

[세미나]

한국기록관리학회 추계학술대회 : 기록시스템의 오픈소스화 전략

발표 1.

서울기록원 디지털 아카이브 : 완성에서 확장으로

원 종 관

서울시 정보공개정책과

IFLA의 동향보고서(2016년 업데이트)는 새롭고도 낯선 질문으로 시작한다.

“200년 전이 아니라 오늘날 도서관이라는 것을 처음 발명하고 제안한다면 그 도서관의 개념이 무엇일 것인가”

기술의 발전속도는 숨이 차게 빠르고, 새로운 세대는 그 기술을 자연스럽게 받아들이는데, 문화적 인프라의 대표 기관인 도서관은 이 변화를 얼마나 수용하고 있다. 아니, 변화를 위협으로만 받아들이고 있는 것은 아닐까. 다행히 전 세계의 도서관은 이 문제를 끌어안고 치열한 토론을 하고 있다. 보고서는 그 과정을 중계한다. 토론은 아직 끝나지 않았다.

아카이브 종사자들은 그게 엄살처럼 보일 수도 있겠다. 사람들은 도서관을 아카이브보다 훨씬 보편적(?)으로 인식하니까. 기회로 생각하든 위협으로 느끼든 저변의 문제의식은 비슷하다.

“책의 보관소가 아니라 정보와 아이디어의 플랫폼”,

“사회와 정보환경의 변화를 이해하는 도구”, “커뮤니티와 광장”

최근 미 의회도서관장에 취임한 Carla Hayden의 취임사는 같은 맥락으로 읽을 수 있다.

(도서관이) 기존의 학자들이 연구를 위해 찾아오는 곳만이 아닌, 새로운 학자들을 길러내고, 어린 작가들에게 영감을 제공하며, 워싱턴 밖의 사람들과 연결하는 곳이 되기를 바란다.

서울기록원은 새로 만들어지는 지방기록물관리기관이다. 법적 정의가 그렇고 ‘서울의 (본격적인)아카이브’가 목적이다. 지향하는 정체성은 현실의 법적 규정을 넘는 더 큰 그림을 그린다. 짧은 발표에서 차이를 모두 밝힐 수는 없지만 일종의 전제로 두 개의 개념은 다르다. 변화를 바라보는 눈, 해석하는 정책과 철학, 실행의 힘까지. 종전의 기록관, 기록물관리기관의 범주를 넘어 어떻게하면 서울의 아카이브를 새롭게 발명할 수 있을까.

2016년 <서울기록원 정보화전략계획ISP 수립 사업>과정에서 향후 서울기록원이 맞닥뜨릴 다양한 이슈를 확인했다. 크게는 서울역사박물관, 서울시립미술관 등 서울시의 여러 문화기관과의 관계부터 이미 각 실, 국, 부서에서 추진하고 있는 아카이빙 사업까지 서울이라는 도시의 크기만큼이나 현황은 복잡했다. 몇 년 간 아카이브(빙)의 정의와 범주는 DB구축 사업부터 업무 정보를 정리하는 시스템 개발까지 폭넓게 확장되고 있다. 각자의 정의는 다르지만 그런 개별적 사업이 아카이브의 범주에 포함되지 않는다고도 할 수 없다.

이와 같은 현실에서 서울기록원은 아카이브를 독점적으로 규정하기보다 수용하고 협력할 수 있는 방법을 찾고자 한다. 물론 서울기록원이 가장 잘 할 수 있는 것에는 집중하고 그것을 특화하려고 한다. 디지털 아카이브가 그 중 중요한 축이다. DB든 정책이든 기반 인프라든 ‘내 것’을 더 많이 공개할수록 더 커질 수 있다는 철학은 디지털 아카이브를 관통한다. 이미 많은 정보 비즈니스에서 확인된 전략과 모델이다.

서울기록원 디지털 아카이브가 완성보다 확장을 지향하겠다는 선언은 개발 소스, 개발 방법론, 배포 전략에 공통적으로 적용된다. 레거시(legacy) 시스템의 저항, 공공기관의 관료적 특성은 예상 가능한 걸림돌이지만, 변화에는 늘 일정한 반동이 따라온다는 것 또한 우리는 잘 알고 있다.

완성에서 확장으로 가는 첫 번째 경로는 오픈소스소프트웨어(OSS)의 채택이다. 점점 더 많은 IT 프로젝트가 이른바 ‘한국형’, ‘서울형’이라는 독자규격을 버리고 호환성과 상호운용성, 확장성을 지향하고 있다. ‘바퀴를 새로 발명하는 일’의 비효율을 우리 모두가 확인한 것은 다행이다. 깃허브(GitHub)는 개발자들에게 가장 중요한 소스 코드의 바다이고, 공개된 API의 적용과 개발은 IT 프로젝트에서 그 효용을 새롭게 확인할 필요조차 없게 되었다.

다만 중요한 것은 OSS의 사용 자체가 아니다. 공공 IT 프로젝트의 짧은 개발 일정과 완고한 행정 프로세스를 고려할 때, 효율은 상용 솔루션이든 OSS든 기존에 존재하는 검증된 ‘바퀴’

의 재활용과 같은 말이다. 다행히 이런 고민과 노력의 역사는 짧지 않다. 우리는 ISP 수립 과정에서 Preservica, Archivemtica, Internet Archive 등 여러 가지 제품을 확인했다. 여기에 더해 도서관, 박물관, 미술관 등의 다양한 정보 시스템 기술을 참고할 수 있다. 물론 다른 정보 비즈니스 분야의 OSS와 API도 우리의 검토 대상이다. 다시 확인하자면 OSS는 채택 그 자체가 목적이 아니라 개발 효율성의 극대화가 가장 중요하다.

두 번째는 애자일(Agile)개발 방법론의 채택이다. 서울기록원 디지털 아카이브는 크게 두 개의 시스템으로 나뉜다. 영구기록관리시스템인 백엔드(back-end)와 웹서비스인 프론트엔드(front-end) 개발이 각각 큰 과업이다. 주어진 시간은 불과 10개월이다. 백엔드는 기록생산시스템과 기록관리시스템(RMS)을 포용하는 아카이브 리퍼지토리(archival repository)이자 서울기록원 운영의 중추로 수많은 기능의 총합이다. 프론트엔드는 서울기록원이 디지털 세계의 고객과 만나는 접점이다. 디지털 정보의 효율적 유통과 배급에 더해 “모바일은 디지털을 먹어치우고 있다”(Mobile is Eating the World)

변화를 예측하는 일은 과업지시서의 기능 명세처럼 명확하지 않다. 우리는 미리 완벽하게 합의하고 정의한 일목요연한 문서를 사랑하지만, 사실 IT 프로젝트에서 그것은 신기루에 가까운 일이다. 애자일 개발방법론은 기존의 순차적 폭포수(waterfall, 문제 정의 - 디자인 - 개발 - 테스트) 모델의 한계를 인정하고 각각의 태스크를 더 작은 프로세스로 나눈다. 작동하는 소프트웨어를 릴리즈(delivery)하고 이 과정을 반복 피드백하는 것이 더 중요하다. 이것은 만능의 도구는 아니지만 미국 정부의 디지털 서비스 랩(lab)인 18F과 영국 정부의 GDS(Government Digital Service)가 소기의 성과를 이룬 사례가 있다.

세 번째는 플랫폼으로서의 디지털 아카이브이다. 아카이브, 도서관, 박물관, 미술관은 사회의 문화적 인프라이다. 디지털에서도 같다. 서울기록원 디지털 아카이브는 기술적 인프라(tech habitat)를 지향한다. 그 동안 공공분야 뿐만 아니라 시민사회에서도 아카이브(빙)에 대한 요구와 인식은 커져왔다. 다만 인프라 사업의 특성 상 비교적 큰 자본과 인력 투입이 늘 과제였다. 공공분야는 예산 투입을 통해 일정하게 진행했지만, 시민사회의 아카이브 사업에는 누구도 또렷하게 답하지 못했다. 몇몇 선구적 노력을 제외한다면, 사실 기록 커뮤니티는 죄책감을 떠안은 채로 이 문제를 외면하고 방치했다고 할 수 있다.

서울기록원 디지털 아카이브가 이것을 한 번에 해결하지는 못할 것이다. 하지만 적어도 인프라와 서비스로써의 디지털 아카이브는 예산과 인력의 높은 진입장벽을 개방할 시작점이 될 수 있다. 강조하자면 이것은 시스템이 아니라 서비스를 지향한다. 특정한 클라이언트를 설치해야만 이메일을 쓰던 시절이 과거가 된 것처럼 시스템과 서버, 네트워크를 별도로 구축하는 디지털 아카이브도 천천히 과거가 될 것이다. 그 형태를 SaaS(Software as a Service)로 단정할 수는 없지만 가장 가까운 모델이 될 수 있을 것이다. IT 플랫폼 비즈니스의 거인

들은 고객 데이터 수집을 극대화하기 위해 플랫폼을 무료로 제공하지만 우리의 목적은 디지털의 힘을 빌린 '망라적 수집'과는 거리가 멀다. 오히려 다양한 영역의 아카이브(빙)가 자생할 수 있는 생태계 구축에 가깝다.

공공기록물관리법이 개정된 지 10년이 지났다. 디지털에서의 10년은 법과 규정보다 훨씬 급진적이었다. 요즘은 누구나 클라우드와 빅데이터를 말하지만, 10년 전 우리에게 클라우드 낯선 말이었다. 순차적 넘버링은 아니지만 공공기록 2.0을 모색할 시기이다. 이것은 디지털 아카이브에도 같은 무게의 변화를 요구하고 있다. 서울기록원은 디지털 아카이브를 어떻게 새롭게 제안할 것인가.