1. 메타데이터 기반 기록관리시스템의 현황과 전망

Barbara Reed
(호주 Recordkeeping Innovation Pty Ltd 대표)

메타데이터 기반 기록관리시스템의 현황과 전망

Barbara Reed(호주 Recordkeeping Innovation Pty Ltd 대표)

도입

오늘 이 자리에서 호주의 기록관리 현황 분석과 달성하고자 하는 방향에 대해 발표할 기회를 갖게 되어 개인적으로 영광으로 생각합니다. 한국 국가기록원에 깊은 감사를 드립니다.

이번 심포지엄을 위해 저에게 미리 보내주신 자료들을 읽으면서, 또한 국제 표준화 작업에서 국가기록원의 직원들과 함께 작업할 기회를 통해, 한국이 종래의 기록관리 방식에서 새로운 방식으로 얼마나 큰 변화를 이루었는지 알게 되었습니다. 우리는 이러한 한국의 경험으로부터 많은 것을 배우고 있습니다.

호주 경험의 배경

레코드키핑은 수많은 조건을 필요로 합니다. 즉 레코드키핑이 수행되기 위해서는 많은 시간과 공간, 기타 인적 물적 자원이 소요됩니다. 우리가 어떻게 레코드키핑을 수행하고 있으며, 어떻게 잘 수행할지는 문화적 풍토에 달려 있습니다.

저의 동료인 Greg Goulding이 잠시 후에 발표하겠지만, 호주와 뉴질랜드는 영어를 모국어로 하는 여타 국가들과 동일한 경험을 공유하고 있습니다. 우리는 멀리 떨어진 식민 지배국 정부에 의해 통치되는, 식민지 경험을 지닌 여느 국가들과 동질성을 공유합니다. 우리들이 지닌 식민 유산은 먼 과거의 것이지만, 이것은 우리의 정부체제, 업무 수행방식(그리고 레코드키핑 방식) 및 공공영역과 민간영역이 상호작용하는 방식에 중대한 영향을 미쳐 왔습니다.

레코드키핑은 정부기관에서 좀 더 체계적으로 수행합니다. 왜냐하면 의무적으로 준수해야 할 기록관리법과 각종 표준이 존재하기 때문입니다. 민간영역의 경

우 우리는 단지 바람직한 기록관리 방식을 도입하도록 권고할 뿐이고, 아울러 우리가 추천하는 표준 및 시스템은 업무를 쉽고 효율적이게 하며 업무상의 이익을 보호하도록 하기 위해 노력할 뿐입니다.

그러나 호주의 정부기관에는 '무엇을 하라고' 지시를 내릴 수 없습니다. 기록관리법은 있지만, 기록관리법을 지키는 것은 전적으로 각 정부기관에 달려 있습니다. 따라서 호주 공공영역의 레코드키핑은 매우 큰 편차를 지니고 있습니다. 어떤 기관들은 그들이 수행하는 업무와 통합된 전자레코드키핑시스템을 가동하고 있습니다. 또 어떤 기관들은 종이기록과 전자기록을 동시에 관리하는 하이브리드 환경 하에 있습니다. 그리고 일부 기관에서는 전자기록을 관리할 어떠한 시스템도 존재하지 않아, 여전히 전자기록을 종이에 출력하여 관리하고 있는 실정입니다. 호주의 기록관리 소프트웨어 및 이들의 수행능력은 바로 이러한 상황을 반영하고 있습니다.

호주의 접근

우리들의 과거와 현재적 실제는 우리들이 기록 및 업무 처리행위를 바라보는 방식에 영향을 주어 왔습니다. 기록은 업무의 일부라는 것이 우리들의 전통이었습니다.

업무 처리행위 및 업무 처리 행위 시 생성된 기록 간에는 매우 밀접한 관계가 존재하며, 전자적 환경에서 이 양자는 보통 동일한 것으로 간주됩니다. 우리에게 기록과 처리행위 간의 관계는 핵심적 사안입니다. 기록은 업무를 수행하는데 필수불가결한 것으로, 행해진 업무와 분리될 수도 또한 더해질 수도 없습니다. 문서(Document)와 기록(Record)은 서로 다른 절차에 있긴 하지만 동일한 것입니다. 만일 우리가 형성 면에서 문서들의 협력적 구축에 대해 이야기 한다면, 문서와 기록 간의 구분은 단지 이해를 위한 것입니다. 그렇지 않다면 문서는 처리행위 내지 특정 유형의 행위에 대한 증거를 제공한다는 점에서, 모든 문서는 기록이 되어야 합니다.

또한 기록은 다양한 목적으로 활용될 수 있다는 점을 우리는 알고 있습니다. 기록은 업무의 시작 단계에서 필요하고, 수행된 업무에 대한 설명책임의 도구로, 처리행위 당사자들의 권리 및 의무 보호수단으로, 그리고 사회 및 문화적 목적을 위해 필요로 됩니다. 이러한 기록이 지닌 다양한 필요성은 기록의 생산단계 이전부터 기록을 관리할 시스템에 반영될 필요가 있습니다. 영구보존 목적(업무 처리행위와 상관없는 외부인들이 접근해 활용할 수 있는)을 위한 기록의 지속성(Sustainability) 및 확고성(Robustness)을 확보할 수 있는 능력은 기록의 생산 시점에서부터 필요합니다.

이런 접근방법은 비단 디지털 기록에만 한정될 수는 없지만, 우리가 디지털 레코드키핑에 대해 사고하는데 도움을 줍니다. 일단 어떠한 기록이 필요한지 안다면, 우리는 시스템이 해당 기록을 생산・획득하도록 할 수 있습니다. 시스템은 통합된 것이거나 별도의 것일 수도, 집중화 된 것이거나 그렇지 않은 것일 수도 있습니다. 또한 시스템은 기록을 장기 보존하기 위한 것일 수도 있으며, 또는 기록이 필요로 되는 한 해당 기록을 유지하기 위해 하나의 시스템에서 다른 시스템으로 이전시킬 수도 있습니다. 이를 감안할 때 기록을 관리하기 위한 컴퓨터 기술 및 옵션은 다양하게 선택할 수 있습니다. 만일 우리가 왜 기록관리를 수행하고 있고 또 얼마나 오래 기록을 유지해야 하는지를 안다면, 기록관리 절차 및 컴퓨터 기술은 주어진 환경에 맞게 창의적이면서도 유연성 있게 선정할 수 있습니다.

하지만 만일 우리가 이러한 큰 그림을 이해하지 못한다면, 왜 우리가 무엇인가를 하고 있고 그것의 목적은 무엇인지 혼돈하게 됩니다. 우리는 21세기 초반 가장 큰 도전이라 할 수 있는 기술적 변화에 의해 위협받고 있습니다. 우리는 새로운 세계를 맞아 기존의 것들을 재검토 할 수 있는 합리적인 개념적 기반이 필요합니다. 이것은 우리가 달성하기 위해 노력하고 있는 것을 수행하기 위해 필수적인 사항입니다. 만일 우리가 그것을 안다면, 우리는 종이 세계에서 우리가 실행해온 방식을 다시 한번 재숙고할 수 있습니다. 즉 종이 환경에서 사용해왔던 것들이 디지털 세계에서도 여전히 유효한지, 아니면 종이 환경에서만 유효했던 방식이었는지에 대해 말입니다. 하지만 디지털 세계에서는 기존의 것과는 다른 방식을 통해 기록관리의 목표를 실질적으로 달성할 수 있습니다.그리고 이것은 바로 현재 우리 모

두가 직면하고 있는 도전입니다. 진정한 디지털 사고로의 전환은 아직까지는 없는 것 같습니다.

디지털 기록

디지털 기록은 종이기록과 동일한 성격을 지니면서도, 다른 한편으로는 매우 다른 특성을 지닙니다. 양자가 동일한 성격을 지닌다는 것은 필요한 기간동안 오래 지속될 수 있는 확고한 기록(처리행위에 대한 증거)을 생성시켜야 한다는 면에서이고, 서로 다른 점은 기록을 관리하는 포맷, 기술, 방법들이 종이 세계에서 그동안 해왔던 방식에 도전을 안겨다 준다는 점입니다. 종이기록의 관리 방법론은 오랜 세월을 거쳐 진화해 왔습니다. 기록은 그 안에 수록된 내용과 함께 기록을 관리하는 절차에 관한 수많은 관련 정보를 포함하고 있습니다. 종이문서를 기록철에 처리행위의 순서대로 축적하는 것, 서명과 같은 주석과 함께 누가 다음 단계를 수행하는지 혹은 어떻게 응답해야하는지에 대해 수기로 작성한 지침 그리고 처리행위의 응답이나 다음 단계는 우리가 관리하는 물리적인 기록 그 자체 내에 모두 포함되어 있습니다. 종이 세계에서 서식(Form), 내용(Content), 절차(Process)는 보통 하나의 동일한 것이었습니다. 하지만 디지털 기록으로 시각을 돌리면, 우리는 이러한 구성요소들을 정의하고 획득하는 방식을 재숙고할 필요가 있습니다. 왜냐하면 디지털 기록은 더 이상 데이터 그 자체에 물리적으로 체화되어 있지 않기 때문입니다.

기록은 서로 다른 유형의 데이터들로 구성된 구조물이라는 점을 이해할 필요가 있습니다. 먼저 기록물의 내용을 들 수 있습니다. 기록물의 내용은 행해진 처리행위에 관한 데이터입니다. 우리가 기록으로서 관리하기 위해 선택하는 데이터는, 기록으로서 관리될 수 있도록 의식적으로 선택하지 않는다면, 기록으로서 존립할 수 없습니다. 그것은 데이터에 불과하며, 쓸모없는 데이터를 줄이고 가장 최신의 값만을 최소한의 사본과 함께 저장하여 데이터관리시스템을 통해 관리되게 됩니다. 하지만 일련의 데이터를 기록으로 관리할 필요가 있다고 결정하게 되면, 그것의 성격 일부를 변화시킬 필요가 있습니다. 우리는 그 기록의 데이터가 행위가 수행되었을 때와 동일하며, 변경 되지 않으며, 행위가 발생한 시간과 공간이 연계되

어 있다는 점을 알 필요가 있습니다. 이것은 본래의 업무 처리행위에 참여하고 기록으로 획득된 데이터의 집합입니다. 현재 우리는 이것을 '기록 객체'(Record Object)라 부릅니다.

하지만 본래의 내용 말고도 기록이 지녀야 할 보다 많은 것들이 있습니다. 우리는 기록 객체를 누가 언제 어떠한 행위의 과정 속에서 생산했는지를 알려주는 업무 맥락 가운데에서 획득할 필요가 있습니다. 본래의 기록에 대한 이러한 데이터는 기록 객체와 연동하여 파악해 관리해야 할 데이터 집합의 시작이라 할 수 있습니다. 우리는 이러한 데이터를 '획득 시점 레코드키핑 메타데이터'(어떤 시스템에서는 이를 '문서 프로파일'(document profile) 메타데이터라고 함)라고 칭합니다.

신뢰성 및 진본성을 지닌 기록을 확보하기 위해서는 이러한 메타데이터만으로는 부족합니다. 우리는 기록이 본래 의도한 취지와 맞는지를 입증할 필요가 있습니다. 아울러 우리는 그것이 적절하게 관리되어 왔는지, 레코드키핑 프로세스가 기록을 보호하기 위해 적용되었는지, 그리고 어떻게 기록이 시간을 지속하며 관리되어 왔는지를 증명할 필요가 있습니다. 우리는 이것을 '레코드키핑 프로세스 메타데이터'라고 부릅니다. 레코드키핑 프로세스 메타데이터는 기록 객체가 존속하는 한 기록관리의 일부로 축적되게 됩니다.

그러나 이러한 기록 메타데이터는 훨씬 복잡한 측면을 지닙니다. 우리는 기록 메타데이터를 일련의 관계(Relationship)로 표현할 필요가 있습니다. 이러한 관계는 기록 객체 및 기록 객체를 가지고 상호작용한 사람(기록을 생산하고 관리한) 간에 설정됩니다.

일단 기록 객체 및 레코드키핑 메타데이터, 그리고 이들 간의 관계의 관점에서 생각해보면, 우리는 레코드키핑의 새로운 세계 안으로 들어가게 됩니다. 이전과는 다른 방식으로 디지털 기록을 획득·관리할 개연성이 창출됩니다. 그동안 종이세계에서 우리가 해왔던 방식인, 레코드키핑에 관해 계승되어 왔던 지식의 대부분에 의문을 제기할 수 있다.

우리는 레코드키핑 메타데이터를 이용하여 기록을 정의하는 세계에 살고 있습니다. 이것은 현재(또는 앞으로) 우리의 핵심 업무이며, 소프트웨어가 자동적으로 메타데이터를 생산하도록 하게 될 것입니다. 그러나 우리는 항시 우리들 스스로에게 도전해야 합니다. 우리는 실제 무엇을 필요로 하고 어떠한 것이 추가적인 것이며, 왜 어떤 데이터는 기록의 메타데이터로 획득해야 하며 다른 것은 아닌지를 말입니다. 현재 구축 중인 시스템이 이러한 메타데이터를 적절히 생산・관리하는지 어떻게 검증해야 할까요? 우리는 실제로 우리의 필요성들을 정의해 왔을까요?

모든 전문영역 내지 학문영역에서는 그 자체의 메타데이터 요건을 지니고 있다는 점 역시 이해할 필요가 있습니다. 이러한 메타데이터 요건들은 해당 영역의 필요에 맞추어 개발된 것들로, 모두 해당 영역에서 유효성을 지닌다고 볼 수 있습니다. 우리는 메타데이터 앞에 붙은 한정어 없이, 즉 어떠한 용도의 메타데이터인지를 모르고서는 메타데이터를 논할 수 없습니다. 메타데이터 자체에 대해서만 이야기하는 것은 혼돈을 일으킵니다. 이 자리에 모인 우리들의 관심은 레코드키핑 메타데이터입니다. 레코드키핑 메타데이터는 기록을 관리하는데 필요한, 동일한 속성을 지닌 세트 내지 데이터의 패키지끼리 연관되어 있어야 합니다. 레코드키핑 메타데이터는 기록으로 관리될 필요가 있는 데이터의 식별 가능한 세트 혹은 패키지와 계속적해서 관련되어야 합니다 이러한 레코드키핑 메타데이터는 생산이나 검색에 초점을 맞춘 메타데이터와는 다릅니다. 물론 이들 메타데이터 요소들과 중첩될 수도 있지만, 아무튼 우리들의 초점은 레코드키핑입니다. 그것은 우리들만의 전문적 영역이라 할 수 있습니다.

우리는 이 영역에서 무엇을 해왔는가?

레코드키핑 메타데이터가 앞으로의 디지털 시대에 핵심적 사안이 될 것이라는 논의는 1990년대부터 있어 왔습니다. 레코드키핑 메타데이터에 초점을 맞춘 첫 번째 프로젝트는 아래에 명시된 1990년대의 전자레코드키핑 관련 연구 프로젝트들입니다.

■ 미국 피츠버그대학교 프로젝트: Functional Requirements for Evidence in

Recordkeeping

- 캐나다 UBC 프로젝트: The Preservation of the Integrity of Electronic Records
- 호주 모나쉬대학교 프로젝트:Recordkeeping Metadata Standards for Managing and Accessing Information Resources in Networked Environments Over Time for Government, Commerce, Social and Cultural Purposes People and Partners

호주에서는 이러한 메타데이터 연구를 진전시키기 위해 그동안 매우 활동적으로 임해왔습니다. 호주는 상위 레벨에서 레코드키핑 메타데이터를 설명해주는 ISO 23081의 개발 및 발간에 적극적으로 참여해 왔습니다. 각 국가별로도 레코드키핑 메타데이터 요소들을 지정해 이를 발간했습니다. 뉴질랜드와 같은 국가들에서는, 조직이 레코드키핑 메타데이터를 적용시키는데 필요한 최소한의 요건들을 표준으로 간행하였습니다.

최근 저는 '호주 표준 레코드키핑 메타데이터요소'(Australian Standard Recordkeeping Metadata Element)를 지정하는 프로젝트의 책임을 맡았습니다. 여기서는 국제표준 요건과 함께 호주의 레코드키핑 전통을 따르고 있습니다. 우리가 개발하려고 하는 메타데이터는 각 조직의 기록시스템・업무시스템・기록관리 전담 시스템・영구보존시스템 등 현 기록시스템 하의 어떠한 레코드키핑 환경에서도 적용될 수 있도록 하는 것입니다. 이러한 메타데이터를 사용하여 우리는 기록을 생산・관리하는 모든 시스템 환경들을 서로 연계시키고, 이를 통해 필요할 때 시스템 간에 기록이 이동될 수 있는 인프라를 만들 수 있기를 기대하고 있습니다.

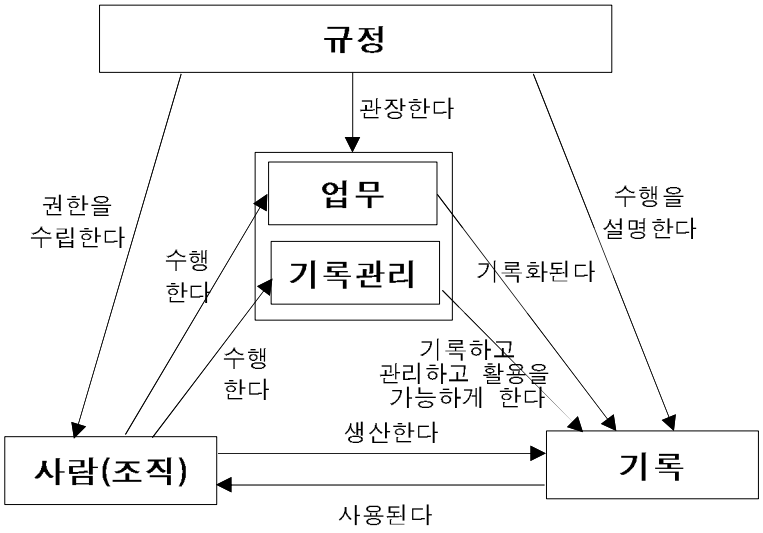
이와 같은 우리들의 구상 현실에 표준은 핵심적인 역할을 담당합니다. 우리는 시스템 설계 및 구축자들에게 어떠한 메타데이터가 필요하며 왜 이것을 획득해야 하는지 주지시킬 필요가 있습니다. 이래야만 오랜 시간에 걸쳐 필요한 메타데이터를 자동적으로 생산해 적절하게 관리하는 시스템을 구축할 수 있기 때문입니다.

우리 사회는 레코드키핑 메타데이터의 중요성을 매우 느리게 이해하고 있을지라도, 우리는 이를 계속적으로 논의할 필요가 있습니다. 레코드키핑 메타데이터는 복잡합니다. 또한 레코드키핑 메타데이터는 개념적입니다. 레코드키핑 메타데

이터는 우리가 달성코자 하는 것을 시스템에 맞는 새로운 언어로 번역해주고 있습니다. 그것은 디지털 세계의 레코드키핑에 결정적인 역할을 담당하게 될 것입니다.

현행 표준의 실제

오세아니아(호주와 뉴질랜드) 지역에서는 국제표준의 기반이 된, 호주의 연구 프로젝트 결과를 사용해 왔습니다. 그 핵심적인 골간은 아래의 그림과 같습니다.



<그림 1> 주요 엔티티와 그 관계[1)]

실제 세계에서는 무엇이 일어나며, 어디서 수행되는가?

호주에서는 지난 20세기 말부터 레코드키핑을 재검토해 왔습니다. 우리는 1990년대 초부터 종이 세계에서 수행해 온 우리의 방식이 디지털 시대에 맞지 않으며 근본적으로 변화되어야 함을 인식해 왔습니다. 레코드키핑 표준개발 과업은 우리가 추진해 온 영역 중의 하나입니다. 이와 맞물려 호주의 공공기록 관련 법령들 역

1) <그림 1>은 *ISO 23081-2, Information and Documentation - Managing Metadata for Records - Part 2: Conceptual and Implementation Issues*로부터 차용함. 원래의 그림은 호주 모나쉬대학교의 연구 프로젝트인 *Recordkeeping Metadata Standards for Managing and Accessing Information Resources in Networked Environments Over Time for Government, Commerce, Social and Cultural Purposes People and Partners*에서 제시하였다.

시 개정작업을 추진해 왔습니다. 레코드키핑 표준 도입을 적극적으로 권장하고 있으며, 호주의 조직들이 레코드키핑에 접근하는 방식을 변화시킬 수 있는 관계 규정 및 업무방식 개발을 위해 노력해 왔습니다. 그리고 이러한 노력들은 일정 부분 성공을 거두어 왔습니다.

주로 우리는 생산 시점부터 기록으로서 관리되도록, 기록을 시스템으로 획득하는 방안에 초점을 맞추어 왔습니다. 이는 일단 기록을 획득한 다음, 오랜 시간동안 기록을 유지할 수 있는 시스템 수행 능력을 확보한다는 발상에 기초한 것입니다. 이러한 소프트웨어 솔루션은 실제 업무가 수행되는 환경에 초점을 맞추어 왔습니다. 이것은 기록관리 및 전자문서기록관리시스템 영역입니다. 우리는 이러한 여정의 시작단계에 있습니다.

소프트웨어 1세대(대략 1985년-1994년)

기록을 관리하기 위해 사용된 1세대 시스템은 종이기록 관리를 자동화하기 위한 전자 메타데이터 시스템 제공을 목표로 하였습니다. 이것은 본질적으로 종이기록 관리의 자동화라 할 수 있습니다. 기록 대부분은 상위의 시리즈로 통합될, 종이 기록철 형태로 존재하였습니다. 전자 메타데이터 시스템은 종이기록의 등록・분류・처분 및 위치확인과 같은 관리조치들을 도왔습니다. 이러한 시스템은 호주 자체 내의 기록관리 업무 토양에서 개발된 것입니다. 호주에서 개발된 이러한 시스템으로는 TRIM과 RecFind가 있고, 캐나다 개발된 것으로는 Provenance를 들 수 있습니다.

이런 시스템들에서 흥미로운 점 중의 하나는 시스템이 등록부나 색인을 대체했다는 점입니다. 이러한 종이 형태의 등록부 및 색인은 전통적으로 영구보존 가치를 지닌 대상으로 간주되어 왔습니다. 등록부와 색인은, 일단 기록이 사무실에서의 현용적 필요가 끝나게 되면 기록의 설명과 향후 이용을 위해 필요로 되었기 때문입니다. 호주의 기록보존소는 모든 형태의 전자기록을 보존할 능력을 지니고 있지 않았고, 그 시기까지 기록은 조직 자체의 보존공간으로 이관되었습니다. 따라서 우리는 시스템을 포기하고 기록의 이관 시 리스트 형태로 동반되는 종이 기

반 산출물을 활용하였습니다. 이것은 기록관리 상의 전체적인 통제도구가 개발되지 못함과 더불어, 기록보존소로의 이관 시점에 종이기록에 대한 핵심적 통제를 중단시켰음을 의미하는 것입니다. 우리는 조직에서의 현용기록관리와 기록보존소에서의 영구보존 대상 관리 사이를 엄격히 분리시키는 라이프사이클의 논리를 강화시켰습니다.

소프트웨어 2세대(대략 1994년-2008년)

아마 1994년 전후로 등장하기 시작한 2세대 소프트웨어는 이전 세대의 소프트웨어에 비해 보다 많은 기능 수행을 의도하였습니다. 2세대 시스템은 기록이 전자적 형태로 생성되는 현실적 문제에 대응하기 위해 개발되었습니다. 우리는 주로 마이크로소프트사의 오피스 프로그램을 이용하여, 워드프로세스라든지 스프레드시트 등으로 문서를 만들었습니다. 이메일은 사용이 보편화되면서 실질적 문제가 되었습니다. 그러한 소프트웨어 패키지는 통제 시스템을 자동화시킴으로써, 또한 전자기록 자체를 위한 저장 능력을 제공함을 통해 1세대 소프트웨어의 기능성을 강화시켰습니다. 이들 시스템을 통해 관리되는 전자기록은 본질적으로 종이기록과 동일하게 관리되었습니다. 통제 메커니즘은 크게 변화하지 않습니다. 종이에 기반을 둔 전통적인 방식으로 행해지는 등록・분류・처분의 자동화가 여전히 레코드키핑상의 핵심적인 기능이었습니다. 이 외 추가적으로 전자기록은 소프트웨어의 통제 하에 개별적인 기록 저장 공간으로 물리적으로 이전되었습니다.

이러한 시스템은 전문적인 기록관리자를 위한 도구는 아니었으며, 개인용 PC에 탑재되어 업무자에게 기록관리의 초기 작업(등록 및 분류 등)을 일임하였습니다. 여기서 가장 핵심적인 문제는 우리들의 전문적인 기록관리 업무를, 기록관리전문직에게는 익숙하지만 최종이용자에게는 익숙하지 않은 서식을 사용하며, 최종 이용자인 업무자에게 이전시켰다는 점입니다.

우리는 최종이용자인 업무자에게 메타데이터를 생산하고 기록에 대한 상세 템플릿을 채우며, 기록 간의 논리적 연계관계를 생성시키도록 요구했습니다. 하지만 최종이용자는 이것들을 싫어합니다. 이러한 작업들은 업무와는 별도의 부담이

고 시스템은 어려우며, 아울러 짜증나는 서식을 채우기 위해 자신의 업무를 중단시켜 업무의 효율성을 저해한다고 불평합니다.

시스템은 협력적인 문서작성 기능을 제공하는 '문서관리시스템'의 속성을 취하기 시작했습니다. 시스템은 매우 유동적입니다. 레코드키핑 관련 메타데이터 요소들을 한정시키고 어떻게 기록 통제도구가 수행되는지 규정짓는 것은 시스템 운영환경에 전적으로 달려 있었습니다. 아울러 기록관리에 대해 무지하고 최종이용자에게만 초점을 맞추어 기록관리 요건들이 무시되는, 깜짝 놀랄만한 잘못된 시스템 설정이 있을 수도 있었습니다. 우리는 이러한 시스템을 전체 조직에 걸쳐 가동시켜야 하는데, 이것은 상당히 최종이용자의 행위를 변화시키게 됩니다. 이 시스템들을 구입하는 데에는 고비용이 소요되고 가동시키기 위해서는 더욱 더 많은 비용이 필요하며, 여전히 이용자들은 시스템을 따라야 할 필요가 있는지 확신하지 못했습니다.

이러한 시스템의 대부분은 여전히 제자리에서 가동되고 있습니다. 시간이 흘러 이들 기록을 시스템 밖으로 이관할 때, 보다 복잡한 문제에 직면하게 됩니다. 시스템에 저장된 전자기록은 동일한 방식으로 관리될 수 없습니다. 우리가 전자기록 및 기록으로서의 속성을 보호하는데 필요로 했던 연계관계 · 링크 · 기타 관리통제를 제거한다면, 우리는 신뢰할 수 있는 기록을 얻을 수 없게 됩니다. 전자기록은 시스템에 동결된 채 남아있게 됩니다. 우리는 이들 시스템 외부로 전자기록을 가져오는 어떠한 효율적인 수단도 가지고 있지 못했으며, 우리들의 영구보존시스템은 디지털 기록을 적절히 관리하지 못하고 있었습니다.

2008년부터 등장하기 시작한 3세대 소프트웨어

최근 호주에서는 3세대 시스템이 등장하였습니다. 아마 2008년 무렵부터 등장하기 시작한 이 시스템은 이용자에게 초점을 맞추고 있습니다. 이 시스템은 전자문서기록관리시스템이 아닌 콘텐츠관리시스템(Content Management System)이라 불립니다. 이 시스템은 호주에서 개발된 것이 아니라 다른 국가들(주로 호주와는 기록관리 전통이 다른 미국)의 회사에서 개발한 것으로, Documentum, Vignette,

Open Text 등의 제품이 바로 그것입니다. 이 제품들에서 레코드키핑은 부차적인 기능으로 간주되며, 보통 처분절차의 관리나 컴플라이언스 이슈 정도만을 다루고 있습니다. 이 시기는 기록관리 표준이 간행된 이후였으며 따라서 시스템 설계자들은 기록관리 표준을 활용하였습니다. 하지만 시스템의 우선적 주안점은 시스템 이용 상의 이용자 고통을 최소화시키는데 있었습니다.

바로 이것이 위와 같은 시스템이 현재 호주에서 이용되게 된 배경으로 순식간에 일어난 일입니다. 우리는 이러한 시스템을 시장에서 구매하거나 조달할 수 있습니다. 그리고 이들 시스템은 구매해 도입하는데 상당한 경비가 소요되며, 조직 전체에 걸쳐 가동하기 위해서는 수행상의 상당한 고통이 수반됩니다. 이들 시스템의 특징은 전자기록관리를 포함한다는 점입니다. 이 시스템은 가지각색의 소위 구조화되지 않은 기록(외부 처리행위 데이터베이스와 같은)과 함께 이메일・블로그・웹페이지 등과 같은 다양한 유형의 전자기록을 획득하는 방식을 조정할 수 있습니다.

이러한 시스템들은 여전히 바람직한 기록시스템이 아니라는 점을 지적하고 싶습니다. 이들 시스템은 진정한 디지털 기록관리를 향한 여정 상의 시작단계에 있는 것으로, 업무상의 당면 목표나 이용자 인터페이스, 시스템과의 상호작용에 초점을 맞추고 있습니다. 이들 시스템은 정보의 공유 및 교환, 연계 및 재사용 기능을 제공해줍니다. 또한 이들 시스템은 기록관리의 일부 측면인 검색 및 탐색 기능에 보다 많은 주안점이 놓여 있습니다.

이 시스템들은 문서 작성 및 공유, 간행 등과 같은 업무적 필요 충족을 일차적 목적으로 한 것입니다. 이 시스템들은 디지털 세상에 초점을 둔 것이지만, 간혹 옵션 형식으로 종이 관리기능을 지니기도 합니다. 우리가 필요로 하는 레코드키핑 기능은 직접적인 업무처리 기능에 파묻혀 버렸습니다. 이들 시스템에서 생성되는 디지털 기록의 관리는 시스템 설정에 따라 좌우되는 상당 정도의 유연성에 달려 있습니다. 하지만 시스템은 민간기업의 제품으로, 기록은 해당 기업 소프트웨어의 통제 하에 있는 저장환경에서 생산・저장되게 됩니다. 기록 객체는 여전히 사적 시스템 내에 갇혀 있고, 기록의 속성 역시 사적 소프트웨어에 의해 좌우됩니다.

기록은 마이그레이션이나 기타 조치를 위해 시스템 밖으로 쉽게 이전되지 못합니다.

이것으로 충분한가?

기록관리 전문가의 관점에서 볼 때, 소프트웨어 시장은 매우 중요합니다. 소프트웨어 개발회사들은 세계 각국에 걸쳐 있지만, 호주 시장에서는 주로 미국 및 영국 등 영어권 국가들의 기술이 지배적입니다. 최근 호주의 시장은 다국적 기업의 패키지가 지배하고 있지만, 호주 자체의 제품 역시 매우 강력합니다. 시스템 설계 시 시스템 설계자들이 참조하는 것은 레코드키핑 관련 표준이 아닌, 각 정부기관에서 요구하는 기능 요건들입니다. 이에 우리는 시스템 설계자들에게 유럽연합의 MoReq이나 미국의 DoD 5015.2 표준, 영국 TNA의 기능규격 활용을 권고하고 있습니다. 그들은 우리들의 국제표준을 따르고 있지 않습니다.

2세대 소프트웨어를 사용하는지 아니면 3세대 소프트웨어를 사용하는지 여부는, 단순한 기술상의 수준 가늠을 넘어 우리 레코드키핑 전문가에게는 매우 중대한 사안입니다. 기록은 업무자들에 의해 획득되어야 하고 업무작업의 진행방식에 따라 수행되어야 한다는 점은 레코드키핑의 중요한 출발점이지만, 또 다른 문제가 있습니다. 우리는 기록이 필요로 되는 한 유지해야 합니다. 기록이 필요로 되는 기간은 보통 장기간으로, 기록을 생산한 시스템의 수명보다 깁니다. 따라서 우리는 디지털 기록의 장기간에 걸친 유지 문제에 대처해야 합니다. 기록은 다른 환경에서 관리될 필요가 있습니다. 우리가 종이 시대에 기록관리를 위해 다른 기관으로 이관했던 것처럼, 변화된 환경에 맞추어 관리될 필요가 있습니다. 그래서 우리의 현 시스템은 디지털 기록의 생산 및 초기단계의 관리를 우리가 원하는 방식으로 접근해 오고 있지만, 디지털 기록의 유지성(Sustainability) 상의 중대한 문제가 있습니다.

이와 유사하게 우리는 이들 시스템의 범위와 관련된 문제를 지니고 있습니다. 예전에 시스템은 중앙 문서수・발신 체계(Central Correspondence System)를 대체하거나, 사무 환경을 관리하며 소위 구조화되지 않은 문서를 생산하였습니다. 하

지만 다른 조직들에서 생산하는 수많은 유형의 기록들이 존재함을 우리는 알고 있습니다. 이 기록 유형들은 각 조직의 업무적 필요를 통해 생성된 것들로, 전통적인 문서수・발신 영역 외부에서 관리되는 것입니다. 이 시스템은 데이터베이스시스템으로 변모됩니다. 하지만 데이터베이스 시스템은 보통 원시 데이터를 관리하며, 이들 데이터를 사용하여 수행된 행위 내역은 공식적인 기록시스템에서 관리됩니다. 그러나 이들 시스템은 업무시스템과 연동되어 너무 자동화되어 있습니다. 그리고 이들 시스템은 데이터 처리행위를 기록으로 생산할 뿐만 아니라, 시스템의 기능성 내에 기록의 생산을 병합시키고 있습니다. 기록은 업무시스템 내 업무 수행의 일부분으로, 보통 업무시스템 내에서 업무 프로세스가 자동화됨에 따라 기록관리 영역 밖의 업무에 의해 결정됩니다. 이러한 기록들은 우리들의 전문적인 레코드키핑 표준에 따라 생산・관리되지 못합니다.

여기서 현재 무엇이 가능한가?

제가 지금까지 말씀드렸던 시스템들에서의 메타데이터 역할은 우리의 전문직 세계에서 잘 이해되지 못하고 있습니다. 레코드키핑을 위한 메타데이터 요소 표준이 수립된 것에서도 그 실제 수행은 빈약하기 짝이 없습니다. 시스템 설계자들은 메타데이터를 시스템에 적용시키지 않습니다. 메타데이터 적용은 시스템 설계상의 옵션 사항이라고 생각들 하기 때문입니다. 메타데이터 표준의 가장 일반적인 사용은 시스템 내에 메타데이터 요소들이 존재한다는 것을 주장하기 위해 일종의 체크리스트로서 활용하는 것이었습니다. 가장 최근의 메타데이터 표준에서는, 시스템 구축 시점에서의 시스템 설정을 통해 표준을 수행하는 방식을 안내해 주고 있습니다. 하지만 시스템 업체들은 여전히 표준의 적용에 저항하고 있습니다.

그러나 연구자 그룹은 메타데이터 수행의 필요성을 더욱 진전시켜 왔습니다. 현행 시스템들은 여전히 어울리지 않는 종이기록 관리절차를 답습하고 있다는 주장을 받아들이며, 다양한 프로젝트들에서는 디지털 레코드키핑 모색을 위한 새로운 기술을 연구하고 있습니다. 여기에는 기록의 개념과 함께, 어떻게 디지털 세계에서는 레코드키핑 산출물을 이전과 다르게 우리에게 전달할 수 있을까에 대한 재

검토가 포함됩니다.

이러한 프로젝트 중의 하나인 모나쉬대학교의 'Clever Recordkeeping Metadata' 프로젝트는 상호운용성 기술로서 다양한 메타데이터 요소의 번역 개념을 증명하기 위해 메타데이터 레지스트리와 웹서비스를 사용하는 메타데이터 이용 방법을 조사하기 시작하였습니다. 이 프로젝트의 결론은 현재의 메타데이터 표준은 충분히 정미하지 못하고 그 프로세스 역시 충분히 구조화되지 못했으며, 우리들의 통제 도구는 인간 해독을 위해 만들어졌지 자동적 실행을 위해 만들어진 것이 아니라는 것입니다. 이 프로젝트의 일부로서 현재 우리는 레코드키핑 수행 골격으로서의 서비스 기반 아키텍처를 위한 기술적 환경을 연구하고 있습니다. 그러한 환경에서 엄격하게 한정된 메타데이터는 있으면 좋은 그 어떤 것이 아니라, 아키텍처 작업 수행을 위해 없어서는 안 될 필수불가결한 것이라 할 수 있습니다.

이 프로젝트에서는 디지털 레코드키핑 도구 개발을 위해 박차를 가하고 있습니다. 이 프로젝트의 일부 예비적인 연구내용은 아래와 같습니다.

우리가 관리하고 있는 레벨에 대한 고려

종이기록의 경우 통제단위는 기록철 및 시리즈와 같은 집합체였다. 왜냐하면 기록철 및 시리즈는 수기 방식으로 실용적으로 관리될 수 있는 레벨이었기 때문이다. 이에 반해 자동화된 디지털 세계에서는 보다 하위의 레벨에서 관리해야 할 필요가 있다. 서비스는 축적되며, 웹상의 서비스 레지스트리를 통해 활용될 수 있는 수많은 서로 다른 원천으로부터 유래할 수 있다. 이들의 각각은 확인되고 인증될 필요가 있으며, 서비스 결과는 입증될 필요가 있다.

저장위치는 상관이 없다-통제가 중요하다

저장위치의 개념은 아무 상관없게 되었다. 디지털 환경에서 저장은 전혀 문제가 되지 않는다. 기록에 대한 접근 및 권한 통제가 핵심적 사안이다. 저장 환경이 인증되고 필요 시 언제라도 기록이 검색될 수 있으며 기록의 속성이 확보되는 한, 기록은 생산 시점부터 생산조직이 아닌 신뢰할 수 있는 제3자에 의해 저장될 수 있

다. 이것은 특정 시점에 기록이 이관되어야 하는 종래의 개념을 급격히 와해시키고 있다.

자동화는 핵심이다

디지털 세계에서는 모든 기록관리 행위들이 자동화됨과 아울러 자동적으로 기록관리 조치들이 개시되어야 한다(승인 절차를 기반으로). 따라서 우리는 개시 일자, 규칙 및 충족될 필요가 있는 조건을 명시한, 컴퓨터 기반 규칙을 통해 적용되는, 모든 트리거 시점(Trigger Point)을 파악해 구축할 필요가 있다. 디지털 기록의 규모 및 양을 감안할 때, 사람의 직접적인 개입은 거의 불가능하다. 우리는 디지털 기록을 손으로 분류할 수도 없고, 종이 시대에 해왔던 수기 방식의 기술도 수행할 수 없다. 평가 또한 명확하게 정의된 수행 가능한 규칙을 기반으로 컴퓨터를 통해 행해지도록 할 필요가 있다.

새로운 프로세스가 필요하다

새로운 프로세스가 필요하다. 보다 명확한 접근 및 보안 승인뿐만 아니라, 마이그레이션 및 진본성 보증과 같은 새로운 프로세스가 신설되어야 한다. 어떠한 수기 방식의 관여도 디지털 기록의 관리에는 활성화시킬 필요가 없다.

디지털 아카이브는 무엇인가?

우리는 디지털 아카이브의 개념을 구축하는 다양한 방법론들을 생각할 수 있다. 디지털 아카이브는 신뢰성 있는 저장소의 개념에 머물지 않는다. 그것은 물리적일 필요가 없으며 이관에 기반을 두지도 않는다. 디지털 아카이브는 서로 다른 시점에 수많은 시스템이 연계되어 접근승인 계층들을 통해 구현되는 가상의 구조화된 공간이다. 디지털 기록은 외부의 이용자가 해당 기록을 활용하기 위해 30년을 기다릴 필요가 없다. 정보공개법과 같은 장치를 기반으로 열람이 승인된다면, 바로 그 시점에서 공공접근시스템을 통해 제공되어야 한다.

레코드키핑에 대한 사고 형성

우리가 사고하기 시작한 것은 일련의 연속적인 프로세스로부터 조직의 모든 업무 절차의 기반을 이루고 업무절차와 상호작용하는 근본적인 조직적 사회적 골격

에 이르는 레코드키핑을 형성시키는 것이다. 우리는 가능한 한 디지털 기록을 예전의 종이기록처럼 기록철 내지 시리즈 수준에서 통제할 필요가 없다. 디지털 기록은 아이템 단위 내지 보다 낮은 집합체 수준에서 통제할 수 있다.

우리들의 통제도구에 대한 재검토

우리들의 통제도구는 보다 낮은 수준의 디지털 기록 집합체를 관리할 수 있도록 재검토할 필요가 있다. 현재 우리들의 분류체계는 너무 높은 레벨을 다루고 있고, 처분 및 보유지침 역시 효율적으로 자동화시키기에는 너무 광범위한 범주로 설정되어 있다. 따라서 여기에는 사람의 개입이 필요하다. 우리는 메타데이터와 기록객체를 통합하여 관리할 수도 또는 분리해서 관리할 수도 있다. 그것은 수행 상의 선택 문제이다. 그리고 만일 우리가 기록객체에 레코드키핑 절차를 적용시킨다면, 메타데이터에도 이와 동일한 절차를 적용시킬 수 있다.

결론

우리는 최근의 디지털 세계에 부합하는 일련의 레코드키핑 프로세스를 형성하기 위한 여정상의 초기 단계에 있습니다. 하지만 분명한 것은 이러한 변이를 관리하기 위해 필요한 두 가지의 근본적인 영역이 존재한다는 점입니다. 하나는 우리들의 레코드키핑 실무가 핵심적이면서도 대체할 수 없는 구성요소라는 점에 대한 강력한 개념적 이해입니다. 또 다른 하나는 디지털 기록이 필요로 되는 한 오래 지속될 수 있도록 하는, 예전과는 다른 기술적 환경에 적용시킬 수 있는 레코드키핑 메타데이터입니다.

디지털 레코드키핑에 대한 전망은 현재 존재하고 있습니다. 호주에서 우리는 그러한 길의 입구를 보고 있습니다. 그것은 매우 먼 여행의 시작단계라 할 수 입습니다. 우리는 이러한 여정에서 실수를 범할 수도, 잘못된 방향으로 선회할 수도 있으며, 그동안 우리는 너무 관례적이며 유연했다는 점을 알게 되었습니다. 최근 한국에서 추진한 경험은 매우 고무적이며, 저는 한국의 디지털 레코드키핑 프로젝트에서 가능한 한 많은 것들을 배우려고 노력할 것입니다.

감사합니다.

The status quo and prospect of RMS based on metadata

Barbara Reed, Recordkeeping Innovation Pty Ltd, Sydney Australia

Introduction

It is an honour to be asked to present an analysis of the records management arrangements in Australiaandwhatweareaimingtoachieve. I thank the National Archives of Korea for such an honour.

In reading material provided to me and having had the pleasure of working with some of your colleagues in the International Standards work, I know how significant the change from old ways of working to new ways of working has been in the Korean government. We have a lot to learn from your experience.

Background to the experience in Australasia

Recordkeeping is contingent – it depends very much on the time, place and conditions in which it is implemented. How we do and how well we do recordkeeping is culturally driven.

Australia and NewZealand, as my colleague Greg Goulding will further explain, share the experience of a number of English speaking countries, but even there, we are more like some than others. We share most commonality with other countries who have a colonial tradition of being ruled by a remote administration. While our colonial heritage is long past, it has profoundly affected our system of government, the way we do business (and therefore recordkeeping) and the way our public and private sector interact.

Recordkeeping is stronger in the government agencies where the requirements to comply with records legislation and standards can be made mandatory. In the private sector, we can only encourage adoption of good practice and try to make the standards and systems we recommend contribute to the ease of doing business, the efficiency of effort and the protection of the business interests.

Even our government agencies cannot be told what to do. The laws are there, but how each government agency complies with the law is up to them. So in implementation we have a great diversity of sophistication. Some are truly embracing electronic working and integrating recordkeeping into the business that they do. Others are still working in a hybrid environment – one which manages both paper and electronic records concurrently. Still others have no formal systems to manage electronic records and are still relying on printing electronic records into paper for the purpose of management. Our records management software systems and their capabilities reflect this diversity of experience.

Our professional approach

Our past and present reality has affected the way we see records and business interacting – our tradition is very much that records must be part of the business.

Business transactions and records of those business transactions exist in a very tight relationship and in the electronic world, they are often the same thing. For us the relationship between the transaction and the record we keep is core. Records are integral in the doing of business, not something separate to or additional to the business that takes place. Documents and records are the same thing at different levels of formality. The distinction between documents and records only makes sense if we are talking about collaborative construction of documents in formation. Otherwise, every document must be a record, in that it evidences a transaction or an action of some kind.

We understand that records serve many purposes at the same time. They are required for the initial doing of the business, accountability for the business undertaken, protecting rights and obligations of the parties in transactions for as long as those rights and obligations exist, and also for social and cultural purposes. These different purposes for records need to be built into the systems that we use to manage records from before the time any one record is created. The capacity to ensure sustainability and robustness of records so they can exist for archival purposes (that is access by people external to the business transaction that created the record) is critical from the point the records are created.

This approach is not restricted to digital records, but it works well to assist us thinking about digital recordkeeping. Once we know what we need the records for, we can implement systems to create and capture the records. The systems can be integrated, or separate, can be centralised or not, can be business systems or specific records systems, can be systems which manage long term preservation of the records or we can move records between one system and another at different points to achieve our goal of sustainable records for as long as they are required. The technology, and the options we have to manage records, therefore is not constrained – if we know why we are doing something and how long we need to ensure the record will survive, we can implement processes and technology creatively and flexibly to suit the circumstances.

But, if we don't understand the big picture, why we are doing something and the purposes of it, we are constrained. We are threatened by change – technological change being our biggest challenge at the beginning of the twenty first century. We need a very sound conceptual understanding to help us think through the new world. This is critical to working out what we are trying to achieve. If we know that, we can look at the ways we used to do things in the paper world. We can examine

whether they are still valid in the digital world, or if they were a way to do something that worked in the paper world, but we can actually achieve the same end by doing something different in the digital world. And that is the current challenge that we all face – the transition to truly digital thinking is not yet here.

Digital records

Digital records are both very different and very much the same as paper records. The same, in that what we are trying to do remains the same – we are trying to create robust records (evidence of transactions) that will last as long as we need them. But very different, because the formats, techniques and methods of managing these records challenges the way we have done things in the paper world. Our paper transactions have evolved over centuries of practice. The actual record embodies both its content and a lot of implicit information about the processes of managing it. The accumulation of paper documents into sequences of transactions, onto files, and with annotations such as signatures, hand-written instructions on who should do the next thing, or how to respond and the response or the next step in the transaction are often all embedded within the way we manage the physical record itself. In the paper world, the form, the content and the process were often one and the same thing. As we move to digital records, we need to rethink the way we define and capture these components, because they are no longer physically embedded in the data itself.

We need to understand that records are a construct – they comprise different types of data. Firstly, there is the content of records – these are the data that make up the transaction or action being done. The data that we choose to manage as records doesn't exist as records – unless we consciously decide that we are going to manage it as a record, it is just data, and managed as data – changing to reflect the latest data values, managed to reduce redundancy, so stored once with minimum duplication, managed with systems that apply data management techniques. Once we decide we need to manage a set of defined data as a record, then we need to change some of its characteristics. We need to know that the data in the record is the same as it was when action was taken, that it is not updated, that it is linked to the time and sequence of actions that took place. This is the group of data that took part in an initial business transaction of some kind and was captured as a record. We are currently calling this the 'record object'.

But there is much more to having a record than just having the initial content. We need to capture the record object into its business context – who created it, when, in the course of what action? This data about the initial record at capture is the beginning of a set of data we need to identify and manage in conjunction with the 'record object'. We have started to call this the 'point of capture recordkeeping metadata' (in some systems this might be called 'document profile' metadata).

Even still, this is not enough to ensure that we have reliable, authentic records. We need to be able to prove that the record is what we say it is. We need to be able to show that it has been managed appropriately, that the processes of recordkeeping have been applied to the record in order to protect it, that we can prove how the record has been managed over time. We have started to call this 'recordkeeping process metadata'. The recordkeeping process metadata keeps accumulating as a part of managing the record for as long as the record object itself exists.

But it is even more complex than that. We want to express records metadata as sets of relationships. Relationships between the record object, the people who interact with the record object (in creating and managing it).

Once we start to think in terms of record objects and recordkeeping metadata and relationships between them, we have entered a new world of recordkeeping. The possibilities of creating different ways of capturing and managing digital records open up. We can question much of the inherited wisdom of recordkeeping that came bound up with the way we did things in the paper world.

So we are in the new world of defining records by using recordkeeping metadata. This is now (or will be) our core business. And it will be embedded into software that will automatically create the metadata. But we need to constantly challenge ourselves. What do we really need, what is extra, why are we insisting on some data about the record (metadata) and not others? How can we test whether the systems that are being built will create and manage this metadata appropriately? Have we really defined our needs?

We also need to understand that every profession/discipline has its own requirements for metadata. They are all valid for the purposes they address to do to ensure the requirements of specific communities of interest are met. We cannot talk about metadata without a qualifier in front – what type of metadata. Just talking about metadata itself will cause confusion. Our business is recordkeeping metadata. That metadata is continuously associated with an identifiable set or package of data which needs to be managed as record. This is different to metadata that is focussed on health outcomes or metadata that is focussed on search and retrieval. There are overlaps, of course, but our discipline is recordkeeping. That is our area of specialty.

What have we done in this area?

Understanding that recordkeeping metadata will be the key to the digital future has been with us for some time. The first projects that focussed on recordkeeping metadata were research projects into electronic recordkeeping in the 1990s. The three projects of note were:

○ The University of Pittsburgh's (US) project 'Functional Requirements

for Evidence in Recordkeeping'
- ○ The University of British Columbia's (Canada) project The Preservation of the Integrity of Electronic Records
- ○ Monash University's (Australia) project 'Recordkeeping Metadata Standards for Managing and Accessing Information Resources in Networked Environments Over Time for Government, Commerce, Social and Cultural Purposes People and Partners'.

Within Australia we have been very active in trying to take these frameworks forward. We have actively worked with the international community to develop and publish 2 international standards, the ISO 23081 series, which explain requirements for recordkeeping metadata at a high level. Within individual jurisdictions, recordkeeping metadata element sets have been specified and published. In some jurisdictions, such as NewZealand, standards establishing minimum requirements for organisations implementing recordkeeping metadata have been defined.

More recently, I have been chairing a project which is defining an Australian Standard Recordkeeping Metadata Element set. In that set, we are following both the international standards requirements and our own recordkeeping tradition. The metadata set we are developing is aimed at being able to be implemented in any recordkeeping environment – in current records systems, in cross organisational records systems, in business systems, in systems dedicated to the management of records, and in archives systems. We are hoping that using this metadata we can make connections with all the different systems environments in which records can be created and managed, and therefore create the infrastructure for ensuring that records can be moved between systems as they need to be.

Standards are crucial to enabling us to take these understandings much further. We have to be able to define for the people who are building and designing systems what metadata we require and why we require it, in order to build systems that will automatically create and appropriately manage the recordkeeping metadata across time.

While our community is very slowly coming to understand the importance of recordkeeping metadata, we need to keep talking about it. Recordkeeping metadata is complex. It is conceptual. It is translating the framework of what we want to achieve into a new language suited to systems implementation. It is critical to the future of recordkeeping in the digital world.

So what do our standards look like?

For the Australasian (Australian and NewZealand) environment, we have used the understandings of the Australian research projects, which have also been embedded in the international standards. The key diagram is

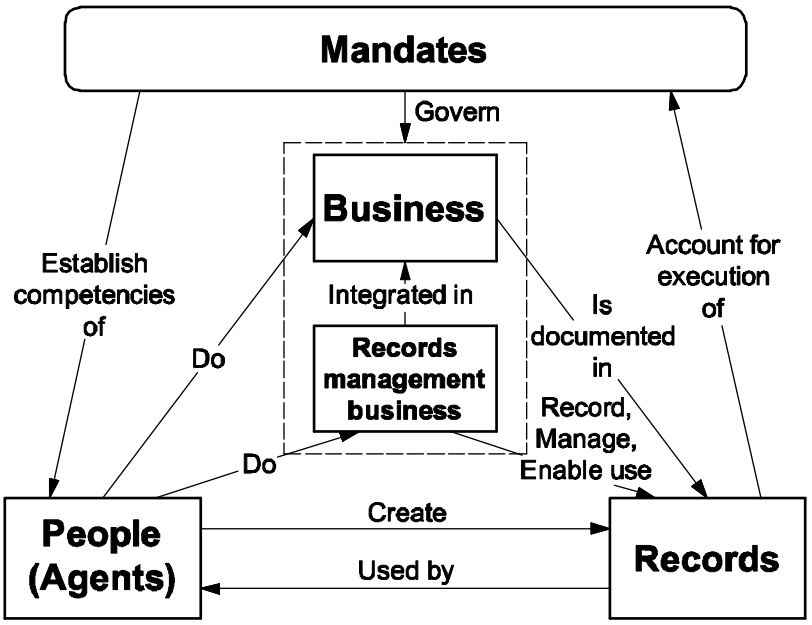


Figure 1 — Main entities and their relationships[1)]

What is happening in the real world, where are the implementations?

In Australasia we have been rethinking recordkeeping for some time. We have been aware since the early 1990s that our ways of doing things in the paper world are not sustainable and must change. The work in developing standards for recordkeeping is one area we have been active. Simiarly most of our public records legislation has been updated and changed. Adoption of standards for recordkeeping has been encouraged and we have been trying to find the right regulatory and business lever to change the way our organisations approach recordkeeping. And we have had some success in doing this.

In the main we have been focussed on capturing records into systems that will manage them as records from the time they are created. The thinking is that once that is achieved, we have the capacity to identify and implement systems that will sustain the records through time. These software solutions have been focussed on the working environment – the environment where business is taking place. This is the domain of records management and of Electronic Document and Records Management Systems (EDRMS). We are on a journey, which is really at its beginning.

First generation of software (perhaps 1985-1994)

---

[1] *Figure 1 from ISO 23081-2, Information and documentation—Managing metadata for records—Part 2: Conceptual and implementation issues, original diagram from Monash University's research project 'Recordkeeping Metadata Standards for Managing and Accessing Information Resources in Networked Environments Over Time for Government, Commerce, Social and Cultural Purposes People and Partners'.*

The first generation of systems used to manage records were really aimed at providing electronic metadata systems to automate our controls to manage paper records. We essentially automated our paper ways of managing records. The records still existed in paper files, which were aggregated into paper series. The electronic metadata system helped us to manage things like the registration, classification, disposal and location of these paper records. These systems derived from our local practices and were home grown. They included systems such as TRIM and RecFind in the Australasian environment and Provenance from the Canadian environment.

One of the interesting things about these systems is that they replaced the previous paper forms of registers and indexes. These paper forms – registers and indexes – have traditionally been regarded as being of continuing value – archival records. We needed them to explain and to facilitate use of records once those records ceased to be used in current offices. Conceptually the profession did not make the analogy. And our archival organisations did not have the capacity to take custody of electronic records in any form, so at the time the records came to be transferred to archives organisations, we discarded the systems and used paper based outputs in the form of lists to accompany transfers of records. This meant that we ignored a whole generation of key control tools and divorced the paper records from their critical controls at the point of transfer to archives. We enforced the life cycle view creating a radical break between the operations of current records in organisations and archival records in archives.

The second generation of software (perhaps 1994-about 2008)

The next generation of software, perhaps from about 1994 onwards, attempted to do more than the previous generation. They were developed to deal with the emerging reality that records were being created in electronic forms. In the main we were in the era of Microsoft, where office documents were created in word processing packages, spreadsheets and the like. Email became a real issue in the time these systems were (aare) around. Such software packages enhanced the functionality of the first generation of software by automating the control systems, but also by providing a storage capacity for electronic records themselves. The electronic records managed by these systems were essentially treated as paper equivalents. The control mechanisms did not change much – still core recordkeeping functions such as registration, classification, automation of disposal delivered in essentially traditional paper based ways – but in addition, electronic records were physically moved to a proprietary records store under the software control.

Rather than these systems being tools for professional records managers, the software was rolled to every desktop, requiring individuals in the business to apply the initial records controls (such as registration and classification). The common criticism is that we transferred our professional concerns for controls to the end users using forms that were familiar to the professionals but not familiar to the end users. We asked the end users to create the point of capture metadata, to fill in templates with details about the records and to

make the logical connections between records. The end users did not like them. They argued that it was burdensome, that the systems were not intuitive, and that they constrained end user efficiency by making them pause in their work to complete tedious forms.

The systems began to take on the characteristics of 'document management systems' providing collaborative authoring tools. The systems are/were highly flexible. It was largely up to the implementing environment to define the metadata elements relevant to recordkeeping and to define how the records control tools would be implemented. And there were some spectacularly bad configurations done by people who didn't always understand records requirements, but were totally focussed on the end user, to the detriment of records requirements. And this is understandable – we had to roll these systems out across the whole organisation, which involved a high degree of change in behaviour of the end user. They were expensive to buy and even more expensive to implement, and still the users weren't really convinced that they needed to follow the systems.

Many of these systems are still in place and still being implemented. When the time came to archive these records and transfer them outside the system, we ran into more problems. The electronic records stored in the system could not be treated in the same way – if we removed the electronic records, the links, relationships and those management controls that we need to protect the characteristics of records – we ceased to have reliable records. The electronic records were frozen into the system. We had no effective means of getting them out of these systems, and our archival systems were still not enabled to deal with digital records properly.

Third generation – emerging around 2008

Very recently, we have seen the evolution of third generation systems. These systems, perhaps dating from around 2008, are focussed on the user. They are often not called EDRMS but content management systems. They are not home grown, but developed by international companies (often from the USA which has a different tradition of recordkeeping to our local conditions). The products in this arena are those like Documentum, Vignette, Open Text etc. The recordkeeping was a secondary notion and often dismissively treated as simply a compliance issue and really only required for the management of the disposal process. By this time the standards for records management were published and available to system designers, but the focus of the systems was very much to minimise user pain in their interaction with the systems.

This is really where the software systems of today are based. It is a volatile time for such systems. We can see it in the marketplace, with consolidation, take overs, buy outs etc. And in implementation, these systems are still very expensive to buy using purchasing models of individual licences and requiring a great deal of implementation pain as we need to embed them in the organisation. The features of these systems include genuine electronic

records management (if configured with thought – not always done). They were geared to identifying ways of capturing the myriad of so called unstructured records (those outside transactional databases) and evolving to cater for ever more formats – email, blogs, web pages etc.

I would argue that these systems are not yet good records systems. They begin the journey to true digital records management, and are focussed on the front end, the user interface and interaction with the system. They provide capacity to share information, exchange information, link and reuse information. They are focussed more on the immediate requirements of retrieval and searching, priviledging one aspect of managing records.

These systems are primarily address a business need – document construction, sharing, publishing etc. They are focussed on the digital world, and often have paper management as an optional add on. The recordkeeping controls we require are buried under the immediate business functionality, and this is great. How well those controls work is still subject to considerable flexibility depending on the configuration decisions made. However the systems are proprietary – records are either created or placed into a storage environment under control of the software. The increasingly complex links are made within the software. The record objects are still locked into a proprietary system and many of the complex links required to ensure their records characteristics are locked into the proprietary software. Records cannot be easily moved out of the systems, either to manage migration to the next generation of software or for

Is this good enough?

From a professional perspective the development in the software market is very important. Software developers come from a variety of countries but for the Australasian market we tend to use the technology that comes from the US or UK–other English speaking countries. Our local products are still quite strong, although the leadering package has recently been purcheased by a muti-national vendor. What drives these software designers is not our professional or international standards, but rather the mandatory functional requirements issued by various government agencies. So we find that these designers are addressing requirements such as MoReq, or DoD 5015.2 or TNA specifications. They are not complying with our international standards.

Whether we are dealing with second or third generation software packages for recordkeeping there are bigger issues than simply those at the functional level of getting the systems to work. While this is a crucial initial starting point – records must be captured and this has to be done by users and be implemented in ways that suit the ongoing functionality of the businesses work - there are other issues. We need to ensure that records will last as long as they are required. This is often for long periods of time – they need to be able to last longer than the systems that create them, so we must address the issues of sustaining digital records over time. The records may need to be managed in different environment – it may be that the records

have to change control, just as in the physical world we transferred records to other agencies through archival arrangements. So, while our current systems may be approaching the creation and initial management of digital records in ways that approximate what we might want, there are crucial issues of sustainability that are not being thought about.

Similarly we have issues with scope of these systems. Previously the systems were replacing the central correspondence systems, or managing the office environment, the so-called unstructured documents. But we always knew that there were many other records created in organisations. These tended to be the records of defined organisational practices which were managed outside the traditional correspondence sections. They transitioned into database systems but these database systems were often for managing the raw data, and the actions that were taken using the data were still managed in the formal records systems. But now these systems, too, are being automated as line of business systems. And increasingly these systems are not only creating data transactions as records, but actually incorporating the creation of records within their functionality. Records are a component part of doing the business in business systems – and often determined by the business remote to the records management as the work process itself is automated in business systems. These records are not being created or managed according to our professional standards for recordkeeping.

What is possible beyond the practical here and now?

The role of metadata in the types of systems I have described is not well understood by our profession. Where metadata element set standards for recordkeeping have been established implementation has been reasonably poor. The system designers haven't adopted them – they seem to think that they are an optional add on. The most common use of the standards has been as a kind of checklist to assert that the elements exist within systems. With the latest work on metadata standards, advice is being prepared on how to implement the standards within the configuration options at the time of setting up some of these systems. But the vendor community is still broadly resistant to the requirements, chosing to interpret the existence of more than one standard as an indication of lack of clarity within the community, thus reinforcing the need for international standardisation.

Within the research arena however, the implementation options have been taken much further. Accepting the proposition that the current systems are still largely replicating paper processes which are not sustainable, many such projects have sought out new and emerging technologies to explore digital recordkeeping. This involves rethinking the notion of records and how the digital world might enable us to deliver recordkeeping outcomes quite differently.

One of the projects, the Monash University 'Clever Recordkeeping Metadata project' began to explore the use of metadata as an interoperability technique to prove the concept of translation of different metadata elements using

metadata registries and web services. The conclusions of that project were that our metadata standards were not rigorous enough, that our processes are not structured enough, our control tools are currently written for human interpretation and not suited to automatic execution. As a part of this project we explored the emerging technical environment of service oriented architectures as a framework for delivering recordkeeping. In that environment, metadata which is rigourously defined is not something that it is nice to have, but an essential prerequisite to making the architectures work.

That project has spurred thought on what digital recordkeeping might involve. Some of the preliminary areas that have arisen are:

Consideration of the level at which we are managing.
In the paper world we put our controls on aggregates (files, series) because that was the management level that was practical in a manual world. In an automated digital world, we can manage (and in fact, have to manage) at much smaller layers of aggregation. Services are cumulative and can be sourced from many different origins made available through service registries on the web. Each of these needs to be identified, authenticated and the results of the service validated.

Storage location is irrelevant – control is important
The notion of a storage location becomes something that is a convenience. In a digital environment, storage doesn't matter at all. Control of access and authority to act on the record becomes the key. Records can be stored with a trusted third party which is different to the creating organisation from the time of the creation of the record – it doesn't matter as long as we can authenticate the environment and retrieve the record when we need it, with its characteristics ensured. This radically decomposes the notion of the archival threashold across which records must pass at some time.

Automation is the key
In the digital future all prompts to action should be automated, enabling processes to trigger automatically (subject to an authorisation process). So, we need to ensure we have identified and built in all the trigger points - those dates, rules and conditions that need to be satisfied and which can be found and applied by machine based rules. With the scale and sheer volume of digital records, manual intervention is not possible. We cannot classify manually, we cannot do archival description in the same handcrafted ways that suited the paper world, our appriasal decisions need to be able to be implemented by a computer based on clearly defined and implementable rules.

Some new processes will be necessary
New processes are necessary – both a much clearer articulation of access and security permissions, but also a new process of migration and authenticity assurance. No manual intervention should be needed to activate many of the desired controls.

What is a digital archive?

We can think of different ways of constructing the notion of digital archives. It doesn't have to be restricted to the concept of a trusted repository. It is not necessarily physical and based on transfer, but virtual and constructed through layers of access permissions connecting many different systems at different points. Records don't have to wait for 30 years to become available to external public users, if they have been released under a form of accelerated access such as Freedom of Information, they should be in the public access system (archival system?) from that time.

Transformation of thinking about recordkeeping

What we can begin to think about is a transformation of recordkeeping from a set of linear or sequential processes to a fundamental organisational and social framework underlying and interconnecting with all points of organisational process. We possibly don't need a digital equivalent of files or series – we can manage at individual item or lower levels of aggregation.

Rethinking our control tools

Our control tools need to be rethought to enable management at these lower levels of decomposition. Our classification schemes are too high level for these requirements, and similarly, our disposal and retention authorities are far too broad to be effectively automated – at the moment they need human translation. We can manage metadata and the record object separately – or together. It becomes an implementation choice. And if we can apply recordkeeping practices to the record object, we can also apply those same processes to the metadata itself.

Conclusion

We are only at the very beginning of a journey to transform our recordkeeping into a set of practices suited for the emerging digital world. But what seems very clear is that there are two fundamental building blocks that we need in place to manage the transition. One is a strong conceptual understanding on which to pin our recordkeeping practice becomes a critical and irreplaceable component. The second is a set of recordkeeping metadata which we can, with confidence, deploy in different technological environments to ensure that digital records are sustained as long as required.

The prospects of digital recordkeeping are exiting. In Australasia we can just see the beginnings of the path. It is a journey that is only just beginning. And we will make mistakes, take wrong turnings and find we have been too prescriptive or too flexible. The experience of your initiatives in Korea are inspiring and I shall be seeking to find out as much as possible of your successful digital recordkeeping projects.