의 가치라 할 수 있다.

최근 국내 기록관리 분야에 오픈소스 소프트웨어에 대한 관심이 커지고 있다. 이러한 관심은 비용절감 외에도 기술종속성을 탈피하고 실무에서의 요구사항을 스스로 해결하려는 측면이 있다. 이 연구는 AtoM 시스템 구축을 위한 범용의 실제적 방법론이 없다는 데 문제의식을 가지고 있다. AtoM을 그대로 사용할 때에도 기존 데이터의 마이그레이션이나 사용자 인터페이스를 개선하는 등의 기본 작업이 반드시 필요하다. 따라서 본고에서는 AtoM 소프트웨어를 기반으로 아카이브 시스템을 구축하고 커스터마이징하기 위한 절차와 방법론을 제시하였다.

주제어 : 에이투엠, 오픈소스소프트웨어, 기록관리시스템

1. 머리말

소스코드의 공개와 자유로운 재사용을 주장하는 오픈소스 운동은 1990년대 말에 시작되었다. 이러한 흐름은 기록관리 분야에도 이어졌다. ICA는 소규모 기록관을 지원하기 위해 2010년 웹 기반의 기록물 기

술 소프트웨어인 AtoM¹⁾ 을 무료로 배포했다. AtoM 프로젝트의 목적은 단일기관, 혹은 복수의 기관이 웹을 통해 ICA 기술표준에 따라 기록물을 기술하고 온라인을 통해 접근을 제공하는 것이다. 기술적 전문성을 공유함으로써 실무경험을 축진하고 사용자와 개발자들의 활발한 커뮤니티를 조성하는 것이 AtoM 프로젝트의 가치라 할 수 있다.

이 연구는 AtoM 시스템 구축을 위한 범용의 실제적 방법론이 없다는 데 문제의식을 가지고 있다. AtoM의 수요기관들은 설치부터 커스터마이징까지 다양한 기술적 지원을 필요로 하고 있으나 국내 기록관리 커뮤니티가 자생적인 해결 능력을 갖추고 있지 못하기 때문이다. AtoM에 대한 보다 심도 깊은 연구가 필요한 시점이다.

국내 선행연구들은 OSS 기반 기록관리시스템 도입의 유용함과 운용가능성을 언급하고 있다. 이보람 외(2014)²⁾는 민간아카이브의 기록시스템 부재에 대한 돌파구로서, 공개 소프트웨어의 도입을 제안하고, 수집형 아카이브와 이관형 아카이브로 나누어 아카이빙을 수행한 결과를 소개하였다. 이 연구에서는 이관형 아카이브를 위한 목록 가져오기 기능과 보안의 취약함을 지적하였다. 송정숙 외(2014)³⁾는 오메카를 이용하여 부산항 사진 아카이브를 구축한 과정과 방법을 소개하고 있다. 수집된 기록을 더블링크어에 따라 기술한 뒤 부산항의 공간적 성격을 교통공간, 역사공간, 산업공간, 생활공간으로 구분하여 11개 전시 주제를 구성하였다. 이러한 개발방법론은 특정 컬렉션의 온라인 전시를 위한 오메카 사이트 개발에 특화된 것으로 AtoM 시스템을 개발하는 경우와

1) AtoM(<https://www.accessmemory.org>)은 'Access to Memory'의 약어로 소규모 기록관을 지원하기 위해 ICA가 주도적으로 개발한 웹 기반 기록물 기술 소프트웨어이다. ICA의 자금지원으로 v1.3까지 출시되었고, 이후 ICA와 함께 개발을 주도했던 Artefactual社가 2.2버전까지 발전시켰다.

2) 이보람, 황진현, 박민영, 김형희, 최동운, 최윤진, 임진희, 「공개 소프트웨어를 이용한 기록시스템 구축가능성 연구」, 『기록학연구』 39호, 2014, 193-228쪽.

3) 송정숙, 허정숙, 이에린, 「오픈소스를 이용한 부산항 사진 아카이브의 구축 방안」, 『한국기록관리학회지』 14(3), 2014, 127-151쪽.

는 구축의 목적과 방법 측면에서 차이가 있다. 또한 시스템 구축 방법 보다는 로컬리티의 재현방법론을 확산하고 공유하고자 하는 목적이 크다. 심갑용 외(2015)⁴⁾는 K-Food 콘텐츠의 체계적 관리와 이용자 열람의 편의성을 위해 AtoM을 이용하여 기록시스템을 구축한 과정을 소개하고 있다. 국가기록원의 ‘기록관리시스템 기능 요건’에 AtoM의 기능을 매핑하고 8가지 기능요건을 AtoM시스템에 어떻게 구현할지 적용방안을 기술하였다. AtoM이 웹 기반 기록물 기술과 접근에 특화되어 있음을 고려했다라면 기능요건이 좀 더 실효성 있게 선정될 수 있었을 것이다. 국가기록원의 8가지 기능요건 중 상당수는 현재 AtoM 시스템의 부차적인 기능이거나 제공되지 않고 있기 때문이다.

피터 밴 가더렌(2009)⁵⁾과 제시카 부시(2012)⁶⁾는 AtoM 프로젝트가 시작된 배경과 개발과정에 대해 상세히 기술하고 있다. 피터 밴 가더렌은 AtoM 프로젝트의 진행 배경을 소규모 아카이브의 예산 부족과 2000년대 초반의 오픈소스 운동으로 설명하고 있다. 소프트웨어 라이선스 비용을 지불하는 것보다는 오픈소스 기반의 시스템을 개발하고 엔지니어를 고용하는 것이 훨씬 비용을 절감한다는 것이다. 또한 AtoM의 개발을 위해 검토한 기존 OSS의 장단점 분석을 통해 AtoM 소프트웨어가 지향하는 바를 더욱 확실하게 이해하도록 해 준다. 블로그엔진, 디지털 리포지터리 시스템, 아키비스트 툴킷 등 기존의 OSS들은 기록물 계층구조 표현의 어려움, 호환성 없는 내부구조, 고비용 발생 등의 이유로 AtoM의 비전과 부합하지 않았다. 결국 웹서비스를 위한 기본 소프트웨어 세

4) 심갑용, 유현경, 문상훈, 이윤용, 이정현, 김용, 「한류문화콘텐츠의 기록화를 위한 AtoM 활용 방안에 관한 연구」, 『기록학연구』 43호, 2015, 5-42쪽.

5) Peter Van Garderen, *The ICA-AtoM Project and Technology*, Presentation at Association of Brazilian Archivists, Third Meeting on Archival Information Databases, Rio De Janiero, Brazil, 2009, pp.1-36.

6) Jessica Bushey, “ICA-AtoM: open-source software for archival description”, *Archivi & Computer article*, 2012, pp.1-16.

트인 리눅스, 아파치 웹서버, MySQL, PHP와 PHP 프레임워크인 심포니(Symfony), 데이터 모델을 위한 큐빗(Qubit) 등이 선택되었다.

본고에서는 AtoM 기반의 아카이브 시스템을 구축하고 커스터마이징하기 위한 절차와 방법론을 제시하였다. 이를 위해 OSS와 AtoM의 특징을 알아보고, 해외의 AtoM 개발사례 중 광범위한 기능개발, 소규모 커스터마이징, 마이그레이션 등 프로젝트 규모별 3개 사례를 선정하여 시사점을 도출하였다. 전통적인 소프트웨어 개발방법론과 OSS 개발 프로세스의 차이점을 연구한 사례는 많다. 이 연구가 그와 구분되는 점은 AtoM 시스템을 구축할 때 가장 신경써야 하는 기술표준 매핑과 마이그레이션, 기록물 이용환경을 향상시키기 위한 AtoM 시스템의 커스터마이징 요소 등을 상세히 다루었다는 것이다. 연구 결과를 일종의 기록시스템 구축 매뉴얼로 제공하기 위해 한국정보화진흥원의 'CBD SW 표준산출물 관리가이드'를 기반으로 프로세스와 산출물을 제시하였다.

이 연구의 목적은 AtoM을 활용하여 저렴한 비용으로 아카이브 시스템을 구축하고 보다 향상된 이용자 경험을 제공하도록 하는 것이다. 이 연구결과가 AtoM시스템 구축을 위한 기초자료로 이용되어 국내 AtoM 커뮤니티가 자생적으로 발전할 수 있는 계기가 되길 바란다.

2. AtoM 도입 시 고려사항

1) OSS의 개념 및 특징

OSS는 소스 코드에 대한 접근, 자유로운 재배포, 파생 저작물의 작성, 제한 없는 사용 등을 허용하는 라이선스와 함께 배포되는 소프트웨어이다.⁷⁾

OSS는 사용자들에게 해당 소프트웨어에 적용되는 라이선스를 준수할 것을 요구한다. OSS 라이선스에는 여러 가지 유형이 있으며 소스코드의 사용, 파생 소프트웨어의 개발, 원본 소프트웨어 및 파생 소프트웨어에 대한 비독점적, 상업적 이용의 허용 여부를 규정하고 있다.⁸⁾ AtoM은 GPL(GNU Public License)기반의 AGPLv3를 적용하고 있다.

모질라와 리눅스의 설계 구조에 대한 연구들은 많은 OSS 프로젝트와 개발방법론들의 근본적인 특징이 분리된 모듈들로 시스템 설계를 분해한 것임을 보여준다.⁹⁾ AtoM이나 오메카와 같은 OSS 역시 핵심 커널을 제외하고 대부분의 기능이 플러그인 모듈 형태로 구성되어 있다. 모듈화는 명확하고 이해할 수 있는 설계를 개발하는 것이 핵심이며, 분리된 개발자 그룹들에 의한 독립적이고, 수용적이며, 자율적인 기여를 가능하게 한다. 즉, 커뮤니티 구성원 개개인의 작은 코드 기여를 통해 매우 가치 있는 합을 만들어낼 수 있다는 것이다.

소프트웨어 개발은 사용자의 요구사항을 파악하고, 이 요구사항을 만족하는 소프트웨어를 설계하고, 구축하며, 테스트하여 고객에게 인도하기 위한 절차이다. 일반적인 개발 프로세스는 요구사항 분석, 설계, 구현 및 테스트, 통합, 문서화 등의 단계가 순차적으로 또는 반복적으로 진행되는 것이다.¹⁰⁾ 반면 오픈소스를 활용한 소프트웨어 개발은 초

7) Androutsellis-Theotokis, Stephanos, et al. "Open source software: A survey from 10,000 feet", *Foundations and Trends in Technology, Information and Operations Management* 4(3/4), 2011, pp.87-347.

8) Laurent, Andrew M. St., "Understanding Open Source and Free Software Licensing", O'Reilly Media, Inc., 2004.

9) MacCormack, Alan, John Rusnak, and Carliss Y. Baldwin. "Exploring the structure of complex software designs: An empirical study of open source and proprietary code", *Management Science* 52(7), 2006, pp.1015-1030.

10) Feller, J. and Fitzgerald, B., "A Framework Analysis of the Open Source Software Development Paradigm", in W. Orlikowski, P. Weill, S. Ang & H. Krcmar (Eds) Proc. of 21st Annual International Conference on Information Systems, (ICIS2000), Brisbane, Australia, 2000.

기의 계획 단계에서부터 오픈소스 활용의 타당성 검토와 활용 범위 분석 등 전략적인 접근이 이루어진다. OSS의 개발은 OSS 프로젝트의 생명 주기에 따라 다양한 역할을 가진 커뮤니티가 주도한다.

성공적인 OSS 프로젝트에는 폭넓은 사람들의 참여를 기반으로 한 커뮤니티가 형성되어 있다. 이러한 커뮤니티의 특징은 다양한 역할, 능력, 참여 정도, 책임감을 가진 멤버들로 구성되어 있다는 것이다. 코어 멤버는 프로젝트의 나아갈 방향에 대한 전략적 의사결정을 하며 액티브 멤버는 정기적으로 버그 수정이나 새로운 기능을 위한 코드 작업을 하는 액티브 개발자와 주로 버그를 수정하는 버그 수정자로 구성된다. 페리페럴 멤버들은 OSS 프로젝트 커뮤니티의 대다수를 차지하며, 프로젝트에 산발적으로 참여하여 지식을 교류한다. 패시브 유저는 커뮤니티 멤버는 아니지만 소프트웨어를 사용해 보고 비평이나 의견을 제공함으로써 결과적으로 프로젝트에 기여한다.¹¹⁾

마지막으로 OSS 운용예산이다. 많은 기업들이 오픈소스 운용에 대한 충분한 이해 없이 쉽게 도입을 결정한다. 가트너의 2008년 조사에 의하면, OSS를 적극적으로 사용하고 있는 전 세계 274개 기업 중 69퍼센트는 그들의 오픈소스 운용 예산을 측정해 보지 않았다고 한다.¹²⁾ 오픈소스 프로젝트의 성공을 좌우하는 핵심적인 요소는 총소유비용(TCO, Total Cost of Ownership)을 측정하는 것이다. 총소유비용은 기업이 특정 기술에 대한 구현비용을 결정하기 위해 고려해야 할 직간접비용의 합계이다.¹³⁾ 총소유비용의 비용요소로는 하드웨어 및 소프트웨어 획득, 설치, 교육 및 지원, 유지보수, 저장장치 등 기반구조, 가동중지비용, 공간 및 에너지 비용 등이 있다. 상용 소프트웨어를 사용하는 경우에 비

11) Androutsellis-Theotokis, Stephanos, et al. pp.187-347.

12) Computer Weekly 2008년 기사, (<http://www.computerweekly.com/news/2240087521/Firms-open-to-huge-open-source-liabilities>), [인용 날짜: 2015. 7. 10].

13) 위키백과 - 총소유비용, (<https://ko.wikipedia.org/wiki/%EC%B4%9D%EC%86%8C%EC%9C%A0%E B%B9%84%EC%9A%A9>), [인용 날짜: 2015. 7. 10].

해 OSS를 사용할 때 라이선스 비용은 획기적으로 줄지만, 교육 및 유지 보수 비용이 더 발생하곤 한다. 커스터마이징할 경우 거의 동일한 수준의 비용이 발생한다.¹⁴⁾ OSS의 투자수익(ROI, Return On Investment)을 극대화하기 위해서는 다음 사항들을 유념해야 한다. 첫째, 양질의 OSS를 선택해야 한다. 둘째, 개발 전문가의 도움을 얻을 수 있어야 한다. 셋째, 업데이트나 패치 등 OSS의 주기적 관리가 필요하다. 넷째, 라이선스를 준수하여 분쟁비용이 발생되지 않도록 해야 한다. 다섯째, OSS 커뮤니티에 전적으로 의존하지 않아야 한다. 통계적으로 39%의 질문자들은 커뮤니티로부터 응답을 받지 못했다.¹⁵⁾ 이를 해결하려면 아카이브에 상주하거나 도움을 얻을 수 있는 개발자가 있어야 한다.

2) AtoM 구조 및 기능

(1) 데이터 모델

데이터 모델이란 엔티티 사이의 관계를 기술하기 위한 개념적 표현이다. AtoM의 데이터 모델은 데이터 프레임워크인 큐빗 도메인 모델¹⁶⁾에 기반하여 기술, 전거레코드, 기록소장기관, 기능, 저작권 레코드, 용어(Terms) 등 6개 유형의 엔티티로 구성된다. ICA기술표준 4종은 기술, 전거레코드, 기록소장기관, 기능 엔티티에 매핑된다. 용어 엔티티는 텍스트노미나 시스템 어휘통제에 쓰인다. 이벤트(Event) 엔티티는 ISAD(G)

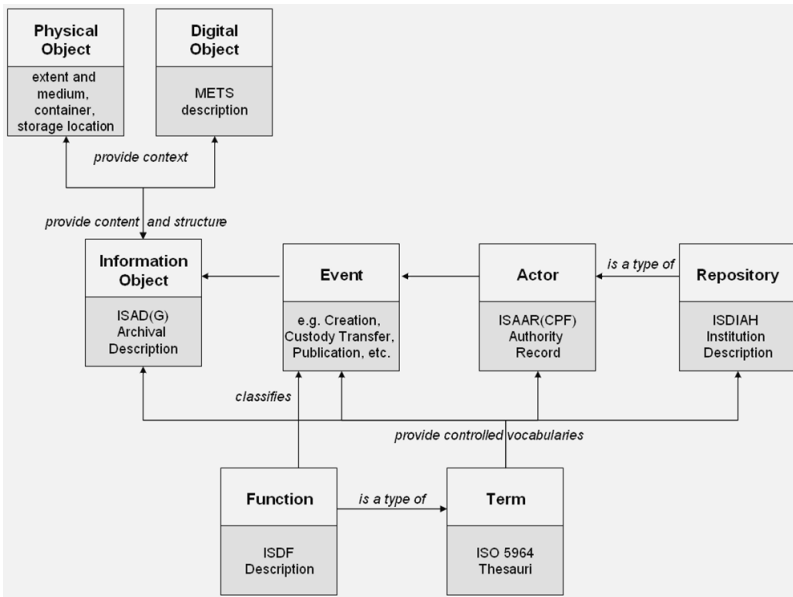
14) ActiveState의 2010년 보고서에 의하면 커스터마이징 등 새로운 기능 개발이 요구될 경우 12만 달러 정도의 비용이 소요된다. 이 금액은 숙련 개발자 1명의 1년치 임금 10만 달러와 상주 오픈소스 개발인력의 1년치 임금 2만 달러를 기준으로 산정한 것이다.

15) "The True Cost of Open Source Software Uncovering Hidden Costs and Maximizing ROI", ActiveState White Paper, February 2010.

16) Domain model - Qubit Toolkit, (https://www.qubit-toolkit.org/wiki/Domain_model), [인용 날짜: 2015. 4. 13].

와 ISAAR(CPF)의 약한 연결관계를 보장해 주는 역할을 한다. 기록학에 서는 기록을 정의할 때 업무 프로세스나 행위(acts), 활동(transactions)을 강조하므로 이벤트 엔티티가 개념적으로 이를 보충하게 된다. 이벤트 엔티티는 호주의 메타데이터 모델인 SPIRT의 ‘Business’ 엔티티와 유사하며¹⁷⁾, 장기보존 메타데이터 표준인 PREMIS에도 존재한다. ‘Business’ 엔티티가 업무 간의 맥락을 보장해 준다면, PREMIS의 이벤트는 법적, 절차적 맥락을 제공해 준다.

〈그림 1〉 AtoM 엔티티 관계 모델



* 출처 : Peter Van Garderen, p. 33.

17) Conceptual and Relationship Models: Records in Business and Socio-legal Contexts, <http://www.infotech.monash.edu.au/research/groups/rcrg/projects/spirt/deliverables/conrelmod.html>, <http://www.infotech.monash.edu.au/research/groups/rcrg/projects/spirt/deliverables/conrelmod.html>, [인용 날짜: 2015. 4. 13].

(2) AtoM 기능

AtoM은 기록소장기관이 기록을 관리하고 서비스하는 데 필요한 대부분의 기능을 제공한다. AtoM 2.1버전의 기술문서를 기준으로 사용자관리 및 접근권한 설정, 등록과 편집, 검색과 열람, 리포팅, 가져오기와 내보내기, 관리 등 6가지 영역으로 크게 구분할 수 있다. 감사증적, 처분, 접근권한 관리 등 일부 기능은 아직까지 기록관리시스템 기능요건 표준들이 요구하는 수준을 만족하지 못하고 있다. 이는 대규모 이관이나 주기적인 처분이 발생하는 공공기관보다 메뉴스크립트 기록관에 더 적합함을 의미한다. AtoM은 웹을 통한 기록물 기술과 접근 제공에 특화되어 있다. 바꿔 말하면 아키비스트가 쉽게 기록물을 기술할 수 있고, 이용자는 기록을 어려움 없이 이용할 수 있도록 하는 데 집중하고 있다. 이는 AtoM의 데이터 모델을 기반으로 가능해진다. ICA 기술표준에 따른 기록물 기술과 데이터 마이그레이션을 제공하고 각 기술표준이 상호 참조되어 아키비스트나 이용자 모두에게 아주 높은 수준의 기록물 이용 및 관리환경을 제공한다.

AtoM은 간편하고 유연하며 효율적인 시스템이다. 분류체계나 텍소노미, 서가정보 등을 쉽게 만들거나 변형할 수 있고, 여러 아카이브를 묶어 아카이브 포털을 만들 수도 있다. 엄격한 기술규칙을 적용하여 애를 먹게 하기보다는 아키비스트가 유연하게 기록을 설명할 수 있도록 해 준다. AtoM으로 기술한 결과물의 가장 큰 장점은 각 기술표준의 특정 필드가 다중 엔티티 속성을 가지고 있다는 점이다. 예를 들어 ISAD(G)의 생산자(Creator) 필드는 ISAAR(CPF)의 이름(Authorized form of name)과 연결되어 이용자가 기록정보와 출처정보를 오가며 소장 컬렉션을 편리하게 브라우징할 수 있도록 해 준다. 또한 웹페이지 개발에 많이 사용되는 PHP와 CSS를 이용하여

쉽게 디자인을 변경할 수도 있다. 기관의 홈페이지가 없을 경우 AtoM 사이트를 홈페이지로 이용하며 기록물 열람 및 검색 서비스를 제공할 수 있다.

3. AtoM 개발사례 분석

AtoM 개발 사례를 광범위한 커스터마이징, 소규모 커스터마이징, 데이터 마이그레이션 등 프로젝트 규모별로 구분하여 살펴보았다. 각 사례는 AtoM 기반의 아카이브 시스템을 구축할 때 수요기관의 AtoM 활용 수준에 따라 시사점을 제공해 준다.

1) AtoM 개발사례 및 특징

(1) 광범위한 기능개발 - BCAUL 프로젝트¹⁸⁾

브리티시컬럼비아기록관협회(AABC, Archives Association of British Columbia)는 브리티시컬럼비아 주의 기록관 종합목록(BCAUL, British Columbia Archival Union List)¹⁹⁾ 시스템을 AtoM으로 교체하고 180개 소속기관 소장목록의 온라인 접근을 제공하고자 이 프로젝트를 수행하였다. 이 프로젝트는 2008년 8월부터 2009년 3월까지 AtoM 개발사인 아티팩추얼의 주도로 5개 기관이 파트너로 참여하여 진행되었다.

18) BCAUL Pilot Project-Artefactual, <https://wiki.artefactual.com/wiki/BCAUL_Pilot_Project>, [인용 날짜: 2015. 4. 13].

19) Gateway to British Columbia's Past - MemoryBC, <<http://www.memorybc.ca/>>, [인용 날짜: 2015. 4. 13].

BCAUL 프로젝트에서는 캐나다 기록물기술표준인 RAD를 AtoM의 ISAD(G)로 매핑하기 위한 템플릿 모형의 작성부터 EAD, MARC21 등 현재 AtoM에 내장되어 있는 표준들의 호환을 위한 크로스워크와 매핑, 안정적인 벌크 임포팅을 위한 성능개선 작업이 이루어졌다. BCAUL 기술 데이터의 33개 필드 중 27개가 RAD로, 기관목록의 14개 필드가 ISDIAH로 마이그레이션되었다. 이러한 매핑과 조정 과정에서는 기존 기술항목의 작성규칙이나 표현방법 등을 AtoM에 맞게 조정하는 작업이 요구된다. 예를 들어 6자리로 이루어진 BCAUL의 기록물 통제번호는 AtoM의 'description_identifier' 필드로 매핑하고 AtoM에서 자동 부여하는 새로운 식별자를 병행하여 사용하였다. 기관목록의 주소 표기방식 또한 AtoM 표기방식과 달라 거리 주소만을 추출하였고, 소장기관 유형을 구분하는 방식도 대학과 교육기관을 통합하였다.

BCAUL 목록의 서비스 사이트인 메모리비씨(MemoryBC)²⁰⁾의 디자인 변경과 이용자 편의를 위한 커스터마이징 작업 또한 병행되었다. 기본 기술 템플릿을 RAD로 바꾸고, 필수 작성 필드가 잘 보이도록 붉은 색으로 처리하고, RAD 템플릿에 온라인 참고자료 링크를 추가하였다. 또한 기본언어를 English로, 국가를 캐나다로 변경하고, 온라인 검색도구 링크가 가능하도록 처리했다. 이후 이용자 도움말, 매뉴얼, 교육자료 등을 새로 작성하거나 수정하는 작업이 이루어졌다. 기술 등록 화면에서 전거통제를 위해 기존 전거레코드를 참조하라는 팁을 제공하거나, 전거레코드의 날짜가 기록 생산일자와 구분되도록 수정하는 등 관리자나 이용자 모두가 사이트를 쉽게 이용할 수 있도록 세세한 가이드가 구축되었다.

또한 AtoM사이트 이용현황을 지속적으로 모니터링하기 위해 구글 분석도구 계정을 생성하고 사이트와 연결시켰다. 이러한 작업은 서

20) MemoryBC, <<http://www.memorybc.ca/>>, [인용 날짜: 2015. 4. 13].

비스 품질을 유지하고 향상시키는 데 아주 중요하다. 사이트 기본작업이 마무리되는 시점에서는 디지털 객체 사본을 생성하고, 각 기관이 직접 업로드하거나 해당 기관의 서버에 업로드한 후 링크하도록 하였다.

이러한 마이그레이션과 커스터마이징 과정을 거쳐 메모리비씨 사이트가 소프트 런치되었고, 2009년 7월부터 11월까지 지속적인 테스트와 기능 개선이 이루어졌다. 스테이징 서버에서 프로덕션 서버로 옮기고, EAD 목록을 가져오고, 새롭게 추가되거나 벌크 임포트된 콘텐츠를 특정 테이블에서 보여주는 기능을 추가하고, 검색결과 필터링 규칙, 디지털 객체 링크 및 업로드 규칙, 전거래코드 공유규칙, 복수기관 풍 입력규칙 등 정책과 가이드라인을 수립하였다. 이러한 과정을 거쳐 2009년 11월 메모리비씨 사이트가 공식 오픈되었다. 이 파일럿 프로젝트의 결과물은 AtoM 소스코드에 대부분 반영되었다.

(2) 소규모 커스터마이징-아카이브즈캐나다(ARCHIVESCANADA.ca)

아카이브즈캐나다라는 캐나다의 아카이브 포털로 800개 이상의 기관이 소장목록을 제공하고 있다. LAC와 CCA는 2012년 3월을 목표로 2007년부터 아카이브즈캐나다 사이트를 AtoM으로 업그레이드하기 위한 테스트를 진행해 왔다. 이 연구는 복수 기록소장기관을 지원하고, RAD 기술을 제공해야 하고, 기존 데이터베이스와 높은 호환성을 제공해야 한다는 데 중점을 두었다. 본고에서는 이 프로젝트의 사용자 인터페이스 개선 부분을 주로 조사하였다.

사용자 인터페이스 개선의 목적은 주로 편리한 기록 검색을 제공하기 위한 것이었다. 이를 위해 확장검색, 퍼시트 검색(Faceted Search), 포털 검색 등으로 구분하여 여러 사이트를 벤치마킹하였다. 우선 확장검색

은 여러 종류의 필터를 이용하여 검색결과를 좁히는 기능과 특정 필드만을 선택하여 검색하는 기능을 중점적으로 조사하였다. 퍼싱 검색은 다차원의 네비게이션 도구를 통해 검색결과를 좁히거나 체크박스, 링크 등을 통해 검색결과를 얻는 기능을 중심으로 벤치마킹하였다. 스미소니언 인스티튜션(Smithsonian Institution) 컬렉션 검색센터가 주로 참고되었다. 포털 검색은 검색엔진, 디렉토리, 데이터베이스, 지도, 포럼 등 여러 유형의 정보를 통일된 방식으로 제공하거나 여러 어플리케이션이나 데이터베이스가 구동되어도 일관된 룩앤필(Look and feel)을 제공하는 것을 기준으로 삼았다. 유로피어나 포털이 주로 참고되었다. 조사 결과 퍼싱 검색이 가능하려면 AtoM이 제공하는 주제, 장소, 인명, 파일 유형 등 접근점 소스 데이터의 인텔싱이 필요했다. 풍과 컬렉션 레벨을 함께 보여주는 것과 아이템 계층 기술 화면의 제한을 없애는 것 등이 지적되었다.

벤치마킹 결과 도출한 시사점을 바탕으로 AtoM 사이트의 목업을 제작하였다. 메인페이지, 확장검색, 퍼싱 검색, 모바일 인터페이스 등 11개의 목업을 통해 사용자 인터페이스의 기본 방향을 설계하였다. 와이어프레임은 목업을 한 단계 발전시킨 것으로 웹디자인을 위한 기본자료가 된다. 아카이브즈캐나다에서 2012년 작성한 검색 및 브라우징 관련 와이어프레임은 'Trillium' 테마로 ICA-AtoM 1.2버전과 함께 출시되었다. 이 와이어프레임은 아카이브즈캐나다 사이트의 콘텐츠와 기능, 네비게이션에 대한 요구사항을 적용한 시각적 가이드이다. 메인화면과 아이콘 등을 그린 사용자 인터페이스 디자인과 검색/열람 등 핵심기능을 그린 와이어프레임, 모바일용 와이어프레임 등을 통해 요구사항에 맞는 사이트를 완성할 수 있게 된다.

〈그림 2〉 아즈캐나다 목업 작성사례 - 퍼싱 검색



* 출처 : https://wiki.artefactual.com/mediawiki/images/thumb/6/6d/AC_FS02.png/732px-AC_FS02.png

(3) 데이터 마이그레이션 - AABC Data Migration Toolkit²¹⁾

브리티시컬럼비아기록관협회와 7개 소속기관은 2010년 11월부터 2011년 4월까지 각 기관 데이터베이스 시스템을 ICA-AtOM으로 마이그레이션하고자 이 프로젝트를 수행하였다. 마이그레이션을 위해 드루팔 기반 툴, 구글 리파인(Google Refine), 펜타호(Pentaho), CSV 등의 툴을

21) https://wiki.artefactual.com/wiki/AABC_Data_Migration_Toolkit.

조사하여 적합성 평가를 한 뒤 펜타호와 CSV를 선정하였고, 각 기관이 틀을 선택하여 마이그레이션을 진행하였다. 펜타호를 이용한 작업과정은 데이터 추출, 매핑, 로딩, 스크립트의 과정을 거치며 EAD를 이용한다. CSV를 이용할 경우 데이터 추출, 검증, 변환의 단계를 거친다. 벤쿠버시 아카이브, UNBC, 사이먼프레이저대학교(Simon Fraser University), 성공회 아카이브(Anglican Archives), 빅토리아주립대학교(University of Victoria)는 기존 기록관리시스템, 데이터베이스 등 다양한 이기종 시스템으로부터 AtoM시스템으로 디스크립션과 디지털 객체를 마이그레이션하는 데 성공했다. AtoM의 데이터 구조로 매핑하는 과정에서 각 필드가 통합, 분산되거나 일부 누락시키는 등의 작업이 이루어졌다.

〈그림 3〉 CSV 템플릿을 이용한 데이터 매핑

ISAD(G)	ISAD	ICA-AtoM / Quilt	csv (ISAD(G) Template)	csv RAD Template
IDENTIFY AREA identifier / reference code*	TITLE & STATEMENT OF RESPONSIBILITY AREA	identifier*	identifier	
level of description*	level of description	level of description/arrangement*	level of description	level of description
title*	title proper / general materials designation	parent level* title proper*	title	title
		general materials designation		general materials designation
	parallel titles / other title information / statement of responsibility	title notes (drop down list: attributions and conjectures, continuation of title, parallel titles and other title information, source of title proper, statements of responsibility, variations in title)		other title information / parallel titles / statement of responsibility
extent and medium*	physical description area extent of descriptive unit	physical description / extent and medium	extent and medium	extent and medium
repository	repository option	repository	repository	repository option
CONTEXT AREA				
name of creator*/ start*		creator start year and year	name of creators start date (creation) and date (creation)	name of creators start date (creation) and date (creation)
end*/ dates of creation*		date display type of creation	dates of creation	dates of creation
type*		place		
	administrative history/ biographical sketch	administrative history/biographical sketch		administrative history/biographical sketch
archival history	custodial history	custodial history	archival history	archival history
immediate source of acquisition	immediate source of acquisition	immediate source of acquisition	immediate source of acquisition and transfer	immediate source of acquisition and transfer
CONTENT AND STRUCTURE AREA				
scope and content	scope and content	scope and content*	scope and content	scope and content
appraisal, destruction, scheduling		appraisal, destruction, scheduling	appraisal, destruction, scheduling	appraisal, destruction, scheduling
Accession 10.3.31 system of arrangement	Accession level (is notes) arrangement	Accession arrangement	Accession system of arrangement	Accession system of arrangement
	CLASS OF MATERIAL SPECIFIC DETAILS AREA			
	cartographic materials – statement of scale	statement of scale		statement of scale
	cartographic materials – statement of projection	statement of projection		statement of projection
	cartographic materials – statement of coordinates	statement of coordinates		statement of coordinates
	architectural and technical drawings – statement of scale	statement of scale [architectural drawings]		statement of scale [architectural drawings]
	philatelic records – denomination/ jurisdiction	issuing jurisdiction/denomination [philatelic records]		issuing jurisdiction/denomination [philatelic records]
		CONTROL AREA		
		description identifier	description identifier	description identifier
		institution identifier	institution identifier	institution identifier
		rules or conventions	rules or conventions	rules or conventions
		status	status	status
		level of detail	level of detail	level of detail
		language of description	language of description	language of description
		scope of description	scope of description	scope of description
		dates of creation, revision, deletion	dates of creation, revision, deletion	dates of creation, revision, deletion
		restrictions on access	restrictions on access	restrictions on access
		sources	sources	sources
		archivist's notes	archivist's notes	archivist's notes
CONDITIONS OF ACCESS AND USE AREA				

* 출처 : <https://wiki.artefactual.com/mediawiki/images/8/89/Crosswalkcapture.png>

2) 시사점

우선 첫째, 마이그레이션을 성공시키는 것이 중요하다. 위 사례들은 기존 데이터베이스에서 AtoM 시스템으로 데이터를 마이그레이션하는 과정에 가장 많은 노력을 들이고 있다. 기록물의 양이 많을 수록 메타데이터 매핑과 데이터 정제, EAD, CSV 등 벌크임포팅하는 방식의 선정과 이를 수행하는 방식을 두루 고민해야 한다. BCAUL 프로젝트의 경우 데이터 마이그레이션 테스트에만 4만 달러의 예산을 책정하였다. AtoM 시스템으로의 안전한 이관을 위해서는 충분한 기간 동안 테스트가 요구된다.

둘째, 섬세한 커스터마이징이 필요하다. 기관에 맞는 메뉴구성과 용어 번역, 도움말 제공 등은 적은 노력으로 큰 효과를 제공한다. 특히 핵심 기능인 기록 검색과 열람 기능에 중점을 두고 기능 개선이 이루어질 필요가 있다. OSS 기반 시스템 개발이라 하더라도 대대적인 기능 개선을 위해서는 엄청난 비용이 발생한다. AtoM이 기본적으로 제공하는 시스템상의 커스터마이징 요소들을 최대한 활용할 필요가 있다.

셋째, 사용자 인터페이스와 디자인 개선이다. AtoM은 2.1 버전부터 모바일 로그인을 제공하며 거의 완벽한 CSS 기반의 레이아웃을 제공한다. 대부분의 AtoM 사이트들이 디자인과 인터페이스 개선을 위해 노력하고 있다. 타겟 이용자를 대상으로 한 AAO의 AtoM 사용성 평가²²⁾와 기존의 AtoM 프로젝트 벤치마킹을 통해 목업과 와이어프레임을 제작해 보고 트위터 부트스트랩 등 CSS 프레임워크 도구를 이용하여 보다 쉽게 디자인을 개선할 수 있다.

22) <https://wiki.artefactual.com/File:AAO_UI_Features.pdf>, [인용 날짜: 2015. 4. 14].

4. AtoM 시스템의 구축과 커스터마이징

OSS 관련 연구들은 성공적인 OSS프로젝트인 리눅스, 모질라 등의 사례 분석을 통해 아래와 같은 OSS 개발방법론을 제시하고 있다. Huang Meng(2005)²³⁾은 OSS 기반 어플리케이션을 개발하기 위한 재사용 프로세스를 후보 OSS 조사부터 제품 완성까지 8단계의 활동으로 제시하고 있다. 아래의 방법들이 전통적인 OSS 개발방법론과 구분되는 점은 새로운 기능을 결합하기 위해 기존 OSS를 활용하거나 저장소 분기를 통해 새로운 프로젝트로 이어진다는 것이다. 이러한 특성은 AtoM 시스템이 제공하지 않는 기능을 추가 개발할 때의 방법론으로 유용할 것이다.

〈표 1〉 OSS 개발 프로세스 비교

개발단계	Stephanos ²⁴⁾	Mozilla 개발 ²⁵⁾	이민석 ²⁶⁾	김종배 외 ²⁷⁾
분석	요구사항 정의	코드 리뷰 -모듈 소유자 리뷰 -슈퍼 리뷰	요구분석 및 기존SW 검색	계획수립 1)OSS전략 2)타당성 검토 3)활용수준 분석 4)소요자원 산정 5)분리계획 수립 6)라이선스 정책수립

23) Huang, Meng, Liguang Yang and Ye Yang, "A development process for building OSS-based applications", *In Unifying the Software Process Spectrum*, Springer, 2006, p.125.

24) Androutsellis-Theotokis, Stephanos, Diomidis Spinellis, Maria Kechagia, Georgios Gousios, Sergei Evdokimov, Benjamin Fabian, Oliver Günther, Lenka Ivantysynova, Holger Ziekow & Michael A. Lapré, "Open source software: A survey from 10,000 feet", *Foundations and Trends in Technology, Information and Operations Management* Vol.4, 2011, pp.187-347.

25) 윤석찬, 「Mozilla를 통해 본 공개 S/W 개발 프로세스」, 『정보과학회지』 26(7), 2008, 44-51쪽.

설계 구현 테스트 품질관리	새로운 기능 결합	저장소 분기	활용방식 설정 1)가치치기 2)신규프로젝트 시작	활용대상 선정 1)후보 OSS식별 2)적합성 평가 3)작업량 산정 4)대상 선정
	코드 통합	알파버전 출시(3~5회) -테스트 및 버그수정	프로토타입 구현	획득 및 변경
	릴리스 관리	베타버전 출시(3~5회) -테스트 및 버그수정	배포	적용 및 검증
	OSS 프로젝트 평가	출시 후보판 품질검증 정식 출시	개발자간 소통 커뮤니티 기반 개발	릴리스 및 사후관리

이 연구의 목표는 AtoM 기반의 아카이브시스템 구축과 안정적 운영을 위한 방법론을 제시하는 것이다. 따라서 2장과 3장의 분석결과와 OSS 개발방법론 연구사례, 한국정보화진흥원의 CBD SW표준산출물관리 가이드를 바탕으로 AtoM 시스템 개발 과정을 요구분석, 적합성 평가, 설치, 마이그레이션, 커스터마이징, 추가기능 개발, 문서화, 기여의 8단계로 구분하였다. 또한 각 단계별 산출물을 제시하고 작성 방법을 설명하였다. 본고에서는 요구분석, 적합성 평가, 마이그레이션, 커스터마이징 과정에 대해 알아보았다. 설치 과정은 오픈소스 기록관리 포럼 (<http://osasf.net>)을 통해 알아볼 수 있다. 추가기능 개발 및 문서화, 기여 등은 후속 연구가 필요하다.

- 26) 이민석, 「공개 소스 소프트웨어 프로젝트의 생명 주기와 품질 유지 방안」, 『정보과학회지』 26(7), 2008, 11-21쪽.
- 27) 김종배, 송재영, 류성열, 「오픈소스 기반의 소프트웨어 개발 모델 연구」, 『한국디지탈콘텐츠학회논문지』 6(4), 2005, 229-234쪽.

〈표 2〉 AtoM 시스템 구축 프로세스 및 산출물(안)

개발단계	수행작업		산출물
요구분석	요구사항 분석	컬렉션 분석	사용자 요구사항 정의서
		기능 요구사항 정의	
		메타데이터 요구사항 정의	
		유스케이스 작성	유스케이스 명세서 유스케이스 다이어그램
적합성 평가	OSS 선별	후보 OSS 조사	OSS 개요서
		적합성 평가	OSS 적합성 평가서
		활용수준 설정	
설치	AtoM 소프트웨어 설치	하드웨어, 서버, 소프트웨어 준비	
		LAMP 환경 설정	
		AtoM 설치	시스템 명세서
마이그레이션	데이터 마이그레이션	데이터 추출(CSV)	CSV 목록
		AtoM 템플릿 매핑	컬럼 매핑 표(CSV)
		CSV 벌크 임포트	마이그레이션 결과보고서
커스터마이징	커스터마이징	기능 개선	소스코드
		UI, 디자인	CSS 파일 사용자 인터페이스 설계서
추가기능 개발	시스템 구조 설계	시스템 구조 설계	아키텍처 설계서
	프로그램 구조 설계	프로그램 구조 설계	클래스 설계서
	UI 설계	사용자 인터페이스 설계서 작성	사용자 인터페이스 설계서
		UI목업, UI와이어프레임 작성	UI 목업, UI 와이어프레임
문서화	매뉴얼 작성	매뉴얼, 튜토리얼, 지침 작성	관리자/이용자 매뉴얼
			튜토리얼
			등록지침
	기술 설계	ISAD(G) 등 기술설계	기술 설계서
기여	AtoM 커뮤니티 참여	AtoM 버그 리포트 및 소스코드 공유	버그리포트, 소스코드

1) 요구분석

(1) 컬렉션 분석

아카이브 시스템 구축을 위해 가장 먼저 해야 할 일은 소장 기록물의 양과 유형, 내용 등을 분석하는 것이다. 일반적으로 컬렉션 분석은 수집정책 작성 및 평가를 위해 수행되지만 본고에서는 AtoM 시스템을 통한 관리 및 서비스 방식을 결정하기 위해 소장 기록물의 양과 계층구조, 매체 유형, 내용 등을 분석하는 것으로 한정한다. 컬렉션 분석을 통해 소장 기록물의 관리와 서비스를 위한 큰 틀의 요구사항을 도출할 수 있다.

우선 첫째, 수량 및 구조 분석을 수행해야 한다. 수량 및 구조 분석은 전체 소장물의 양과 기록 계층구조를 파악하기 위한 것이다. 이를 통해 필요한 하드웨어나 소프트웨어 규모를 산정하고, AtoM 시스템이 제공하는 계층구조를 그대로 사용할 것인지, 추가하거나 변경할 것인지를 결정하게 된다. 전체 기록물 목록이 작성되면 그 중 AtoM 사이트에 등록하여 관리할 대상을 정하게 되고, 향후 생성될 디스크립션 수 등을 고려하여 데이터베이스의 규모와 소프트웨어 설정, 메모리 크기와 네트워크 속도 등 전반적인 시스템 인프라를 확정하게 된다. AtoM 시스템에서 기본적으로 제공하는 기록 계층은 품, 서브품, 컬렉션, 시리즈, 서브시리즈, 철, 파트, 건 등 8개이다.²⁸⁾ 목록작업을 통해 소장기관 등 출처 및 사용할 기록계층, 계층명을 확정하고 필요한 경우 ‘품’을 ‘레코드 그룹’으로 명칭 변경하거나 ‘컴포넌트’ 등의 새로운 기록계층을 추가해야 할지 여부를 결정할 수 있다. AtoM 2.x 버전을 사용하는 18개 AtoM 사이트를 조사한 결과 하나의 AtoM 사이트를 한 기관이 단독으로 사용하

28) <<https://www.accesstomemory.org/en/docs/2.1/user-manual/data-templates/isad-template/#level-of-description>>, [인용 날짜: 2015. 4. 13].

는 경우가 8개로 가장 많았으며 평균적으로 한 AtoM 사이트당 22개 소장기관이 등록되어 있었다. AtoM 시스템 개발 당시 복수 소장기관을 지원할 수 있어야 한다는 요구조건에 맞게 동일한 지역이나 주제분야에 속해 있는 복수의 기관이 AtoM 사이트 포털을 구축한 경우가 많았다. 캐나다 온타리오 주의 기록 포털사이트인 아르케이온(Archeion)에는 171개 기관의 소장목록을 통합하여 제공하고 있다.

둘째, 매체 유형분석이 필요하다. 본 디지털 형태의 기록을 주로 보유하고 있는 기관은 AtoM 시스템에 디스크립션과 디지털 객체를 함께 등록하여 원본 열람 서비스를 제공하기가 수월할 것이다. 웹을 통한 원본 열람을 제공하는 경우 대용량 파일 전송이 요구되는 오디오나 비디오보다는 문서나 사진 등의 기록 유형이 적합하다. AtoM 시스템에 업로드 가능한 디지털 객체 1건의 최대 용량은 64MB 로 설정되어 있으며, PHP 설정파일의 실행시간과 메모리 설정을 수정하여 이를 변경할 수 있다. 하지만 이러한 방식은 AtoM이 설치된 시스템의 성능이나 가용 메모리에 따라 변화하므로 정확한 업로드 용량을 확정하기는 어렵다.²⁹⁾ 디지털 객체의 해상도(픽셀 수)에 따라 필요한 입출력 메모리 크기는 13메가픽셀당 100MB 수준임을 감안하여 시스템 메모리를 증설하거나 디지털 객체의 용량을 조정해야 한다.

〈표 3〉 디지털 객체 해상도별 입출력 메모리

메모리 크기(MB)	디지털 객체 해상도(Megapixels)
100	13
128	18
256	35
512	70

* 출처 : <https://www.accesstomemory.org/en/docs/2.1/admin-manual/installation/execution-limits/>

29) 〈<https://www.accesstomemory.org/en/docs/2.1/admin-manual/installation/execution-limits/>〉, [인용 날짜: 2015. 4. 13].

여러 기관이 하나의 AtoM 포털사이트를 구축한 경우 디지털 객체를 등록하지 않고 각 기관의 소장목록만을 제공하기도 한다. 이러한 AtoM 사이트는 특정 지역이나 주제의 연구자들에게 접근점을 제공하기 위한 목적으로 구축되었다. 따라서 실제 기록물 열람 서비스는 해당 기록관 사이트나 오프라인을 통해 제공한다. 뉴펀들랜드와 래브라도 아카이브의 종합목록(ARC: the Archival Resource Catalogue of the Association of Newfoundland and Labrador Archives) 사이트에서는 70개 소장기관의 52개 컬렉션 목록만을 제공하고 있다.

매체 유형분석을 통해 원문 열람 서비스 제공 범위를 설정할 수 있다. 본 디지털(Born digital) 문서와 사진은 원문 열람을 제공하고 실물기록 중 일부를 디지털화하여 AtoM 사이트 구축 시 제공하는 등의 전략을 설정할 필요가 있다. 인간과기억아카이브의 경우 본 디지털형태의 문서와 디지털 이미지 1,337건을 가장 먼저 등록하였고, 5월 12일 일기 등 종이 기록을 스캔한 이미지 1,471건을 추가로 등록하였다. 알버타 온레코드의 경우 전체 17만 2천여 건의 90퍼센트에 해당하는 15만 4천여 건의 기록을 디지털 객체와 함께 제공하고 있다.

셋째, 내용 분석을 통해 대표 컬렉션을 선별하고 이에 대한 기술을 보강하거나 메인화면에 노출하는 등의 서비스 전략을 설정할 수 있다. 이 과정에는 AtoM 시스템을 통해 일반 이용자들에게 공개할 컬렉션과 관리자나 연구자만 이용 가능한 컬렉션을 구분하는 작업이 포함된다. 내용 분석 과정은 전통적 컬렉션 분석의 두 번째 단계인 질적 평가에 해당한다. 해당 컬렉션의 수집 완결성, 향후 수집 가능성, 타 보존소에서 해당 영역이 얼마나 잘 도큐멘테이션되고 있는지의 정도, 내용적 중요성, 서비스 적합성 등의 요소를 평가할 수 있을 것이다.³⁰⁾ 인간과기

30) F. Gerald Ham, *Selecting and Appraising Archives and Manuscripts*, 강경무, 김상민 역, 『아카이브와 매뉴스크립트의 선별과 평가(SAA 기록학 기초시리즈)』, 2002, 서울: 진리탐구.

역아카이브는 전체 8개 컬렉션 중 '5월 12일 일기수집', '개인기증', '고대 검우회', '명지대' 등 4개 컬렉션을 일반 이용자들에게 공개하고 있다. '5월 12일 일기수집' 컬렉션은 일상 영역의 기록화를 목표로 하고 있는 기관의 사명과 가장 잘 부합하므로 이용자가 '소장기록물 정보'를 클릭했을 때 가장 상단에 배치하였다. 그리고 각 컬렉션의 이해를 돕도록 이미지를 링크하였다.

이처럼 컬렉션 분석은 AtoM 사이트를 기획할 때 가장 먼저 수행되어 서비스 요건을 수립하기 위한 기본 정보를 제공해야 한다.

(2) 기능 요구분석

AtoM 시스템을 구축하려는 기관은 가장 먼저 기록관리시스템이 반드시 제공해야 할 기능을 정의하는 요구분석 단계를 거치게 된다. AtoM 기반의 아카이브 시스템 구축을 위한 요구분석 단계에서는 신규 기능을 추가하거나, 기존 기능을 개선할 경우에만 요구사항 정의서를 작성하는 방식도 효율적일 것이다. 그러기 위해서는 현재 버전의 AtoM 시스템이 제공하는 기능을 충실히 파악하여 기능 목록을 작성하는 작업이 반드시 선행되어야 한다.

AtoM 시스템 개발 초기인 2008년에 작성된 기능 요구사항은 통제시스템, 콘텐츠 추가/삭제, 번역, 접근, 가져오기/내보내기, 시스템 관리 등 6개 상위 기능요건으로 이루어져 있다.³¹⁾ AtoM 프로젝트 관리 페이지를 참고하여 이미 개발되었거나 현재 개발 중인 기능 목록과 완료일정을 파악하고 기능요건을 정리해 볼 수 있다.³²⁾ 또한 최신 버전의 이

31) <https://www.ica-atom.org/doc/Functional_requirements>, [인용 날짜: 2015. 4. 13].

32) <https://projects.artefactual.com/projects/atom/issues?set_filter=1&tracker_id=2>, [인용 날짜: 2015. 4. 13].

용자 매뉴얼이나 AtoM 이용자 포럼³³⁾, 개발자 포럼³⁴⁾을 참고하여 문제점이나 주요 이슈들을 파악할 수 있다.³⁵⁾ AtoM의 데이터 모델을 제공하는 큐빗의 기능 요구사항을 참고한다면 가능한 개발 범위를 설정하는데 도움을 얻을 수 있다.³⁶⁾ 무엇보다 AtoM을 직접 설치하거나 데모 사이트에 접속하여 제공 기능을 충분히 테스트하는 과정이 필요하다. 이후 상위 기능요건을 구분한 뒤 그에 따른 중기능, 소기능으로 상세화하여 기능요건표를 작성해야 한다.

요구사항 정의서는 요구사항 ID(FR-x.x.x), 요구사항명, 요구사항 설명 등의 요소로 구성된다. 상위 기능요건과 하위 기능요건을 계층적으로 둘 수 있으며, 관련된 표준, 관련 메타데이터 요구사항, 해당 기능과 관련한 문제나 제약사항 등을 포함하면 상세한 기능요건이 완성된다.

〈표 4〉는 한국정보화진흥원의 ‘CBD SW표준 산출물 관리 가이드’의 서식 중 ‘사용자 요구사항 정의서’ 양식을 이용하여 AtoM 추가기능 개발을 위한 기능 요구사항을 작성한 것이다. 기록에 대한 디스크립션을 입력할 때 연관된 기록을 선택하여 서로 링크되도록 하는 경우를 상정하였다. 요구사항에 대한 설명과 제약사항, 중요도, 해결 방안, 검수기준 등의 항목을 포함하여 해당 요건을 명확하게 기술해야 한다.

33) <<https://groups.google.com/forum/?hl=ko#!forum/ica-atom-users>>, [인용 날짜: 2015. 4. 13].

34) <<https://groups.google.com/forum/?hl=ko#!forum/qubit-dev>>, [인용 날짜: 2015. 4. 13].

35) <<https://www.accesstomemory.org/ko/docs/2,1/contents>>, [인용 날짜: 2015. 4. 13].

36) <https://www.qubit-toolkit.org/wiki/Functional_Requirements#Menu_26_Navigation>, [인용 날짜: 2015. 4. 13].

〈표 4〉 사용자 요구사항 정의서 작성 사례

요구사항 ID	요구사항명	구분	요구사항 설명	요구사항 출처	제약사항	중요도	해결 방안	검수 기준	비고
FR-2.5.5	관련 디스크립션 링크 제공	기능, 기능 개선	ISAD(G) 템플릿 입력 시 관련자료영역(Allied materials area)의 'Related descriptions' 필드에 특정 디스크립션의 링크 제공. 드롭다운메뉴로 기술항목을 검색하여 선택할 수 있도록 해야 함	Feature #5435 https://projects.artefactual.com/issues/5435	이미 입력된 기록이어야 한다 해당 기록의 URL이 바뀌면 자동으로 링크도 변경되어야 한다	중	필드 자체의 속성 변경과 함께 해당 기록의 제목을 URL로 변경하는 기능 추가	드롭다운 메뉴 통합 입력 제공 및 링크 정상작동 100회	버그확인 필요 Bug #7154 https://projects.artefactual.com/issues/7154

(3) 메타데이터 요구분석

AtoM 시스템은 웹 기반의 기록물 기술이 가능한 간편한 프로그램을 목표로 개발되었다. 따라서 기록을 효율적으로 기술하고 열람, 검색을 제공하기 위해 메타데이터에 대한 요구분석이 상당히 중요해진다. AtoM 시스템을 개발할 때에도 다양한 메타데이터 표준의 반영은 가장 중요한 요소 중 하나였으며 ICA 기술표준 4종 뿐만 아니라 DACS, RAD 등 국가표준과 더블린코어, MODS 등 다양한 메타데이터가 적용되었다.

메타데이터 요구사항을 도출할 때 고려해야 할 사항으로는 해당 기관의 기록, 출처, 소장기관, 기능 등을 기술하기 위해 AtoM에 내장된 ISAD(G), ISAAR(CPF), ISDIAH, ISDF 등이 적합한지, 또는 PREMIS나 InterPARES 등 장기보존을 위한 메타데이터 요소가 요구되는지 등을 결정하는 것이다.

AtoM에 내장된 ICA 기술표준 4종은 기록물의 보존기간 설정, 폐기 등 처분 이벤트에 대한 필드가 상대적으로 부족하여 주기적인 평가 및 폐기가 이루어지는 기록관의 경우 이를 위한 별도의 메타데이터 추가를

고려할 필요가 있다. 또는 ISAD(G)의 특정 필드를 이벤트 기술을 위해 활용하도록 설계할 수 있다. AtoM 데이터 모델 설계 시 이벤트(Event) 엔티티는 기록 레코드의 생성, 관할 변경, 보존 처리, 디지털 객체의 파일포맷 변환, 접근 로그, 기술정보 수정 등 기록의 생애주기 전반에 걸친 문맥을 제공하는 역할로 설정되었으나 ISAD(G)나 ISAAR(CPF)와의 연결 요소가 거의 없는 실정이다. AtoM 베타 테스트 기간에 4종의 ICA 기술표준과 이벤트 엔티티를 어떻게 연결할지 논의되었으나, 결국 1.1 버전에 이벤트 엔티티가 완전히 숨겨졌으며, ISAD(G)의 기록 생산일자(Creation date) 항목으로만 제한되었다.

ICA 표준위원회(CBPS, Committee on Best Practices and Standards) 역시 보고서를 통해 기록(ISAD)과 행위자(Actors)의 약한 연결고리를 지적하고 있다. 위원회는 이벤트(Event) 엔티티가 기록과 행위자의 관계를 규정하는 확실한 개념 모델이 필요함을 강조하였다.³⁷⁾ 여기서 이벤트는 행위(acts), 트랜잭션(transactions), 업무 프로세스(business process) 등을 의미한다. 이러한 현재의 상황을 반영하여 이벤트 엔티티에 해당하는 메타데이터 템플릿을 별도의 플러그인으로 추가하는 것 또한 고려해 볼 만하다.

AtoM에서 제공하고 있는 메타데이터 기능 명세는 기록, 행위자, 기록 소장기관, 용어, 디스크립션 정보, 디스크립션 간의 관계, 이용자 등 7 가지 항목에 대한 것이다. AtoM 개발 초기에 AtoM시스템의 메타데이터 기능요건 목록 또한 작성되었다. 위에서 지적한 바와 같이 처분과 관련된 'Event' 엔티티를 보충하고자 한다면 최상위 요구사항에 이벤트 레코드를 추가해야 할 것이다.

37) <<http://www.ica.org/13155/standards/cbps-progress-report-for-revising-and-harmonising-ica-descriptive-standards.html>>, [인용 날짜: 2015. 4. 13].

2) 적합성 평가

적합성 평가는 기관의 요구사항에 가장 부합하는 OSS를 선별하는 과정이다. 기록관리 분야의 OSS를 조사하여 후보를 선정한 뒤 각 OSS의 품질이나 기관의 요구사항 부합성 등을 점수로 환산하게 된다. 이후 활용수준을 결정하게 된다. 적합성 평가는 시스템 도입을 위해 수행할 수도 있고, 일부 기능을 추가하기 위해 이루어질 수도 있다. 즉 어떤 단위에서건 OSS를 도입하고자 할 때 해당 제품의 도입여부를 판가름하는 지침으로 이용하게 되는 것이다. 만약 AtoM 소프트웨어의 기능이나 품질요건이 기관의 요구와 상이하거나 상당 부분 부합하지 못할 경우 AtoM이 아닌 다른 OSS의 도입을 고려해야 한다. 이 경우에도 동일한 방식으로 OSS 적합성 평가를 수행하게 된다.

(1) 후보 OSS 조사

기존 OSS를 조사하여 활용 가능성을 분석하는 과정은 OSS 기반 시스템 개발 과정의 가장 큰 특징이다. 밴쿠버 디지털 아카이브는 3개월에 걸쳐 후보 OSS를 선별하고 테스트하였다. 리눅스 우분투 등 운영체제부터 큐빗, 드로이드(DROID) 등 데이터모델과 파일포맷 레지스트리까지 다양한 규모와 수준에서 OSS를 조사하였다. OSS개요서에는 조사한 툴의 명칭과 링크, 테스트한 버전정보, 해당 툴에 대한 간략한 설명, 적용된 OSS 라이선스 유형 등을 기재하였다.

이러한 사례를 바탕으로 AtoM시스템 추가기능 개발을 위한 OSS 개요서에 작성할 항목을 아래와 같이 산출하였다. 링크가 포함된 OSS 명칭, 최신버전 정보, 소프트웨어의 용도 및 특징에 대한 설명, 적용된 OSS 라이선스 유형, 구동환경, 개발문서 등 도큐멘테이션 페이지 링크, 데모 사이트 링크, 관련 요구사항, 접근한 날짜 등으로 구성하였다. OSS 개요

서는 1~3개월의 기간을 두고 해당 제품의 특징 및 기능을 충분히 테스트하여 작성할 필요가 있다.

〈표 5〉 OSS 개요서 작성 사례

OSS 이름 (링크 포함)	버전	설명	라이선스	구동환경	도큐멘테이션 (링크)	데모사이트 (링크)	관련 요구사항	접근 날짜
Omeka	2.2.2	더블린코어 기반 기록 퍼블리싱 소프트웨어. 조지메이슨대학에서 개발하여 미국 내 커뮤니티 활성화됨 OAI-PMH, EAD, CSV 등 AtOM 디스크립션 호환기능 제공	GPL v3	Windows, Linux, Mac OS X	http://omeka.org/codex/Documentation	http://omeka.org/sandbox/	FR-4.5 기록 전시 템플릿 제공	2014.12.17

(2) 적합성 평가

적합성 평가는 OSS 개요서를 바탕으로 요구사항을 충족하는 OSS를 선별하는 과정이다. OSS를 선택하는 과정에서는 상위 수준의 사용자 요구사항 작성과 후보 OSS 프로젝트의 식별 작업이 반드시 선행되어야 한다. 요구사항 분석의 초기에는 사용자의 요구사항이 명확하지 않고 추상적인 수준인 경우가 많으므로 키워드 중심으로 조사한 뒤 상세한 기능 정의로 보완해 나가야 한다. 심사기준으로는 프로젝트의 목적과 요구사항의 만족 정도, 호환성, 지속적인 업그레이드 계획, 지원 받을 수 있는 개발 그룹과 사용자 층의 존재 유무, 문서화 정도, 오류의 정도, 적용 사례, 평판 등을 적용해 볼 수 있다.³⁸⁾ 기관의 요구사항이 아직 상세화되지 않은 단계라면, AtOM 기능명세서의 최상위 기능요건 등을 기준으로 후보 OSS 목록을 작성해 볼 수 있다. MoReq 등 기록시스템 관련 표준을 이용해도 좋을 것이다.

38) 김종배, 송계영, 류성열, 앞의 글, 2005, 232쪽.

〈표 6〉 OSS 적합성 평가서 - 1)상위 요구사항별 후보 OSS목록

요구사항 ID	상위 요구사항 모듈	후보 OSS
FR-1	지적 통제	ArchivesSpace DSpace
FR-2	콘텐츠 등록/수정	Omeka Drupal
FR-3	번역	Transifex
FR-4	열람/검색	Solr Elastic Search
FR-5	가져오기/내보내기	..
FR-6	시스템 관리	..

그 다음으로는 하위 요구사항의 각 항목에 대한 가중치를 부여하게 된다. 후보 OSS 목록 중 'FR-2 콘텐츠 등록/수정'의 하위 요구사항을 아래와 같이 나열한 뒤 가중치를 부여해 보았다. 하위 요구사항 또한 AtoM 사이트의 기능 요구사항을 기준으로 하였고, 기술정보 작성과 열람시 편의성 등에 가중치를 더 부여하였다. 가중치는 해당 기관의 정책과 전략에 따라 달라질 수 있다.

〈표 7〉 OSS 적합성 평가서 - 2)하위 요구사항별 가중치

No.	요구사항 ID	하위 요구사항	가중치
1	FR-2.1	기록 선별	10
2	FR-2.2	기록 입수	10
3	FR-2.3	기록 저장	10
4	FR-2.4	기록 보존	10
5	FR-2.5	기록 기술	15
6	FR-2.6	전자레코드 기술	15
7	FR-2.7	기록소장기관 기술	15
8	FR-2.8	기술정보 통제	10
9	FR-2.9	기록 접근점 할당	15
10	FR-2.10	접근권한 관리	10
11	FR-2.11	기록 가시성 확보	10

그 다음으로는 일반적인 OSS의 속성에 따른 평가기준별로 가중치를 부여하게 된다.³⁹⁾ 기능성 및 기능 정확도, 아키텍처 안정성 등의 순으로 가중치를 부여하는 것이 바람직하다. OSS 평가요소에는 해당 OSS가 조직의 특성과 OSS 사용경험, 관리예산 등에 비추어 적합한지가 반영되어야 한다. 특정 OSS를 도입한다는 것은 향후 해당 커뮤니티가 사라지거나 여러 라이선스의 중복으로 저작권 침해가 발생하는 등의 갖가지 리스크를 떠안아야 함을 의미하기 때문이다. 이러한 측면에서 고려해야 할 재사용의 기준으로는 조직의 OSS 경험과 성향을 고려한 사용 용이성, 기존 시스템과의 호환성, 관리비용, 적합한 개발도구 사용성 등이 있다.

〈표 8〉 OSS 적합성 평가서 - 3)OSS 평가요소 및 가중치

No.	OSS 평가요소	가중치
1	기능 정확도	150
2	인터페이스 유연성	80
3	이용가능성/강건함	100
4	설치/업그레이드 용이성	60
5	보안성	120
6	이동성	80
7	퍼포먼스	60
8	기능성	200
9	인터페이스 이해의 용이성	80
10	인터페이스 사용의 용이성	60
11	소프트웨어 성숙도	80
12	인터페이스 버전 호환성	100
13	내부 컴포넌트 간 호환성	80
14	교육 용이성	100
15	제품 획득 및 관리비용	80
16	개발자그룹 지원	60

39) Huang, Meng의 2005년 연구에 제시된 OSS 평가요소 20가지를 그대로 반영하였다. Huang, Meng, Liguang Yang and Ye Yang. "A development process for building OSS-based applications", *In Unifying the Software Process Spectrum*, 2006, pp.122-135.

17	아키텍처 안정성	150
18	개발도구 사용성	80
19	소스코드 합법성	60
20	소스코드 이해성 및 수정가능성	100
	합계	1,880

마지막으로 요구사항별 후보 OSS를 OSS 평가요소에 대입하여 20개 항목의 점수를 매긴다. 이 점수에 하위 요구사항별 가중치를 곱하여 전체 OSS에 대한 적합성 평가점수를 도출할 수 있다. 평가점수를 바탕으로 각 요구사항에 적합한 OSS를 선별한다.

〈표 9〉 OSS 적합성 평가서 - 4)OSS 적합성 평가표

No.	OSS 평가요소	후보 OSS 평가점수		
		Atom	Omeka	Drupal
1	기능 정확도	140	140	150
2	인터페이스 유연성	70	50	80
3	이용가능성/강건함	90	80	100
4	설치/업그레이드 용이성	50	50	30
5	보안성	100	90	110
6	이동성	50	50	50
7	퍼포먼스	50	50	60
8	기능성	120	120	190
9	인터페이스 이해의 용이성	70	70	30
10	인터페이스 사용의 용이성	60	60	30
11	소프트웨어 성숙도	70	70	50
12	인터페이스 버전 호환성	80	70	60
13	내부 컴포넌트 간 호환성	80	80	70
14	교육 용이성	100	100	50
15	제품 획득 및 관리비용	80	80	60
16	개발자그룹 지원	50	40	30
17	아키텍처 안정성	130	120	140
18	개발도구 사용성	70	60	50
19	소스코드 합법성	60	60	60
20	소스코드 이해성 및 수정가능성	90	80	70
	합계	1,610	1,520	1,470

(3) 활용수준 설정

OSS의 활용 수준을 계획하는 것은 프로젝트의 방향 수립과 프로세스에 중대한 영향을 준다.⁴⁰⁾ 대개 OSS를 그대로 적용하거나 재사용하는 경우로 나눌 수 있다. 예를 들어 문서 작성을 위해 전사적으로 오픈 오피스를 적용한 조직의 경우 내부 업무 기능의 향상만을 추구하고자 하기 때문에 OSS 제품을 그대로 사용하면 된다. 소스코드를 재사용하는 경우도 있다. OSS의 일부를 제품에 그대로 포함시키거나, 일부 수정하여 사용하는 경우이다. 재사용의 목적은 새로운 소프트웨어 개발의 양을 줄이고자 함에 있다.⁴¹⁾ 그 활용수준에 따라 구분하자면, 원래의 소프트웨어를 그대로 또는 약간 수정하여 사용하는 블랙박스 재사용, 인터페이스 수준에서만 코드를 수정하는 그레이박스 재사용, 소스코드에 많은 수정을 가하는 화이트박스 재사용 등이 있다.

패키지 단위로 사용할 것인지, 라이브러리나 컴포넌트 단위인지 또한 결정해야 한다. 개발 인프라가 부족한 기관은 패키지 단위나 완전한 블랙박스 라이브러리 또는 컴포넌트로 사용하여 시간과 비용을 줄이는 것이 바람직하다.

3) 마이그레이션

마이그레이션은 한 세대의 컴퓨터 기술로부터 다음 세대로, 또는 한 가지 조합의 하드웨어, 소프트웨어 설정으로부터 다른 것으로 정기적으로 디지털 자료를 옮기는 것을 말한다. 마이그레이션은 현재로서는 가

40) 김종배, 송재영, 류성열, 앞의 글, 2005, 232쪽.

41) Huang, Meng, Liguang Yang and Ye Yang, "A development process for building OSS-based applications", p.127.

장 선호되는 디지털 데이터의 보존 전략이지만 대용량 자동 처리 등의 문제를 해결하기 위한 실험이 아직 진행 중이다.

AtoM 시스템으로의 마이그레이션은 기존 기록관리 시스템에서 AtoM 시스템으로 어플리케이션과 데이터 모두를 이주시키는 것과, 데이터를 이동시키는 것으로 구분할 수 있다. 기존 기록관리시스템의 종류에 따라 차이가 있겠지만 기존 시스템을 AtoM시스템으로 마이그레이션한다는 것은 이전의 하드웨어·소프트웨어 설정 등을 LAMP⁴²⁾ 기반의 클라이언트 서버 모델로 완전히 변경한다는 의미이다. 기존 시스템 운용을 위해 호스팅 서비스를 이용하고 있었거나 리눅스 서버 환경을 갖춘 상태라면 AtoM시스템으로의 마이그레이션이 좀더 수월할 것이다.

데이터 마이그레이션은 소장 기록물 목록이나 데이터베이스를 AtoM 소프트웨어에서 인식할 수 있는 데이터 구조로 매핑하는 작업이다. 기존 시스템이나 DBMS를 운용하고 있는 기관이라면 소장목록, 전거레코드, 입수기록 등의 데이터베이스 테이블로부터 컬럼명 기준으로 CSV 파일을 추출해야 한다. 이후 AtoM이 제공하는 CSV 형태의 ISAD(G), ISAAR(CPF), Accession records 등의 템플릿에 해당 기관의 목록 컬럼명을 연결하고 일대일로 매핑되지 않거나 누락되는 데이터의 처리 여부를 결정해야 한다. AtoM 시스템 설치 후 실제로 매핑된 데이터를 벌크 임포트하여 품질 테스트를 진행한다. 이 절에서는 데이터 마이그레이션 방법 및 프로세스에 대해 알아보았다.

42) LAMP는 웹 서비스를 위한 소프트웨어의 묶음으로, 리눅스 운영체제(Linux), 아파치 웹서버(Apache), MySQL 데이터베이스 서버(MySQL), 그리고 스크립트 언어인 PHP의 앞글자를 따서 만든 약어이다. ([http://en.wikipedia.org/w/index.php?title=LAMP_\(software_bundle\)&oldid=634938968](http://en.wikipedia.org/w/index.php?title=LAMP_(software_bundle)&oldid=634938968)), [인용 날짜: 2015. 4. 12].

(1) 데이터 추출

AtoM 시스템으로 소장목록 등 기존 데이터를 마이그레이션하기 위해서는 우선 데이터베이스나 액세스 프로그램으로부터 CSV 형태로 데이터를 추출해야 한다. 엑셀 목록을 사용하는 경우 CSV 형태로 내보내기 기능을 이용하면 된다. AtoM 시스템으로 불러오기 가능한 CSV 파일은 UTF-8 형태의 인코딩, 유닉스/리눅스 스타일의 라인 엔딩 기호(line feed, $\backslash n$, 0x0A) 사용 등의 조건에 부합해야 한다.

(2) AtoM 컬럼 매핑

데이터 추출이 완료되면 AtoM에서 제공하는 CSV 템플릿의 데이터 구조와 해당 기관의 컬럼명을 매핑해야 한다. AtoM에서 제공하는 CSV 템플릿은 ISAD(G), ISAAR(CPF), RAD, Accessions, Events 등 5종이다.⁴³⁾

AtoM의 ISAD(G) 템플릿은 50개 컬럼으로 구성된다. 각 컬럼은 AtoM 사이트 상에서 하나의 필드를 의미하며, ISAD(G) 표준의 26개 항목과 AtoM 시스템이 관리를 위해 추가된 필드들로 구성된다. ISAD(G) 표준의 각 항목 중 기술계층(Level of description)의 경우 각 기록 계층의 포함관계를 표현하기 위해 계층ID(legacyID)와 상위계층ID(parentID), 큐비트 상위계층주소(qubitParentSlug) 등의 3개 컬럼이 추가로 요구된다. 이와 같이 특정 ISAD(G) 기술요소의 경우 여러 컬럼으로 구분되어 입력되므로 이를 정확히 파악하여 기존 소장목록과 매핑해야 한다.

우선 ‘인간과기억아카이브’의 명지대 컬렉션 목록을 AtoM 시스템으로 마이그레이션하는 상황을 가정해 보았다. 실제 각 기관들이 보유하고 있는 소장목록의 기술요소는 더 많은 항목을 관리하는 경우

43) <https://www.qubit-toolkit.org/wiki/CSV_import>, [인용 날짜: 2015. 4. 13].

가 많으므로 이를 참고하여 AtoM의 50개 컬럼에 매핑하여 마이그레이션 가능한 컬럼과 매핑되지 않는 컬럼의 처리방안을 고민해 보아야 한다. 어떤 컬럼은 여러 필드로 분산되기도 하고, 매핑되지 않는 컬럼을 ISAD(G)의 특정 필드에 임의로 매핑해야 하는 경우도 있을 것이다.

〈표 10〉 소장목록 CSV 예시

기록 계층			기술(description)								
컬렉션	시리즈	파일	아이템	식별자	생산자	생산 연도	유형	수량	보관 위치	공개 여부	설명
강의	2014년 2학기	기록학개론	출석부	MJ0001	교수	2014	종이	2장	보존서가가 열2단	비공개	16주차 20명 수강생의 출석사항을 체크한 출석부
			강의 계획서	MJ0002	교수	2013	디지털 파일, hwp	1개	스토리지서버 >001	비공개	각 주차별 강의내용과 참고문헌을 기술한 강의계획서
			1주차 사진	MJ0003	수강생	2014	디지털 파일, jpg	1개	스토리지서버 >001	공개	1주차 오리엔테이션 사진
			2주차 사진	MJ0004	수강생	2014	디지털 파일, jpg	1개	스토리지서버 >001	공개	2주차 학생발표 사진
			3주차 사진	MJ0005	수강생	2014	디지털 파일, jpg	1개	스토리지서버 >001	공개	3주차 견학 사진

〈표 13〉은 AtoM시스템의 ISAD(G) 템플릿에 위 소장목록을 매핑해 본 것이다. 좌측에 ISAD(G) 템플릿 컬럼명을 나열한 뒤 우측에 마이그레이션 대상 소장목록의 컬럼명을 매핑해하였다. 일대일로 컬럼이 매핑되면 마이그레이션 가능여부에 가능함을 표기하고, 이전에는 기술하지 않았지만 AtoM시스템으로 목록을 이전하며 새로 기술해야 할 항목들도 체크하였다. 참고사항에는 일대일로 매핑되지 않는 데이터를 처리하기 위한 해결방안이나 기타 사항들을 기술하였다.

〈표 11〉 ISAD(G) 컬럼 매핑표 작성사례

영역구분	AtoM의 ISAD(G) 컬럼명	기관 소장목록 컬럼명	마이그레이션 가능여부	추가기술 여부	참고사항
Identity area	legacyId	기록계층	Yes		
	parentId	기록계층	Yes		
	identifier	식별자	?		기존 식별번호를 AtoM의 식별자 구조로 전환하는 방안 고려.
	title	컬렉션, 시리즈, 파일, 아이템			legacyID와 parentId를 이용하여 모든 기록계층 구조 표현해야 함
	creators	생산자	Yes		생산자1 생산자2 복수입력 가능
	creatorDates	생산연도	Yes		시작일과 종료일이 있을 경우 2000 2014 등으로 표기 가능
	creatorDatesStart	생산연도	Yes		
	creatorDatesEnd	생산연도	Yes		
levelOfDescription	기록계층	Yes		fonds/collection/series/file/item 등으로 명칭 변경 필요	
extentAndMedium	유형, 수량	Yes		유형과 수량 컬럼을 병합해야 함	
Context area	repository			Yes	소장기관이름 표기
	acquisition			Yes	기증, 구매, 위탁 등 표기
Content and structure area	scopeAndContent	설명	Yes		
Allied materials area	relatedUnitsOfDescription			Yes	관련기록 연결
Access points	subjectAccessPoints			Yes	AtoM의 접근점 적극 활용 필요 Subject1 Subject2 등으로 복수입력 가능
	placeAccessPoints			Yes	
	nameAccessPoints			Yes	
Description control area	revisionHistory			Yes	기술 생성이력 관리 필요
Administration area	publicationStatus	공개여부	Yes		'공개'를 'Published'로, '비공개'를 'Draft'로 명칭변경 필요
	physicalObjectName	보관위치	?		보관위치 컬럼에서 보존용기이름 분리 필요 (보존서가, 스토리지서버)
	physicalObjectLocation		Yes		보관위치 컬럼에서 위치정보 분리 필요(가열2단)
	physicalObjectType		?		서가, HDD 등 구분 필요

컬럼 매핑이 완료되면 AtoM 템플릿 양식에 맞춰 기존 소장목록을 재작성해야 한다. 아래는 작성 사례이다. 각 계층의 고유번호를 의미하는

레거시ID(legacyID)와 상위계층을 의미하는 페어런트ID(parentID) 컬럼을 이용하여 각 기록계층의 포함관계를 표현하였다. 레거시ID를 사용하지 않고 'qubitParentSlug'컬럼에 직접 각 디스크립션의 URL 뒷자리를 입력하여 기록계층을 표현할 수도 있다. '강의'컬렉션의 생산자는 '교수'와 '수강생'이 되는데 이처럼 복수의 데이터가 입력될 경우 '!' 기호를 이용하여 구분한다. 원래 목록의 '유형'과 '수량'은 'extentAndMedium' 컬럼에 콤마로 구분하여 넣었고 '보관위치'는 'physicalObjectName'과 'physicalObjectLocation'으로 분리하여 입력했다. '공개여부'의 '공개' 및 '비공개'는 각각 'Published'와 'Draft'로 변경하여 표기하였다.

〈표 12〉 Atom 컬럼 매핑 결과 - ISAD(G) 템플릿 (소장목록.csv)

legacy ID	parent ID	Identifier	title	creators	creator Dates	levelOf Description	extentAnd Medium	physical ObjectName	physical Object Location	publication Status	scopeAnd Content
1		C01	강의	교수 수강생	2013-2014 2014	collection	총이 2장, 디지털파일 4개	보존서가 스토리지서버		Published	명지대학교 강의를 모아놓은 컬렉션
2	1	S01	2014년 2학기	교수 수강생	2013-2014 2014	series	총이 2장, 디지털파일 4개	보존서가 스토리지서버		Published	2014년 2학기 수업 시리즈
3	2	F01	기록학개론	교수 수강생	2013-2014 2014	file	총이 2장, 디지털파일 4개	보존서가 스토리지서버		Published	기록학 개론 과목을 아카이빙한 철
	3	IT01	출석부	교수	2014	item	총이, 2장	보존서가	가열2단	Draft	16주차 20명 수강생의 출석사항을 체크한 출석부
	3	IT02	강의계획서	교수	2013	item	디지털 파일, pdf, 1개	스토리지서버	001	Draft	각 주차별 강의내용과 참고문헌을 기술한 강의계획서
	3	IT03	1주차 사진	수강생	2014	item	디지털 파일, jpg, 1개	스토리지서버	001	Published	1주차 오리엔테이션 사진
	3	IT04	2주차 사진	수강생	2014	item	디지털 파일, jpg, 1개	스토리지서버	001	Published	2주차 학생발표 사진
	3	IT05	3주차 사진	수강생	2014	item	디지털 파일, jpg, 1개	스토리지서버	001	Published	3주차 견학 사진

(3) CSV 임포트(CSV import)

AtoM시스템은 EAD 형태의 XML과 UTF-8 인코딩 형태의 CSV파일 목록을 한꺼번에 가져올 수 있다. 보통 이러한 작업은 한번에 대규모의 데이터를 가져오게 되는데, 그래서 벌크 임포팅이라 부르기도 한다.

CSV 임포트 기능은 관리자 로그인 후 웹페이지 메뉴를 이용하거나 커맨드라인 인터페이스로 AtoM이 설치된 웹서버에 접속하여 명령어를 입력하여 수행할 수도 있다. 벌크 임포팅 작업은 많은 시간이 소요되고 한 번 입력되면 삭제하는 데 어려움이 발생하므로 샘플목록을 테스트하여 대략의 문제점을 파악할 수 있다. 대용량 목록일 경우 분할하여 가져오는 방법이 바람직하다. 예를 들어 20,000 건이 수록된 CSV 목록의 경우, 1,000 건을 먼저 테스트해 보고 이상 없음을 확인한 후 다시 1,000 건씩 20회 수행하는 방법이 권장된다. 테스트하기 전에 AtoM 인스턴스를 복제하여 에러가 발생할 경우에 대비하는 것도 유용한 방법이다. 만약 이미 업로드된 데이터가 있다면 이러한 벌크 임포트 작업으로 인해 기존 데이터가 유실되는 것을 막아야 하기 때문이다.

컬럼 매핑한 '소장목록.csv' 파일을 AtoM 시스템으로 가져온 결과 최상위에 '강의' 컬렉션이 생성되었고 하위계층의 시리즈, 파일, 아이템 5개와 기술항목, 그리고 우측의 보존서가까지 모두 정상적으로 등록되었다. 아마존 웹서비스(AWS) EC2에 설치된 AtoM 인스턴스에서 테스트했으며, 소요된 시간은 1.54초이다.

목록 뿐만 아니라 디지털 객체 또한 벌크 임포팅할 수 있다. 만일 각 목록에 링크될 디지털 객체가 있다면 AtoM이 설치된 웹서버의 특정 폴더에 디지털 객체를 업로드한 뒤 디지털 객체 링크를 위한 CSV파일을

별로 만들어 한꺼번에 임포트할 수 있다. 목록의 수가 아주 많다면 디지털 파일을 하나씩 수동으로 업로드하는 작업보다는 한꺼번에 마이그레이션하는 방법이 효율적이다. 마이그레이션 작업이 완료되었으면 이에 대한 보고서를 작성한다. 몇 개의 목록이 몇 회에 걸쳐 임포팅되었고 어떠한 문제점이 발생했는지, 어떻게 조치했는지 등을 기술한다. 마이그레이션 결과보고서에 지적된 필드별 이슈들은 CSV 컬럼 매핑에 반영하여 템플릿을 업데이트해야 한다.

4) 커스터마이징

일반적으로 시스템 커스터마이징은 수요기관의 요구사항에 맞춰 소프트웨어의 기능, 성능, 이용자 인터페이스 등을 대폭 변경하거나 추가로 개발하는 것을 말한다. AtoM 명세서에서 말하는 커스터마이징⁴⁴⁾은 세 가지 행위를 의미한다. 시스템과 어플리케이션 서버, 테마, 언어 및 타임존과 관련된 설정파일의 소스코드를 변경하는 행위이다. 이러한 커스터마이징 요소들은 일반적으로 선호하는 시스템 옵션을 설정하는 프리퍼런스(Preference)에 해당할 수 있으나, AtoM의 경우 광범위한 소스코드 변경이 이루어지므로 상당한 수준의 개발지식이 요구된다.

AtoM 시스템은 관리자 페이지와 서비스 페이지가 동일한 웹 기반 사용자 인터페이스를 제공한다. 따라서 AtoM을 사용하는 대부분의 기관들은 AtoM 웹페이지의 디자인을 커스터마이징하여 사용하고 있다. 이러한 디자인적 요소 외에도 이용자에게 가장 핵심적인 기능인 기록 열람 및 검색 등을 편리하게 하기 위한 기능적, 비기능적 요소들이 AtoM 시스템 구석구석에 산재해 있다. 새로운 기능을 추가하지 않더라도 관

44) <<https://www.accesstomemory.org/en/docs/2,1/#customization>>, [인용 날짜: 2015. 4. 14].

리자의 노력으로 이용자 편의를 제공할 수 있는 AtoM의 커스터마이징 요소들에 대해 알아 보았다. 또한 기존 기능을 개선하거나 사용자 인터페이스 디자인을 변경하는 방법에 대해 알아 보았다.

(1) 기능적 요소

AtoM의 관리자 메뉴를 통해 일반적인 기능의 커스터마이징이 가능하다. AtoM시스템은 엔티티의 명칭을 변경하거나 식별자 규칙을 변경하는 등 상당한 수준의 커스터마이징 요소를 관리자 인터페이스를 통해 제공한다. 시스템 구축 단계에서 관리자 인터페이스의 커스터마이징 항목을 조정하려면 해당 기관의 보존 및 서비스 정책이 수립되어 있어야 한다. 만약 정책이 마련되어 있지 않다면 충분한 시간을 갖고 AtoM 설정을 변경해 가며 아키비스트와 일반 이용자의 의견을 청취해야 한다. BCAUL의 사례처럼 복잡한 기능 개발보다 이용안내 설명문을 곳곳에 배치하여 기록을 쉽게 이용할 수 있도록 하는 것이 중요하다.

시스템 퍼포먼스 및 어플리케이션 기능 관련 요소는 설정 파일(Configuration files)의 소스코드 변경을 통해 커스터마이징된다. 커맨드 라인 인터페이스를 통해 AtoM 이 설치된 웹서버의 해당 파일에 접근하여 명령어를 수행하게 된다. 비용 효율성을 위해 AtoM 시스템의 퍼포먼스 성능 최적화는 필수적이다. 벌크 임포팅 등 시스템 부하가 큰 작업이 아니더라도, 기술정보 및 디지털 객체를 등록하고 검색 및 브라우징 할 때 안정적으로 쾌적하게 수행되는 것이 중요하다. 각 기관의 하드웨어, 소프트웨어, 인터넷 환경이 제각각이므로 IT전문가이 도움을 받거나, 전문 업체의 호스팅 등을 통해 지속적으로 퍼포먼스 향상을 위한 테스트가 요구된다.

〈표 13〉 AtOM 설정파일 커스터마이징 요소 및 방법

설정파일명	커스터마이징 요소 및 방법
CONFIG/APPS.YML	파일열기: nano config/apps.yml 디지털 객체 업로드 용량 제한 설정 외부로부터 디지털 객체 불러오는 타임아웃 시간 설정 (기본 10초) 캐쉬엔진 선택 (예를 들어 배포를 위한 Memcached 사용여부 등 설정, 기본은 sfAPCCache로 설정됨) Gearman 작업서버의 IP주소 설정(비동기식 작업을 위한) 구글맵 API key 설정 (기록소장기관 정보에 구글맵 위치정보 디스플레이 가능) 구글 애널리틱스 API key 설정 (사이트 접속통계 제공)
APPS/QUBIT/CONFIG/SETTINGS.YML	Symfony와 PHP 전반적 설정 ⁴⁵⁾ - 에러, 로그인, 보안, 모듈 등을 다룸 기본 국가(culture), 기본 타임존 등 설정 - 기본 국가를 미리 설정하지 않으면 english로 자동 설정됨 에러 발생시 처리방법을 설정 (전문가만 수정 요망)
CONFIG/FACTORIES.YML	Symfony 1.x 팩토리 설정 파일 ⁴⁶⁾ 디버깅 위한 로그 옵션 ⁴⁷⁾ 및 사용자 인증(LDAP) 등 설정
CONFIG/CONFIG.PHP	데이터베이스 접근정보가 저장되는 기본 설정파일 데이터베이스 접속자, 비밀번호, 이름 등 저장
/etc/php5/fpm.d/pool.d/atom.conf	AtOM시스템의 내부 요청은 모두 php5-fpm으로 전달됨 타임아웃, 환경변경 등 PHP 설정

(2) 이용자 인터페이스 요소

웹사이트 명칭과 사이트 설명문구, 사이트 로고 변경 등 AtOM 사이트의 특엔필을 변경하는 것 또한 중요하다. 로고나 색상 등 디자인의 변경은 해당 기관의 정체성을 드러내고 이용자들에게 디자인적 요소로

45) <http://symfony.com/legacy/doc/reference/1_4/en/04-Settings>, [인용 날짜: 2015. 4. 13].

46) <http://symfony.com/legacy/doc/reference/1_4/en/05-factories>, [인용 날짜: 2015. 4. 13].

47) <https://www.accesstomemory.org/en/docs/2_1/admin-manual/maintenance/logging/#maintenance-logging>, [인용 날짜: 2015. 4. 13].

통해 가시성, 편리함 등의 가치를 제공할 수 있다. 이와 함께 메인페이지의 사이트 소개(Static page), 도움말 설정 등 사이트 이용자들이 AtOM 사이트를 이용하는 데 가장 필요로 하는 유용한 정보들을 미리 파악하여 제공하는 것 또한 필수적이다. AtOM은 현재 두 가지 테마를 제공하므로 테마 변경만으로 디자인적 차별화를 꾀하기는 힘들다. AtOM 테마 커스터마이징 역시 테마를 구성하고 있는 CSS 및 PHP 파일의 소스코드 변경이 요구된다. 커스터마이징 가능한 요소들 중 대부분의 명세는 PHP와 CSS를 다루는 웹 개발업체에 제공하여 AtOM 사이트 디자인 커스터마이징의 참고자료로 활용될 것이다. 사이트의 각종 엔티티명과 드롭다운 메뉴의 옵션을 설정하는 것은 아키비스트의 몫이다. 예를 들어 'Archival descriptions' 엔티티명을 '기록물 기술'로 할지 '기록 정보'로 할지를 결정하는 것 등이다.

〈표 14〉 AtOM 테마 커스터마이징 요소 및 방법

설정 파일(디렉토리명)	커스터마이징 요소 및 방법
apps/qubit/templates/layout.php	웹사이트명, 태그라인(웹사이트 소개문구), 사이트 로고 수정
web/images	해당 디렉토리에 새로운 로고 이미지 업로드
graphic.css	website name & logo 섹션 수정
/web/favicon.ico	파비콘 파일 변경
apps/qubit/config/view.yml	웹사이트 메타태그 변경 (타이틀, 설명, 키워드)
staticPages.yml 또는 Static page	홈페이지의 고정페이지와 소개(About) 페이지 문구 수정 staticPages.yml을 수정하거나 홈페이지 메뉴를 통해 가능
siteSettings.yml 또는 User interface labels	사이트의 각종 엔티티명 변경 siteSettings.yml을 수정하거나 홈페이지 메뉴를 통해 가능
siteSettings.yml 또는 Default templates	각 모듈의 기본 템플릿을 변경 siteSettings.yml을 수정하거나 홈페이지 메뉴를 통해 가능
drop-down/picklist	메뉴 옵션에서 드롭다운 메뉴별 기본값 수정

본격적인 디자인 개선을 위해서는 사용자 인터페이스 요구사항을 작성한 후 각 요구사항을 목업이나 와이어프레임으로 작성하여 웹 개발 작업을 수행하게 된다. 목업은 종이에 스케치하듯이 원하는 화면을 그려놓은 것이고 와이어프레임은 레이아웃 화면과 각종 버튼 등을 제공하는 와이어프레임 작성 도구를 이용하여 실제와 유사하게 화면과 기능을 구현한 그림이다. 이러한 작업은 아키비스트가 원하는 바를 가장 정확하게 웹 개발업체에 전달함으로써 개발의 완성도를 높이고 결과적으로 사용자 만족도를 향상시키는 결과를 가져올 것이다. 우선, AtoM의 각 페이지를 메인화면, 검색, 브라우징 등 핵심 기능으로 구분한 뒤 아키비스트나 이용자가 기록을 쉽게 관리하고 이용할 수 있는 방향으로 스케치를 작성해 보는 것이 필요하다. 이를 바탕으로 웹 개발업체와 커뮤니케이션하며 사이트 디자인 커스터마이징을 진행할 수 있을 것이다.

5. 맺음말

오픈소스 생태계는 최근 크게 성장하고 있다. 기업들은 오픈소스 기반 비즈니스 모델의 성공을 위해 다양한 역할로 OSS 커뮤니티에 참여하고 있다. 사용자들은 뛰어난 제품을 무료로 사용할 수 있고 더 높은 성능이나 서비스에 대해 비용을 지불하는 것을 합리적인 조건으로 받아들이고 있다.

기록관리 분야에도 2000년대 초반의 오픈소스 운동에 힘입어 AtoM, 오메카, 디스페이스, 아카이브즈스페이스(ArchivesSpace) 등 다양한 OSS 프로젝트들이 등장하였다. OSS의 사용은 소규모 아카이브가 적은 비용으로 아카이브 시스템을 이용할 수 있도록 해 주었다. 또한 특정 소프트웨어에 종속되지 않고 자유롭게 여러 OSS 도구들을 선택하여 사용할 수 있는 환경을 만들어 주었다. AtoM 기반의 아카이브 시스템을 구축하

고 안정적으로 운영하기 위해서는 국내 AtoM 커뮤니티 활성화가 시급하다. 개발자 뿐만 아니라 적극적인 사용자 그룹이 대폭 유입되어야 한다. 최근 AtoM 호스팅 서비스를 통해 국내 여러 기관 및 기록관리대학, 교육원 등에서 적극적으로 활용하기 시작한 것은 커뮤니티가 활성화되기 위한 좋은 출발점이 될 것이다. 개발 인력들이 기록관리 분야 OSS 커뮤니티에 유입되도록 하기 위해서는 AtoM 시스템의 실질적 사용자 그룹이 실무적 관점에서의 요구사항을 만들어 낼 수 있어야 한다. OSS 기반 아카이브 시스템 구축 사례를 하나씩 만들어 간다면 국내 민간 기록관리가 한층 발전할 수 있는 계기가 될 것이다.

AtoM 배포버전이 출시된지 5년째 접어든 지금 초기 프로젝트의 산출물이 모두 집적되어 기능 면에서 상당한 발전을 이루었다. 그럼에도 불구하고 배포 버전을 그대로 사용하기엔 이용자들의 관점에서 부족한 점이 많다. 아키비스트들이 만들어 낼 실무 관점의 요구사항은 버그나 추가기능으로 개발일정에 등록되어 AtoM 시스템이 보다 편리한 기능을 제공하는 선순환 구조를 만들 것이다.

광범위한 커스터마이징은 투자 대비 효과가 떨어진다. 즉 다양한 OSS를 활용하는 것도 중요하지만 기관의 요구사항에 가장 적합한 툴을 선별하고 효율적으로 활용할 수 있는 방안을 고민하는 것이 중요하다. OSS 또한 관리 및 유지보수, 이용자 교육 등의 비용이 필연적으로 발생하므로 안정적 운영을 위한 지속적인 예산 배정 또한 이루어져야 한다.

AtoM 시스템 구축과 커스터마이징 방법에 대한 이 연구가 국내 AtoM 커뮤니티 활성화에 작은 보탬이 되길 바란다.

ABSTRACT

A Study on Constructing and Customizing an AtoM System

An, Dae-Jin · Kim, Ik-Han

In 2012, the International Council on Archives(ICA) distributed AtoM public release, an archival description software based on web, for free to support small archival institutions. The purpose of AtoM project was to let a single or multiple repositories describe archival materials based on ICA descriptive standards and access via a web browser. The value of AtoM project is to promote actual business experience by sharing technical expertise and form active community between users and developers.

Recently, there is a growing interest in open source software in the field of domestic record management. This interest is to not only reduce cost, but to escape technology dependencies and to address requirements on the field. The critical mind of this study lies in the fact that there is no practical methodology to establish AtoM system. Even when using AtoM itself, basic tasks such as the migration of the existing data or the improvement of user interface are required. That is why this study suggests the process and methodologies to establish and customize archival information system based on AtoM software.

Key words : AtoM, open source software, record management system