

공공기록관리정책 평가모형 개발을 위한 국제기준 및 평가 사례 분석

An Analysis of International Standards and Evaluation Practices for the Development of an Evaluation Model for Public Records Management Policy

서지인(Jiin Seo)¹, 노지현(Jee-Hyun Rho)²

E-mail: archivething@naver.com, jhrho@pusan.ac.kr



¹ 제 1저자 부산대학교 일반대학원 문헌정보학과 기록관리학전공 박사과정 수료
² 교신저자 부산대학교 사회과학대학 문헌정보학과 교수

논문접수 2026-07-17
최초심사 2026-07-23
게재확정 2026-08-10

ORCID

Jiin Seo
https://orcid.org/0000-0002-3448-8927
Jee-Hyun Rho
https://orcid.org/0000-0002-2760-8732

초 록

이 연구는 공공기록관리정책 평가가 기관별 실적 확인과 법령 준수 점검에 머무르고 있다는 문제의식에서 출발하여 향후 공공기록관리정책 평가모형을 개발하는 데 필요한 참조 준거를 도출하기 위해 국제기준과 해외 사례를 분석하는 데 목적이 있다. 평가모형이 성립하기 위해서는 평가의 원칙, 세부 평가기준, 성과관리와 환류를 포괄하는 평가체계가 갖추어져야 한다. 이에 따라 이 연구에서는 국제적으로 표준으로 인정받는 OECD DAC 평가기준, UNEG 평가 규범 및 표준, 세계은행 IEG 평가 원칙을 검토하고, 기록관리정책의 집행성과를 보여주는 평가 사례로 미국 NARA, 영국 TNA, 호주 NAA의 평가제도를 비교 분석하였다. 기록관리정책의 평가제도는 평가의 목적과 성격, 구성요소, 평가방식, 활용과 한계라는 공통 분석기준을 설정하여 분석하였다. 분석 결과, 정책목표와 평가기준의 정합성, 정책논리와 성과경로의 명확성, 증거기반 평가자료의 확보, 평가절차의 독립성·투명성·책임성, 디지털 기록환경과 장기 지속가능성, 평가결과의 환류라는 여섯 가지 설계원칙을 사전평가와 사후평가의 관점에서 도출하였다. 이 연구는 공공기록관리정책 평가를 단순 행정점검에서 통합적 정책평가로 확장하고 향후 평가모형을 개발하기 위한 이론적·제도적 참조 기반을 제공한다.

ABSTRACT

Starting from the recognition that public records management policy evaluation has remained largely limited to verifying institutional performance and checking legal compliance, this study analyzes international standards and overseas cases to derive reference criteria for developing a public records management policy evaluation model. Proceeding from the view that a viable evaluation model must encompass evaluation principles, detailed assessment criteria, and a performance management and feedback system, the study reviews the Organisation for Economic Co-operation and Development's Development Assistance Committee (OECD DAC) criteria, the United Nations Evaluation Group (UNEG) Norms and Standards, and World Bank Independent Evaluation Group (IEG) principles, and examines the evaluation framework of the National Archives and Records Administration (NARA) of the United States, the National Archives of the United Kingdom, and the National Archives of Australia (NAA) using a common analytical framework consisting of evaluation objectives and characteristics, components, evaluation methods, utilization, and limitations. The analysis identifies six principles: policy-goal alignment, policy logic, evidence-based data, procedural accountability, digital sustainability, and feedback. This study expands public records management policy evaluation beyond administrative inspection toward integrated policy evaluation and provides a theoretical and institutional reference base for developing future evaluation models.

Keywords: 공공기록관리정책, 정책평가, 평가모형, 평가기준, 평가표준

Public Records Management Policy, Policy Evaluation, Evaluation Model, Evaluation Criteria, Evaluation Standard

© 한국기록관리학회

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

<https://jksarm.koar.kr>

www.kci.go.kr

1. 서 론

1.1 연구배경 및 목적

정부는 어떠한 문제를 해결하거나 목표를 달성하기 위하여 분야별 정책을 수립하고 발표한다. 대체로 새로운 정부가 출범하면 국가비전이라는 간결한 문구를 통해 정부의 국정방향을 제시하고, 그 기초를 달성하기 위하여 하위로 국정목표와 목표별로 세부적인 국정과제를 설정한다. 국정과제 추진에서 더 나아가 국가비전 달성을 위하여 주관부처인 중앙행정기관은 법률·제도·예산 등 여건을 검토하여 정책을 설계한다. 즉, 정책은 정부 및 행정기관이 문제해결 혹은 목표 달성을 위해 모든 행정적인 요인을 고려하여 산출된 결과물이라 할 수 있다.

정책평가는 정책의 설계단계부터 종료단계까지를 아울러 해당 정책을 통해 이루고자 하는 목표를 달성하였는가를 판단하는 행위이다. 정책평가를 위해서는 대상, 시점 및 주체 등 목적에 맞게 평가절차와 모형을 설계할 수 있으며, 그 구성은 평가의 의도에 따라 달라질 수 있다. 그러나 모든 정책평가의 목표는 정책이 문제해결 및 목표달성에 부합하는 정도를 측정하고 미비점을 파악하여 향후 더 나은 정책을 위한 기반을 마련하는 데에 있다.

대한민국 정부는 매년 정부업무평가를 통해 중앙행정기관 뿐만 아니라 지방자치단체, 공공기관에서 수행하는 정책, 사업, 업무에 대한 계획 수립, 집행과정 및 결과 등을 점검, 분석하고 평가를 실시한다. 기록관리 평가 역시 정부업무평가에 포함되어 있으며 행정안전부 주관으로 연 1회 자체평가를 수행하고 있다. 국가기록원의 「2025년 기록관리 평가 계획」에 따르면, 기록관리 평가는 기록관리 기본계획, 전담조직·인력, 생산현황 통보, 이관, 평가·폐기, 공개재분류 등 기관의 관리기반과 업무이행 정도를 중심으로 이루어진다(국가기록원, 2024). 그러나 업무기반·업무추진·기타로 구분한 3개의 영역과 10~16개로 구성된 평가항목이 대부분 전년도 업무추진 여부 및 성과 중심으로 이루어져 있고, 항목별 평가결과는 각 기관의 실적평가에 그치고 있다. 다시 말해 현재의 기록관리 평가는 주로 기관의 기록관리 운영 실태 및 업무 수행 수준을 측정하는 데 초점을 두고 있어, “기록관리정책 자체가 어떤 목표를 가지고, 어떤 투입·과정·산출·결과를 만들어냈는지”를 거시적인 관점에서 종합적·체계적으로 평가하는 데에는 한계가 있다.

2024년 「공공기록물 관리에 관한 법률」 일부개정은 기록관리 체계의 제도적 정비와 운영 효율성 제고라는 측면에서 한국의 기록관리정책을 재구축하기 위한 중요한 계기로 볼 수 있다. 그러나 해당 개정이 실제 기록관리 현장에 미치는 영향과 공공기관의 기록 생산·관리·보존 과정에서 발생할 수 있는 제도적 한계, 그리고 국민의 알 권리와 행정 투명성 보장이라는 기록관리의 본질적 가치에 부합하는지에 대해서는 보다 면밀한 검토가 필요하다. 특히 개정 내용이 행정 편의나 특정 기관의 관점에 치우치지 않고, 기록관리의 전문성, 공공성, 책임성 및 이용자 권리 보장이라는 기준에 따라 실효성 있게 작동하는지를 평가할 필요가 있다. 이러한 평가는 개정 법률의 정책적 타당성과 현장 적용 가능성을 종합적으로 검토하고, 향후 기록관리 제도가 민주적 책임성과 공공성을 강화하는 방향으로 발전할 수 있도록 하는 중요한 기반이 될 수 있다.

따라서 이 연구에서는 기록관리정책의 설계에서부터 결과까지 아우르는 거시적인 시각에서의 정책평가가 필요하다는 문제의식에서 출발하여, 정책평가를 위한 국제기준과 주요국의 평가사례를 분석하고 평가모형 개발을 위한 설계원칙을 도출하고자 한다. 정책평가에 있어 과정평가는 정책 집행 중의 운영 문제와 개선 가능성을 확인하는 데 중요한 의미를 가지지만, 이 연구는 기록관리정책의 형성 단계와 결과 단계에 초점을 두고 사전평가와 사후평가의 분석에 집중하였다.

1.2 연구내용 및 방법

이 연구에서는 공공기록관리정책 평가모형과 관련한 국제기준과 주요국의 평가 사례를 비교·분석하고, 이를

토대로 향후 평가모형 개발을 위한 설계원칙을 도출하였다. 분석 대상 자료는 해당 기관의 웹사이트에 공개된 자료를 활용하였다. 분석의 주요 내용은 다음과 같다. 먼저 정책평가와 기록관리정책에 관한 선행연구를 검토하여 공공기록관리정책 평가의 필요성과 분석 범위를 정리하였다. 정책평가 분야의 선행연구를 통해 사전평가와 사후평가의 개념과 평가시점별 정책목표·수단·집행·산출·성과·영향·환류의 분석 틀을 도출하였다.

다음으로, 분석 대상인 국제기준과 해외 기록관리 평가 사례를 선정하였다. 평가모형을 개발하기 위해서는 무엇을 기준으로 판단할 것인가에 관한 세부 평가기준, 어떤 조건에서 평가가 정당성과 신뢰성을 갖는가에 관한 평가규범, 그리고 성과를 어떻게 입증하고 환류할 것인가에 관한 평가체계가 필요하다. 본 연구에서는 공공기록관리정책 평가모형 설계를 위한 참조 기준으로 국제적으로 널리 활용되는 세 가지 평가기준을 선정하였다. 첫째, OECD DAC 평가기준은 평가의 기본 판단기준과 평가질문을 제시한다. 둘째, 유엔평가그룹(UNEG) 평가 규범 및 표준은 평가 운영의 독립성·공정성·신뢰성·윤리성을 확보하기 위한 원칙을 제공한다. 마지막으로, 세계은행 독립평가그룹(Independent Evaluation Group, IEG) 평가 원칙 및 품질기준은 결과사슬과 성과관리, 증거기반 학습을 통해 정책개입의 논리와 환류를 설명하는 평가체계를 제공한다. 이 세 가지 평가기준은 평가모형의 개발에서 평가기준, 평가 운영 원칙, 성과 환류 체계라는 측면에서 상호보완적인 참조 준거로 활용될 수 있다.

해외 사례로는 주요국 국가기록관리기관에서 운영하는 기록관리 평가제도를 선정하였다. 구체적으로, 미국 국립기록관리청(National Archives and Records Administration, 이하 NARA)의 기록관리 자체평가(Records Management Self-Assessment, RMSA) 및 고위 기록관리 책임자(Senior Agency Official for Records Management, SAORM) 보고 체계, 영국 국립기록보존소(The National Archives, 이하 TNA)의 정보관리평가(Information Management Assessment, IMA), 호주 국가기록원(National Archives of Australia, 이하 NAA)의 디지털 역량진단(Check-up Digital) 및 디지털 연속성(Digital Continuity) 정책을 사례로 선정하여 비교분석하였다. 선정기준은 다음과 같다. NARA는 연례 자기평가와 고위책임자 보고를 결합한 중앙집중형 보고·평가 체계이며, TNA는 성숙도와 위험관리, 개선목표 설정을 중심으로 한 진단형 평가체계를 갖고 있다. NAA는 디지털 자원관리에 집중한 디지털 정보관리와 디지털 연속성을 중심으로 한 자기진단형 평가체계에 해당하여 분석대상으로 선정하였다. 이들 세 사례는 기록관리정책에 대한 직접적인 평가는 아니지만, 각각 보고·평가의 조직화 방식, 기록관리 역량 진단의 방법, 디지털 환경에 대한 대응이라는 서로 다른 측면에서 공공기록관리정책 평가모형 설계에 활용할 수 있는 중요한 실천적 참조요소를 제공한다.

국제기준과 해외 사례는 목적과 성격, 운영방식이 서로 다르기 때문에 이를 분석하기 위해서는 동일한 항목으로 분석할 공통의 틀이 필요하다. 이에 이 연구는 평가의 목적과 성격, 구성요소(원칙·기준·도구), 평가방식, 활용과 한계라는 네 가지 분석기준을 설정하고, 이를 바탕으로 국제기준과 해외 사례를 검토하였다. 이러한 분석기준은 정책평가 관련 선행연구와 국제 평가기준에서 공통적으로 강조되는 요소를 재구성한 것으로, 평가제도를 ‘평가의 목적은 무엇인가—무엇을 활용하여 평가하는가—어떠한 방식으로 평가하는가—평가결과를 어떻게 활용하고 해석하는가’라는 연속적 구조로 비교·분석하기 위한 틀이다. 이를 통해 도출된 분석결과를 종합하여 공공기록관리정책 평가모형 설계시 고려해야 할 사항을 사전평가와 사후평가의 관점에서 도출하고, 각 고려사항에서 참조할 수 있는 국제기준과 해외 사례를 함께 제시하였다.

1.3 선행연구

이 연구에서는 정책평가과 기록관리 정책을 중심으로 선행연구를 분석하였다.

먼저, 정책평가에 관한 선행연구에서는 정책평가를 정책의 성과를 사후적으로 확인하는 절차에 한정하지 않고, 정책의 필요성, 설계, 집행, 성과 및 영향을 종합적으로 분석하는 체계적 과정으로 이해해 왔다. Vedung(1997)은

정책평가를 공공개입의 내용, 집행, 산출, 결과를 일정한 기준에 따라 평가하고 그 결과를 향후 정책적 실천에 활용하는 과정으로 보았다. 이러한 관점은 정책평가가 공공정책의 결과를 판단하는 것에 머무르지 않고, 정책 개선과 환류를 위한 실천적 지식을 제공한다는 점을 강조한다. Weiss(1998)는 평가가 정책결정에 대한 직접적인 판단 근거를 제공할 뿐만 아니라 정책문제에 대한 이해를 확장하고 정책학습과 환류를 가능하게 하는 지식 활용 과정이라고 보았다. 이는 평가 결과가 단순히 정책의 성공 또는 실패를 판정하는 데 사용되는 것이 아니라 정책문제에 대한 사회적 이해를 형성하고 장기적으로 정책담론과 제도 변화를 이끄는 역할을 한다는 점을 의미한다.

Rossi et al.(2019)는 평가를 사회적 개입 프로그램의 필요성, 설계, 집행, 효과성 및 효율성을 체계적으로 판단하는 응용사회과학적 활동으로 설명하였다. 이들의 논의는 정책평가가 단순한 행정적 점검이나 실적 확인이 아니라 정책 또는 프로그램이 왜 필요한지, 어떻게 설계되었는지, 실제로 어떻게 집행되었는지, 그리고 의도한 효과를 달성하였는지를 경험적 근거에 따라 판단하는 분석 과정임을 보여준다.

다음으로, 기록관리정책에 관한 국내의 연구는 기록관리정책의 개념과 범주를 정립하려는 논의에서 출발하여 국가 및 공공기록관리체계의 제도적 발전, 참여정부 시기의 기록관리혁신과 지방 공공기록관리, 디지털 전환기 공공기록정책으로 확장되어 왔다. 광동철(2001)은 기록관리정책을 정책학적 관점에서 분석하면서 정책목표와 수단의 부적절성, 정책 참여자 간 갈등, 기록관리 체계의 모호성, 집행환경의 미흡을 지적하고, 기록관리정책의 발전전략으로 정책목표와 수단의 적절성 확보, 정책 참여자 간 상호작용 촉진, 기록관리 체계 정비, 정책환경 개선을 제시하였다. 이 연구는 기록관리정책을 독립적인 정책연구 대상으로 설정한 초기 연구였다고 할 수 있다.

이후 국가 및 공공기록관리정책의 역사적 전개와 제도적 과제를 분석한 연구가 발표되었다. 서혜란(2009)은 정부 수립 이후 한국 공공기록관리정책의 변천을 도입기, 준비기, 발전기, 전환기 등으로 구분하고 전환기 이후 지속적 발전을 위해 쌍방향적 네트워크형 발전모델이 필요하다고 보았다. 이영학(2009)은 1999년 기록물관리 법제 도입 이후 국가기록관리제도가 발전했음에도 현실의 기록관리 수준은 여전히 낮다고 평가하면서 중앙기록물관리기관의 정치적 중립성과 독립성, 기록관리 거버넌스, 기록관리시스템의 확산, 다양한 아카이브의 설립, 기록물관리 전문요원 배치를 향후 국가기록관리정책의 핵심 과제로 제시하였다.

이원규(2010)는 지방 공공기록관리의 정책과제를 제기함으로써 기록관리정책 논의를 중앙정부 중심에서 지방 자치단체와 지역 기록관리체계로 확장하였다. 이영학(2012)은 참여정부 기록관리정책의 특성을 분석하면서 대통령비서실, 기록관리혁신전문위원회, 국가기록원 기록관리혁신단 등 정책 추진 주체의 역할, 시민사회와 학계 전문가가 참여한 거버넌스, 기록물관리 전문요원 배치, 공공기록물법과 대통령기록물법의 정비를 주요 성과로 평가하였다.

최근에는 디지털 전환에 따른 공공기록정책의 패러다임 전환이 중요한 쟁점으로 제기되고 있다. 설문원(2020)은 기존 공공기록정책이 행정정보시스템에서 생산되는 디지털정보를 관리하기에 미흡하다고 지적하고, 영국과 호주 사례를 통해 정보관리로의 확장, 디지털 연속성 관리, 기록자산으로서 정보의 관리, 정보거버넌스 체계 구축을 공공기록정책의 새로운 방향으로 제시하였다.

이상의 선행연구는 기록관리정책 연구가 개념 정립과 법제 분석에서 출발하여, 영역적으로 국가기록관리체계가 지방 공공기록관리로, 시기적으로 디지털 전환기 정보거버넌스 논의로 확장되어 왔음을 보여준다. 그러나 기존 연구를 살펴보면 일반 공공정책이나 프로그램 평가의 이론과 방법을 제시하는 데 중점을 두어 기록관리정책의 특성을 반영한 평가모형에 대한 논의는 상대적으로 충분하지 않았음을 확인할 수 있다. 따라서 이 연구에서는 기록관리정책 평가모형을 설계하기 위해 일반 정책평가 이론을 토대로 하되, 기록물의 진본성, 무결성, 신뢰성, 이용가능성 등 고유의 가치와 함께 공공성, 전문성, 책임성 등의 정책적 가치를 반영한 평가모형의 설계 방향을 도출하고자 한다.

2. 정책평가 관련 국제기준 분석

2.1 정책평가의 개념과 유형

정책평가는 공공정책이 의도한 목표를 달성하였는지, 정책수단이 적절하게 설계·집행되었는지와 그 결과가 사회적으로 어떠한 효과와 영향을 주었는지를 체계적으로 판단하는 분석 과정이다. 정책평가는 단순한 실적 점검이나 행정적 확인에 그치지 않고 정책의 필요성, 설계, 집행, 산출, 결과 및 영향을 경험적 자료와 명시적 기준에 따라 분석하는 응용사회과학적 활동으로 이해된다(Rossi et al., 2019). 우리나라 「정부업무평가 기본법」에서도 평가는 정책·사업·업무 등에 관하여 계획의 수립, 집행 과정 및 결과를 점검·분석·평정하는 것으로 규정된다. 이는 정책평가가 정책 시행 이후 성과만을 확인하는 절차가 아니라 정책의 계획 단계부터 집행과 결과에 이르는 전 과정을 대상으로 하는 종합적 판단 과정임을 의미한다.

정책평가는 정책의 효과성과 타당성을 검증하고 공공부문의 책임성과 투명성을 확보하며 정책학습과 환류를 가능하게 한다는 점에서 필요하다(Weiss, 1998). 공공정책은 국민의 세금과 행정권한을 바탕으로 사회문제에 개입하는 활동이므로 정책목표가 현실의 문제 구조에 부합하는지, 선택된 정책수단이 목표 달성에 적합한지와 실제 집행 결과가 의도한 성과로 이어졌는지를 검증할 필요성이 있다(Rossi et al., 2019). 또한 평가결과는 정책의 지속, 수정, 확대, 축소 또는 폐지 여부를 판단하는 근거로 활용되어 향후 정책 설계와 제도 개선에 반영된다. 따라서 정책평가는 정책과정의 합리성, 책임성, 투명성, 민주적인 통제 가능성을 확보하는 제도적 장치라고 할 수 있다(「정부업무평가 기본법」 제7조; Vedung, 1997).

정책평가는 평가가 이루어지는 시점에 따라 사전평가, 과정평가, 사후평가로 구분할 수 있다. 이 중 사전평가와 사후평가는 정책의 설계 단계와 결과 단계에서 각각 핵심적인 역할을 수행한다(박홍윤, 박규동, 2023). 사전평가는 정책이 본격적으로 시행되기 전에 수행되는 평가로 정책문제의 성격, 정책목표의 적절성, 정책수단의 타당성, 예상되는 효과와 부작용, 집행 가능성 등을 검토한다. 즉 사전평가는 정책 시행 이후의 결과를 기다려 판단하는 방식이 아니라 정책 설계 단계에서 정책의 필요성, 적합성, 실현 가능성을 미리 분석하는 예방적이자 진단적 평가이다. 이를 통해 정책결정자는 정책목표가 명확하게 설정되었는지, 목표가 사회적 필요와 우선순위에 부합하는지, 선택된 수단이 목표 달성에 적합한지 그리고 정책 집행에 필요한 자원·조직·법적 기반이 충분한지를 점검할 수 있다.

사전평가의 핵심 기능은 정책결정의 합리성을 높이는 데 있다. 정책은 한정된 공공자원을 배분하는 과정이므로 정책 시행 전 단계에서 정책대안의 타당성과 효과를 비교하고 검토하는 절차가 필요하다. 사전평가에서는 정책문제의 정확한 진단, 목표와 수단 간의 논리적 연계, 예상 효과와 위험요인의 분석, 이해관계자에 대한 영향, 다른 정책 및 제도와의 일관성 등이 중요한 평가 항목이 된다(OECD, 2010). 이러한 검토가 충분히 이루어질 때 정책은 시행 이후 발생할 수 있는 시행착오와 부작용을 줄일 수 있으며, 정책의 사회적 수용성과 집행 가능성을 높일 수 있다. 따라서 사전평가는 정책을 ‘할 것인가’와 ‘어떻게 설계할 것인가’를 판단하는 데 필요한 근거를 제공한다.

반면 사후평가는 정책이 시행된 이후 수행되는 평가로 정책이 애초에 설정한 목표를 달성하였는지, 정책의 산출과 결과가 무엇인지, 정책이 사회적·제도적으로 어떠한 영향을 미쳤는지를 분석한다. 사후평가는 정책의 효과성, 영향, 효율성, 지속가능성 등을 검토함으로써 해당 정책의 지속, 수정, 확대 또는 축소, 폐지 여부를 판단하는 근거를 제공한다. 이때 사후평가는 단순히 성과지표의 달성 여부를 확인하는 데 그치지 않는다. 정책 시행 이후 나타난 변화가 실제로 정책개입에 의해 발생한 것인지, 예상하지 못한 부작용이나 과급효과가 있었는지, 투입된 비용에 비해 성과가 적절했는지 그리고 정책효과가 장기적으로 유지될 수 있는지를 종합적으로 검토한다(OECD, 2010; Rossi et al., 2019).

사전평가와 사후평가는 평가 시점과 목적에서 차이가 있지만 상호 보완적인 관계에 있다. 사전평가는 정책 시행 이전에 정책 필요성과 가능성을 검토하여 정책 설계의 오류를 줄이는 역할을 하고, 사후평가는 정책 시행 이후 실제 성과와 영향을 분석하여 정책의 책임성과 개선 가능성을 확보한다. 사전평가가 정책의 ‘타당성’과 ‘예측 가능성’에 초점을 둔다면 사후평가는 정책의 ‘효과성’과 ‘영향’에 초점을 둔다. 따라서 동일한 정책이라도 사전평가에서는 문제 진단, 목표의 적절성, 대안 비교, 예상 효과 및 위험요인 분석이 중요하게 다루어지고, 사후평가에서는 목표 달성도, 실제 성과, 예상하지 못한 영향, 비용 대비 효과, 지속가능성이 주요 평가 항목이 된다(ERA-LEARN, n.d.; OECD, 2010).

종합하자면 정책평가는 정책의 전 과정을 대상으로 하되 평가가 이루어지는 시점에 따라 분석대상과 평가기준이 달라진다. 특히 사전평가와 사후평가는 정책평가의 핵심 축으로서 정책의 설계와 결과를 각각 점검한다. 사전평가는 정책 시행 이전에 정책문제의 적절성, 목표의 명확성, 수단의 타당성, 집행 가능성, 예상 효과와 위험요인을 검토함으로써 정책결정의 합리성을 높인다. 사후평가는 정책 시행 이후 목표 달성도, 실제 성과, 정책의 영향, 효율성, 지속가능성을 분석함으로써 정책의 책임성과 환류를 강화한다(OECD, 2010; Weiss, 1998). 따라서 정책평가모형을 구성할 때 사전평가와 사후평가의 구분은 단순한 형식적 분류가 아니라 평가기준, 평가문항, 자료수집 방법, 정책 개선 방향을 결정하는 핵심적인 이론적 기준이 된다(박홍윤, 박규동, 2023). 앞서 연구배경과 목적에서 언급한 바와 같이 이 연구에서는 과정평가는 제외하였다. 다만 사전평가의 실현가능성과 사후평가의 성과원인 및 환류 여부를 판단하기 위한 보조적인 요소로 집행역량, 개선계획 이행 등 집행과정과 관련한 사항을 활용하였다.

2.2 정책평가 관련 국제기준

2.2.1 OECD DAC 평가기준

OECD 개발원조위원회(DAC)의 평가기준은 국제개발협력 분야에서 출발하였으나 국제협력사업, 원조사업 뿐만 아니라 공공정책, 정부 프로그램 평가 등 다양한 영역에서 널리 활용되는 대표적 평가기준이다. 전통적으로 DAC 기준은 적절성(Relevance), 효과성(Effectiveness), 효율성(Efficiency), 영향(Impact), 지속가능성(Sustainability)의 다섯 가지 기준으로 구성되었으나, 2019년 개정 과정에서 일관성(Coherence)이 추가되어 총 여섯 개의 기준으로 정비되었다. 이러한 개정은 개별 정책이나 사업이 자체 목표를 달성했는지를 넘어서 다른 정책, 제도, 개입과 어떤 관계를 맺는지를 평가할 필요성이 커졌음을 반영한다(OECD DAC, 1991; OECD, 2019; OECD, 2021).

DAC 기준의 구성요소는 여섯 개의 평가기준과 각 기준에 대응하는 평가질문이다. 이 기준은 정책이나 사업이 해결하고자 한 사회문제와 수요에 부합했는지, 설정된 목표를 달성했는지, 투입 대비 성과가 적절했는지, 의도하거나 의도하지 않은 장기적 영향이 있었는지와 성과가 지속될 가능성이 있는지를 체계적으로 질문하도록 한다. 평가질문은 기준을 실제 평가에 적용하는 방법에 해당하며 평가자가 각 기준에 따라 무엇을 관찰하고 판단해야 하는지를 구체화한다. <표 1>은 여섯 개 기준의 정의와 기준별 핵심 평가질문을 정리한 것으로, 이는 특정 정책영역에서 평가지표를 개발할 때 출발점으로 활용될 수 있다. 이 기준의 가장 큰 특징은 평가자가 무엇을 관찰하고 판단해야 하는지를 명료하게 제시한다는 점에서 정책평가의 기본 틀로 활용하기 쉽다는 것이다.

〈표 1〉 OECD DAC 평가기준의 정의와 핵심 평가질문

평가기준	정의	핵심 평가질문
적절성(Relevance)	정책이나 사업이 수혜자와 기관의 수요, 우선순위, 정책·제도적 맥락에 부합하는 정도	정책이 해결하려는 문제와 현장의 수요에 부합하는가?
일관성(Coherence)	정책이나 사업이 다른 정책, 제도, 개입과 조화를 이루는 정도	다른 정책·제도와 상충하지 않고 정합적으로 작동하는가?
효과성(Effectiveness)	정책이 설정한 목표와 산출·결과를 달성한 정도	설정된 목표를 달성하였으며 의도한 결과가 나타났는가?
효율성(Efficiency)	투입된 자원 대비 성과가 경제적·적시적으로 산출된 정도	투입 대비 성과가 적절하며 자원이 경제적으로 사용되었는가?
영향(Impact)	정책이 가져온 의도적·비의도적, 직접적·간접적 장기 효과	정책이 가져온 장기적·구조적 변화는 무엇인가?
지속가능성(Sustainability)	정책의 순편익이 지속되거나 지속될 가능성이 있는 정도	정책 종료 이후에도 성과가 유지될 수 있는가?

평가방식 측면에서 DAC 기준은 기준별 평가질문을 설정한 뒤 정량·정성 자료를 결합하여 정책의 성과와 가치를 종합적으로 판단하도록 하여 성과와 가치 판단의 균형을 지향한다. 정책평가가 단순히 산출(Output)의 양적 달성 여부에 머문다면 정책이 실제 사회문제 해결에 기여했는지를 파악하기 어려울 것이다. 이에 비해 DAC 기준은 결과(Outcome), 영향, 지속가능성을 함께 고려함으로써 단기성과와 장기성과를 구분하고 정책개입의 가치와 사회적 효과를 보다 종합적으로 분석할 수 있도록 한다. 특히 효과성과 영향을 구분하는 방식은 단기 목표 달성과 장기 사회변화를 분리하여 분석할 수 있게 한다는 점에서 학술적·실무적 의의가 있다.

다만 활용 측면에서 DAC 기준은 일반성이 높은 만큼 구체적인 평가방법이나 지표체계를 직접 제공하지는 않는다. 따라서 특정 정책영역에 적용하기 위해서는 정책목표, 대상집단, 정책수단, 자료 접근성, 이해관계자 구조 등을 고려하여 기준별 평가질문과 지표를 재구성해야 한다. 또한 효과성과 영향, 적절성과 일관성 등 일부 기준은 실제 평가과정에서 경계가 모호해질 수 있으므로 평가설계 단계에서 각 기준의 조작적 정의를 명확히 할 필요가 있다.

2.2.2 UNEG 평가 규범 및 표준

UNEG의 「평가 규범 및 표준(Norms and Standards for Evaluation)」은 UN 산하 기구 및 전문기관 등 UN에 포함된 조직을 대상으로 제정된 평가의 원칙, 제도적 조건, 수행 기준을 제시하는 대표적 국제기준이다. OECD DAC 기준이 평가자가 정책이나 사업에 대해 어떤 질문을 제기해야 하는지에 초점을 둔다면, UNEG 기준은 평가가 어떤 제도적이고 윤리적 조건 아래에서 수행되어야 정당하고 신뢰성있는 평가로 인정받을 수 있는지를 강조한다. 이러한 점에서 UNEG 기준은 평가항목의 체계라기보다 평가 거버넌스의 규범적 기준으로 이해할 수 있다.

UNEG 기준의 핵심은 평가의 독립성(Independence), 공정성(Impartiality), 신뢰성(Credibility), 유용성(Utility), 투명성(Transparency)을 확보하는 데 있다. 평가가 정책 수행기관의 이해관계에 종속된다면 평가결과의 객관성과 신뢰성이 약화될 수 있으므로 평가기능의 독립성과 평가자의 공정성은 평가제도의 핵심 조건이 된다. 또한 평가가 정책결정과 조직학습에 활용되기 위해서는 평가질문, 방법, 자료, 판단근거, 한계가 투명하게 제시되어야 하며, 평가결과가 실제 이용자의 의사결정 수요와 연결되어야 한다(UNEG, 2016; UNEG, 2020).

UNEG 기준의 또 다른 특징은 인권 기반 접근(Human rights-based approach)과 성평등(Gender equality)을 평가의 핵심 원칙으로 포함한다는 점이다. 이는 정책평가가 단순한 효과 측정이나 비용 대비 성과 분석에 머물

수 없으며, 정책이 취약집단, 권리보유자, 성별·계층별 집단에 미치는 차별적 영향을 분석해야 한다는 점을 강조한다. 평가의 독립성, 공정성, 신뢰성, 투명성뿐 아니라 인권과 성평등 관점을 평가의 핵심 원칙으로 제시하고 있다는 사실은 최근 공공정책 평가에서 형평성, 포용성, 민주적 정당성의 중요성이 커지는 흐름과도 연결된다.

학술적으로 UNEG 기준은 평가를 기술적 측정행위가 아니라 공공책임성과 민주적 정당성을 확보하는 제도적 장치로 이해할 수 있다. 평가가 정책과정의 권력관계, 이해관계자 참여, 취약집단 보호, 윤리적 책임과 분리될 수 없다는 점을 제시한다는 점에서 의의가 있다. 그러나 UNEG 기준은 평가의 원칙과 제도적 조건을 강하게 제시하는 반면, 구체적인 성과지표 설계나 비용효과 분석들은 상대적으로 제한적이다. 따라서 실제 정책평가에 적용할 때에는 DAC 기준이나 논리모형 기반 평가방법과 결합하여 활용하는 것이 적절할 것이다.

2.2.3 세계은행 IEG 평가 원칙 및 품질기준

세계은행 IEG는 세계은행 사업, 국가협력전략, 개발정책 및 프로그램의 성과를 독립적으로 평가하는 기관이다. IEG의 평가체계는 개발정책 평가에서 결과사슬(Results chain), 성과관리(Results-based management), 독립평가(Independent evaluation), 증거기반 학습(Evidence-based learning)을 강조한다. 이는 정책개입이 어떠한 논리로 문제를 해결하고자 했으며 투입과 활동이 산출 및 성과로 어떻게 연결되었는지를 검증하는 데 초점을 둔다.

IEG 기준의 첫 번째 특징은 정책개입의 논리적 구조를 중시한다는 점이다. 정책이나 프로그램은 투입(Input), 활동(Activity), 산출(Output), 결과(Outcome), 영향(Impact)의 단계로 구성되어, 평가자는 각 단계가 논리적으로 연결되어 있는지를 검토해야 한다. 이러한 접근은 논리모형(Logic model)과 변화이론(Theory of change)에 기반한 평가실무와 밀접하게 관련된다. 정책 실패의 원인을 명확하게 분석하기 위해서는 목표의 부적절성, 집행과정의 문제 혹은 외부요인에 의한 것인지 개입논리를 명료하게 하는 것이 필수적이다.

두 번째 특징은 독립평가를 통해 자기평가의 한계를 보완한다는 점이다. 정책 수행기관의 자기평가는 내부 정보를 풍부하게 활용할 수 있다는 장점이 있으나, 성과를 과대평가하거나 실패요인을 축소할 가능성이 있다. IEG의 독립평가는 이러한 편향을 줄이고 평가결과의 신뢰성을 높여 정책결정자와 이해관계자에게 보다 객관적인 판단근거를 제공하는 역할을 한다.

세 번째 특징은 평가결과의 환류와 학습을 강조한다는 점이다. IEG 기준에서 평가는 사업의 사후관정에 그치지 않고 향후 사업설계, 자원배분, 제도개선, 조직학습으로 연결되어야 함을 밝히고 있다. 이러한 접근은 정책평가를 일회적 감사나 점검이 아니라 성과관리와 정책학습의 순환 과정으로 이해하게 한다. 특히 대규모 개발사업이나 국가전략과 같이 장기성과가 중요한 정책영역에서는 산출지표보다 결과와 영향, 지속가능성을 추적하는 평가체계가 요구될 때 유용하게 활용할 수 있다(WBG, n.d; WBG, 2015; WBG, 2019).

다만 IEG 기준은 개발금융과 국제개발 프로그램의 맥락에서 발전하였기 때문에 일반 행정정책에 직접 적용할 경우 정책유형별 조정이 필요하다. 또한 결과지표와 정량성결과를 강조할 경우 제도변화, 역량강화, 거버넌스 개선과 같은 장기적·질적 성과가 충분히 포착되지 못할 수 있다. 따라서 IEG 기준은 논리모형과 성과관리 체계를 설계하는 데 강점을 가지되, 질적 평가와 제도적 맥락 분석을 병행할 때 보다 효과적으로 활용될 수 있다.

세 기준은 모두 정책평가의 국제적 준거로 활용될 수 있으나, 기준의 성격과 활용 목적은 상이하다. OECD DAC 기준은 평가항목과 평가질문을 구성하는 데 강점을 지니는 평가기준이며, UNEG 기준은 평가의 독립성, 윤리성, 신뢰성을 보장하는 평가운영 원칙을 제공하는 평가규범이다. 세계은행 IEG 기준은 정책개입의 논리와 성과, 영향, 환류를 연결하는 성과관리형 평가체계를 제시한다. 따라서 세 기준은 상호 대체적인 관계라기보다 상호보완적 관계로 이해하는 것이 타당하다. <표 2>는 분석기준에 따라 세 국제기준을 비교한 것이다.

〈표 2〉 정책평가를 위한 국제기준 비교

구분	OECD DAC 평가기준	UNEG 평가 규범과 기준	세계은행 IEG 평가 원칙 및 품질기준
주요주체	OECD 개발원조위원회(DAC) DAC Network on Development Evaluation	유엔평가그룹(UNEG), 유엔에 소속된 평가기능 및 평가부서	세계은행 독립평가그룹(IEG), 세계은행 그룹 및 개발정책 평가기관
특징	정책·사업의 성과와 가치를 판단하기 위한 평가질문 및 판단기준 중심	평가의 독립성, 윤리성, 신뢰성, 인권·성평등을 강조하는 평가 거버넌스 기준	결과사슬, 성과관리, 독립평가, 증거기반 학습을 강조하는 성과관리형 평가체계
정책초점	정책이나 사업이 사회문제와 수요에 부합하고 목표를 달성했는지, 장기적 영향과 지속가능성이 있는지에 초점	평가가 정당하고 신뢰롭게 수행되는 제도적 조건, 이해관계자 참여, 윤리성, 형평성에 초점	정책개입의 투입-활동-산출-결과-영향 간 논리적 연결성과 성과의 입증 및 정책학습에 초점
평가기준	6개 평가기준(적절성, 일관성, 효과성, 효율성, 영향, 지속가능성)	8개 평가규범(독립성, 공정성, 신뢰성, 유용성, 투명성, 윤리, 인권 기반 접근, 성평등 기준)	논리모형, 변화이론, 결과사슬, 성과지표, 독립평가, 증거기반 판단, 환류체계
평가 방식	기준별 평가질문을 설정하고 정책성과를 정량·정성 자료로 종합 판단	평가절차, 평가자의 독립성, 자료와 판단근거의 투명성, 윤리적 수행 여부 점검	정책개입 논리를 설정한 뒤 성과지표와 증거를 통해 결과사슬의 타당성과 성과 검증
한계	일반성이 높아 구체적 평가방법·지표체계를 직접 제공하지 않으며 일부 기준 간 경계가 모호	평가 원칙과 제도적 조건에 비해 성과지표 설계나 비용효과 분석들은 제한적으로 제공	개발금융 맥락에서 발전하여 정책유형별 조정이 필요하고 정량성과 강조시 질적·제도적 성과 포착에 한계

3. 기록관리정책과 주요국의 기록관리 평가제도

3.1 기록관리정책의 정의와 특징

곽동철(2001)은 기록관리정책을 국가적 차원의 기록문화 정착이라는 목표를 달성하기 위한 정부활동이자 기록물의 수집·정리·보존·유통 등을 촉진하거나 규제하는 계획과 활동으로 정의하였다. 이를 미루어보면 기록관리 정책은 공공 및 사회 영역에서 생산·접수·관리·보존·평가·활용되는 기록의 신뢰성과 지속가능성을 확보하고 행정책임성, 공공투명성, 시민권리, 사회적 기억의 보존을 실현하기 위해 국가, 공공기관, 지방자치단체, 기록관리기관, 전문직, 시민사회 등이 수립·집행·평가하는 법제도적·조직적·기술적·거버넌스적 방침과 수단의 총체라고 정의할 수 있다.

기록관리 정책은 기록물을 관리하기 위한 단순한 행정절차나 기술적 방침에 한정되지 않으며, 기록의 생산·관리·보존·활용을 둘러싼 공공정책의 한 영역으로 이해되어야 한다. 정책평가 이론에서 정책은 목표, 수단, 참여자, 집행환경, 산출과 결과가 결합된 공공적 개입으로 파악되므로 기록관리 정책 역시 기록관리 실무의 운영지침을 넘어 공공기록의 신뢰성과 지속가능성을 확보하기 위한 제도적·조직적·기술적 정책활동으로 볼 수 있다.

선행연구와 관련 법령을 종합하면 기록관리정책은 다음의 특징을 가지고 있다. 첫째로 기록관리 정책은 기록 생애주기 전반을 대상으로 한다(곽동철, 2001; Shepherd & Yeo, 2003). 기록관리 정책은 기록의 생산, 등록, 분류, 이관, 평가, 보존, 공개, 활용 등 기록이 생성되어 관리·활용되는 전 과정을 포괄한다. 따라서 기록관리 정책은 단일 사업이나 특정 시스템 구축에 머무르지 않고 기록의 생성 단계부터 보존·활용 단계까지 연속적으로 작동해야 한다.

둘째로 기록관리 정책은 행정의 책임성, 투명성 및 시민권리 보장과 밀접하게 연결된다(이영학, 2009). 공공부문 기록은 정부가 수행한 행위와 의사결정의 근거이므로 기록관리 수준은 행정책임성, 국민의 정보 접근권, 행정서비스의 지속성과 직접적으로 관련된다. 기록관리의 비효율성이 해소되지 않으면 행정서비스 제공이 지연되고 시민 권리 보장에도 한계가 발생할 수 있다.

셋째로 기록관리 정책은 중앙기록물관리기관의 조정 기능과 연결된다(「공공기록물법」 제9조; 이승일, 2014). 기록관리 실무는 개별 행정기관에서 이루어지지만, 공공부문 전체의 기록관리 기준과 방향은 중앙기록물관리기

관 또는 국가단위의 기록물관리기관을 통해 제시되는 경우가 많다. 이는 국가단위 기록물관리기관이 단순히 중요 기록물을 보존하는 기관이 아니라, 표준, 지침, 정책, 매뉴얼, 시스템, 교육, 평가를 통해 공공부문 기록관리 수준을 조정하는 정책 주체임을 의미한다.

넷째로 기록관리 정책은 법제도, 조직, 기술, 전문인력, 거버넌스가 결합된 복합 정책이다(설문원, 2020; 이영학, 2009). 선행연구에서는 중앙기록물관리기관의 독립성, 기록관리 거버넌스, 기록관리시스템의 확산과 고도화, 기록물관리 전문요원 배치, 기록물관리기관의 활성화 등을 기록관리 정책의 핵심 조건으로 제시하였다. 디지털 전환 이후에는 정보관리, 디지털 연속성, 기록자산 관리, 정보거버넌스의 차원까지 정책 범위가 확장될 필요가 있다.

마지막으로 기록관리 정책은 정책평가의 관점에서 투입, 과정, 산출, 결과의 논리적 연계를 통해 분석되어야 한다. 정책평가는 정책의 필요성, 설계, 집행, 효과성, 효율성을 경험적 자료와 명시적 기준에 따라 판단하는 과정이므로 기록관리 정책도 제도 도입 여부나 사업 실적만으로 평가해서는 충분하지 않다. 기록관리 정책의 효과를 판단하기 위해서는 투입된 자원, 집행 과정, 산출물, 실제 결과와 영향이 어떻게 연결되는지 종합적으로 검토해야 한다. 이러한 특징은 기록관리정책에 대한 평가가 일반 정책평가의 틀을 토대로 하되 기록관리 분야의 고유한 가치와 환경을 반영해야 함을 시사한다.

이러한 관점에서 주요국의 기록관리 평가 사례를 분석기준에 따라 분석하였다. 정책평가는 국가 또는 부문 수준에서 정책목표, 정책수단, 자원배분, 집행, 산출, 결과와 영향을 판단하는 반면, 기록관리 평가는 개별 조직의 준수 여부, 역량, 절차, 위험 및 성숙도를 진단한다는 점에서 차이가 있다. 이러한 점에서 각국의 기록관리평가 사례를 분석하여 다른 정책평가에는 없으나 기록관리 정책에 존재하는 특징적인 요소들을 도출할 필요가 있다.

3.2 국가별 기록관리 평가제도 분석

3.2.1 미국 NARA 기록관리 자체평가(RMSA) 및 고위 기록관리 책임자(SAORM) 보고 체계

미국 NARA는 연방기관의 기록관리 현황을 평가하기 위해 체계적인 기록관리 보고 체계를 운영한다. 이 체계는 기록관리 자체평가(RMSA), 고위 기록관리 책임자(SAORM) 연간 보고서, 연방정부 전자기록 및 이메일 관리(Federal Electronic Records and Email Management, FEREM) 보고서 등 서로 다르지만 상호 연계된 자료를 활용한다. NARA는 이러한 자료를 통해 기관별 기록관리 수준과 전자기록관리 현황을 수집하고, 이를 연례 연방기관 기록관리 보고서(Federal Agency Records Management Report)로 종합하는 중앙집중형 보고·평가 체계를 구축하고 있다(NARA n.d.a; NARA n.d.b).

NARA의 평가체계는 법령·규정 준수, 연방기관의 기록관리 수준, 전자기록 및 이메일 관리, 고위책임자의 책무성에 초점을 맞추고 있다. 먼저, 구성요소 측면에서 RMSA는 연방기관이 매년 자체평가를 실시하여 법령 및 규정상 기록관리 요구사항을 준수하고 있는지를 확인하도록 하는 제도로, 기관이 기록관리 요구사항을 어느 정도 충족하고 있는지를 스스로 점검하게 하는 장치이다. SAORM는 기관의 고위책임자가 전자기록 전환과 전략적 기록관리 목표 달성 상황을 설명하도록 하는 보고 체계로, 기록관리 책임을 실무부서 차원에 한정하지 않고 기관 고위관리자의 책임으로 제도화한다는 점에서 의미가 있다.

평가방식 측면에서 NARA의 평가체계는 각 기관에서 매년 제출한 자기평가와 보고자료를 취합, 분석하여 연방정부 전체의 기록관리 상태를 종합하여 판단한다. 따라서 NARA 사례는 기록관리정책 평가가 단순한 업무점검이 아니라 기관 차원의 책무성과 중앙기록물관리기관의 감독 기능을 결합하는 방식으로 운영될 수 있음을 보여준다.

활용성 측면에서 NARA의 평가체계는 연례 보고와 자기평가 자료가 축적되면 기관별 기록관리 수준, 전자기록 관리 전환 정도, 공통 취약영역, 개선 필요사항을 지속적으로 추적하며, 정기적이고 비교 가능한 자료를 통해 중앙기관이 연방정부 전체의 기록관리 상태를 파악할 수 있다는 장점이 있다. 다만 자체평가에 기반하기 때문에 실제 업무관행과 보고내용 간 차이를 교차 검증할 보완장치가 필요하다는 한계점도 있다. 따라서 표본기록 점검,

시스템 로그, 이관·폐기 통계, 감사자료와 같은 객관적 증거를 함께 활용해야 한다.

3.2.2 영국 TNA 정보관리평가(Information Management Assessment)

영국 TNA의 정보관리평가(IMA) 프로그램은 공공부문 기관이 정보를 얼마나 잘 관리하는지에 대한 정보관리 역량과 함께 관련 위험을 완화하고 있는지에 대한 위험 관리 수준을 독립적으로 평가하기 위한 제도이다. IMA의 목적은 기관의 정보·기록관리 현황을 진단하여 잠재적 위험을 식별하고, 기관 스스로 개선 목표를 설정하여 그 이행 정도를 추적하도록 지원하는 데 있다. IMA는 기관 주도의 자기평가를 우선하는 방식을 강조하면서도 필요시 독립적 평가를 결합함으로써 자기평가와 외부 검증을 결합한 평가 체계이다.

TNA 평가체계의 구성요소를 이해하기 위해서는 먼저 성숙도 모델의 개념을 살펴볼 필요가 있다. 성숙도 모델(Maturity model)은 조직의 관리 역량을 초기 단계에서 최적화 단계에 이르는 여러 수준으로 구분하고, 조직이 현재 어느 수준에 있는지를 진단하여 다음 단계로 나아가기 위한 개선 과제를 도출하도록 하는 진단 틀이다. TNA는 이러한 성숙도 모델을 기록·정보관리 영역에 적용한 성숙도 평가도구를 제공하며, 기관은 이 도구를 사용하여 성숙도 진단을 수행한다. 즉 진단의 틀인 성숙도 모델, 진단의 수단인 평가도구, 진단의 결과물인 성숙도 진단이 하나의 체계로 연결되어 있다.

핵심 평가도구로 제46조 실무규약 성숙도 평가도구(Section 46 Code Maturity Assessment Tool)와 전자기록물 평가·선별 및 민감도 검토 성숙도 평가도구(Digital Appraisal, Selection and Sensitivity Review Maturity Assessment Tool, Digital ASSR Maturity Assessment Tool), M365 지식정보관리 성숙도 자체평가도구(M365 KIM Maturity Self-Assessment Tool) 등 여러 성숙도 평가도구를 제공한다. 제46조 실무규약 성숙도 평가도구는 「영국 2000년 정보자유법(Freedom of Information Act 2000)」 제46조에 따른 기록관리 실무규범을 기준으로 기관의 정보·기록관리 정책과 절차를 점검하고 개선목표를 설정하며 진전 상황을 추적하도록 지원한다. 전자기록물 평가·선별 및 민감도 검토 성숙도 평가도구는 디지털 기록의 평가·선별·민감정보 검토 역량을 5단계 성숙도와 거버넌스, 기술, 방법론, 기술역량, 처리능력, 참여, 투명성·책임성 등 핵심 영역을 기준으로 진단한다. 이러한 도구들은 기록관리 실무규범, 디지털 평가·선별·민감정보 검토 역량, M365 환경의 정보거버넌스 평가에 활용된다(TNA, n.d.).

TNA의 평가방식은 독립적 평가와 자기평가 우선 방식을 결합한 성숙도·위험관리 평가 체계로 정리할 수 있다. IMA는 기관 방문, 문서 검토, 관계자 면담, 개선계획 확인 등을 통해 기관의 정보관리 성숙도를 평가한다. NARA가 연방기관 전체의 표준화된 보고와 중앙집중형 모니터링에 강점을 갖는다면, TNA 평가방식은 기관의 현재 역량 수준, 정보관리 위험수준, 개선 목표를 질적으로 진단하는 데 초점을 둔다. 특히 성숙도 모델을 활용함으로써 기관이 단순히 기준을 충족했는지 여부를 넘어 현재 어느 단계에 있으며, 어떤 개선과제를 우선 추진해야 하는지를 파악하는데 유용한 평가체계이다.

활용성과 한계 측면에서 TNA의 평가체계는 평가결과가 기관별 개선목표 설정과 역량 강화로 직접 연결된다는 점에서 장점을 가진다. 또한 정보관리 성숙도, 기록관리 실무규범 준수, 디지털 평가·선별 역량, M365 정보 거버넌스 등의 평가영역은 현대 공공기관의 기록관리 환경을 구체적으로 반영하고 있다. 다만 성숙도 진단과 개선 권고가 실질적 변화로 이어지기 위해서는 기관의 실행역량과 후속관리 체계가 뒷받침되어야 한다. 따라서 성숙도 기반 평가를 도입할 경우에는 평가 자체뿐만 아니라 평가 이후의 실행계획, 진전 추적, 컨설팅, 재평가 절차를 함께 설계할 필요가 있다.

3.2.3 호주 NAA 디지털 역량진단(Check-up Digital) 및 디지털 연속성(Digital Continuity)

호주 NAA는 공공기관의 디지털 정보관리 역량을 점검하고 디지털 연속성을 확보하기 위한 정책적 도구와 지침을 제공해 왔다. NAA는 호주 정부기관을 대상으로 디지털 정보관리 수준을 점검하고, 디지털 정보가 필요한 기간 동안 신뢰할 수 있고 접근 가능하며 이용 가능한 상태로 유지될 수 있도록 관리하는 데 초점을 둔다.

NAA의 핵심 도구와 기준은 디지털 역량진단(Check-up Digital), 2020년 디지털 연속성(Digital Continuity 2020) 정책 및 관련 디지털 정보관리 지침이다. 이를 통해 기관은 자신의 디지털 기록관리 상태를 진단하고 개선 과제를 도출할 수 있다. 디지털 역량진단은 2014년부터 2018년까지 진행되었던 운영 프로그램으로, 기관이 자체적으로 디지털 정보관리 상태를 점검하고 취약한 영역을 식별하도록 지원하는 자기진단형 평가도구이다. 이 도구는 현재 플러스 역량진단(Check-up PLUS)로 개편되었다. 디지털 연속성 정책은 업무수행에 필요한 디지털 정보가 필요한 기간 동안 신뢰할 수 있고 이용 가능하며 접근 가능한 상태로 유지되도록 하는 데 초점을 둔다. 이는 기록관리정책이 종이기록의 보존 중심에서 벗어나 디지털 기록의 생성, 관리, 접근, 재사용, 장기보존 가능성까지 포괄해야 한다는 점을 보여준다(NAA, 2016; NAA, 2021).

평가방식 측면에서 NAA의 평가체계는 기관의 자기진단과 디지털 연속성 기반 평가를 결합한 방식으로 요약할 수 있다. 여기서 중요한 점은 기록관리정책 평가의 초점이 종이기록의 보존이나 규정 준수에 한정되지 않고 디지털 정보의 신뢰성, 접근성, 이용가능성, 업무연속성, 디지털 기록관리 역량으로 확장된다는 점이다. 즉 디지털 기록이 업무시스템, 협업도구, 데이터세트, 전자문서관리 환경에서 생산·관리되는 현실을 반영하여, 평가도 디지털 환경의 구조적 조건을 함께 고려해야 함을 시사한다.

활용성과 한계 측면에서 NAA의 장점은 디지털 전환 환경에서 기록의 지속적 이용가능성과 업무연속성을 평가할 수 있다는 데 있다. 메타데이터 품질, 시스템 간 연계, 접근권한, 로그관리, 장기보존 전략, 데이터세트 관리, 전자적 이관 가능성 등을 함께 점검해야 기록의 신뢰성과 지속가능성을 판단할 수 있다. 다만 디지털 연속성 평가는 기술, 조직, 예산, 책임체계가 함께 확보되지 않으면 지속되기 어렵기 때문에 자기진단 결과를 객관적 운영자료와 후속 개선계획으로 연결하는 평가체계가 필요하다.

이상에서 살펴본 세 가지 사례는 기록관리정책평가의 구조적·내용적 측면에서 반영할 수 있는 특징을 보여준다. 구조적으로 NARA의 연례 보고 표준화와 고위책임자의 책무성을, 내용적으로 TNA의 성숙도 기반 심층진단이나 NAA의 디지털 기록관리 역량과 디지털 연속성 평가를 결합하는 방식을 고려할 수 있다. 이를 통해 기록관리 정책 평가는 법령 준수 중심의 행정점검을 넘어 기관의 책임성, 정보관리 역량, 디지털 전환대응, 정책 환류를 종합적으로 판단하는 체계로 발전할 수 있다. <표 3>은 분석기준에 따라 세 사례를 비교한 것이다.

<표 3> 국가별 기록관리 기관평가 사례 비교

구분	미국 NARA 기록관리 자체평가 및 고위 기록관리 책임자보고 체계	영국 TNA 정보관리평가	호주 NAA 디지털 역량진단 및 디지털 연속성
주요 주체	NARA, 연방기관, 기관 고위기록관리책임자(SAORM)	TNA, 영국 공공부문 기관	NAA, 호주 정부기관
특징	연례 자기평가와 고위책임자 보고를 결합한 중앙집중형 보고·평가 체계	독립적 평가와 자기평가 우선 접근을 결합한 성숙도·위험관리 평가 체계	디지털 정보관리 수준을 점검하는 자기진단 및 디지털 연속성 기반 평가 체계
정책 초점	법령·규정 준수, 연방기관 기록관리 수준, 전자기록·이메일 관리, 고위책임자의 책무성	정보관리 성숙도, 기록관리 실무규범 준수, 디지털 평가·선별·민감정보 검토 역량, M365 정보거버넌스	디지털 정보의 신뢰성, 접근성, 이용가능성, 업무연속성, 디지털 기록관리 역량
핵심 도구	기록관리 자체평가(RMSA), 고위 기록관리 책임자(SAORM) 연간 보고서, 연방정부 전자기록 및 이메일 관리 보고서, 연방기관 기록관리 보고서	제46조 실무규약 성숙도 평가도구, 전자 기록물 평가·선별 및 민감도 검토 성숙도 평가도구, M365 지식정보관리 성숙도 자체평가도구	디지털 역량진단, 2020년 디지털 연속성 정책, 관련 디지털 정보관리 지침
평가 방식	기관별 연례 제출자료를 NARA가 취합·분석하여 종합 보고	공공기관의 자기평가를 기반으로 독립적 평가와 개선목표 설정을 지원	기관이 디지털 정보관리 상태를 자체 점검하고 디지털 연속성 기준에 따라 개선과제 도출
한계	자체평가에 기반하므로 보고내용과 실제 업무관행 간 차이를 검증할 보완장치 필요	성숙도 진단과 개선권고가 실질적 변화로 이어지기 위한 실행역량과 후속관리 체계 필요	기술·조직·예산·책임체계가 함께 확보되지 않으면 지속되기 어려워 후속 개선 계획과의 연결이 필요

4. 공공기록관리정책 평가모형 개발을 위한 설계원칙

앞서 살펴본 바와 같이 공공기록관리정책은 정책의 형성 단계와 결과 단계에 대응하는 사전평가와 사후평가를 필요로 한다. 사전평가는 정책 설계의 타당성과 실행가능성을 검토하여 정책 시행 과정에서의 시행착오를 줄이는 데, 사후평가는 정책의 실제 성과와 영향을 분석하여 책임성과 개선 가능성을 확보하는 데 목적이 있다. 이 장에서는 2장의 국제기준과 3장의 해외 사례 분석 결과를 토대로 공공기록관리정책의 사전평가와 사후평가 모형을 설계하기 위한 원칙을 도출하고, 각 설계원칙별로 참조할 국제기준과 해외 사례를 구체적으로 제시하였다.

4.1 평가모형 개발을 위한 주요 원칙 도출

공공기록관리정책에 대한 평가에서 첫 번째로 고려해야 할 사항은 정책목표와 평가기준 간의 정합성이다. 기록관리정책은 단순히 기록물의 보존이나 관리절차의 준수만을 목적으로 하지 않는다. 행정의 책임성 확보, 국민의 알 권리 보장, 공공정보의 신뢰성 유지, 디지털 기록의 장기적 이용가능성 확보 등 복수의 공공적 가치를 포함한다. 따라서 평가기준은 정책이 해결하려는 문제가 무엇인지 그리고 해당 문제가 기록관리의 공공적 가치와 어떻게 연결되는지를 확인할 수 있도록 설계되어야 한다.

OECD DAC 평가기준 중 적절성은 정책이나 사업이 수요, 우선순위, 제도적 맥락에 부합하는지를 평가하는 기준이다. 공공기록관리정책이 평가 대상이라면 특정 정책이 실제 기록관리 현장의 위협과 기관의 수요를 반영하고 있는지, 국가기록관리체계의 장기적 목표와 조화를 이루는지를 검토할 필요가 있다. 예를 들어 전자기록 장기 보존 정책을 평가한다면 단순히 보존시스템을 구축했는지가 아니라 디지털 기록의 훼손 위험, 포맷 노후화, 메타데이터 결손, 접근권한 관리 문제 등 실제 정책문제가 적절히 반영되었는지를 확인해야 한다.

NARA의 RMSA나 SAORM 체계는 연방기관이 기록관리 책임을 어떻게 이행하고 있는지를 점검하도록 설계되어 있다. 이는 기록관리정책의 목표가 개별 기관의 행정적 편의가 아니라 연방정부 차원의 책임성, 투명성, 전자기록관리 역량 확보와 연결되어 있음을 보여준다. TNA의 IMA 역시 기관의 정보관리 성숙도와 위험관리 수준을 점검함으로써 기록관리 평가가 정책목표와 기관의 실제 관리역량 사이의 부합성을 확인하는 절차임을 보여준다. NAA의 디지털 역량진단은 디지털 업무환경에서 기록이 적절히 포착되고 지속적으로 이용될 수 있는지를 점검한다는 점에서 정책목표와 평가기준을 디지털 연속성이라는 핵심 가치에 맞추고 있다.

따라서 공공기록관리정책 평가모형에서는 정책목표가 기록관리의 공공적 가치와 연결되어 있는지와 평가기준이 그 목표를 실제로 판단할 수 있는 방식으로 구성되어 있는지를 사전에 검토해야 한다. 이는 평가가 단순한 집행점검이나 성과보고에 머무르지 않고 정책의 필요성과 방향성을 판단하는 분석도구가 되기 위한 전제이다.

둘째로 정책수단과 정책성과 사이의 논리적 연결관계를 명확히 제시해야 한다. 공공기록관리정책은 일반적으로 법·제도 개선, 표준 수립, 시스템 구축, 기관 점검, 교육·컨설팅, 보존·복원, 기록정보서비스 확대 등 다양한 수단을 결합하여 추진된다. 이러한 정책수단이 어떤 과정을 거쳐 기록관리 역량 향상, 기록의 신뢰성 확보, 국민의 기록정보 접근성 확대와 같은 성과로 이어지는지 명확히 제시되지 않으면 사후평가 단계에서 정책의 성과를 해석하기 어렵다.

성과경로와 관련하여 세계은행 IEG의 평가원칙은 중요한 시사점을 제공한다. IEG는 정책이나 사업을 평가할 때 투입, 활동, 산출, 결과, 영향으로 이어지는 결과 사슬을 명확히 설정하도록 하며, 실제 성과가 이러한 경로를 통해 발생했는지를 검토한다. 예를 들어 전자기록관리시스템 고도화 정책에 적용해보면 시스템 구축이 아닌 시스템 구축 이후 메타데이터 품질의 개선 정도, 장기보존 가능성의 증대나 이용자 접근성의 확대 여부 등을 성과로 연결하여 평가한다.

해외 사례에서 이러한 성과경로의 논리를 확인할 수 있다. NARA의 RMSA는 기관이 기록관리 요건을 충족하고 있는지 자체적으로 평가하도록 하며, SAORM은 고위책임자가 기관의 기록관리 책임 이행 상황을 보고하도록 한다. 성과경로 측면에서 이러한 구조는 기록관리정책이 단순히 실무부서의 절차 준수에 그치지 않고 기관 전체의 책임체계와 전자기록관리 역량 강화로 이어져야 한다는 논리를 포함한다. TNA의 IMA는 성숙도 평가를 통해 기관의 현재 수준을 진단하고 개선계획을 도출하도록 하는데 평가가 현재 상태의 확인에서 끝나는 것이 아니라 위험 진단과 개선목표 설정을 거쳐 향후 성과경로를 구성하는 과정임을 보여준다. NAA의 디지털 연속성 정책 역시 디지털 기록의 장기적 이용가능성을 확보하기 위해서는 기록이 생산되는 업무시스템, 메타데이터, 접근권한, 보존포맷, 시스템 간 연계가 함께 고려되어야 한다고 밝히고 있다. 따라서 전자문서의 보존 여부가 아니라 업무과정에서 기록이 적절히 획득되고, 필요한 맥락정보가 유지되며 시간이 지나도 이용가능한 상태로 관리되는지를 디지털 기록관리정책의 성과경로로 인식하고 있음을 알 수 있다.

따라서 공공기록관리정책 평가모형은 정책수단과 성과 사이의 논리적 연결을 명확히 제시해야 한다. 사전평가에서는 정책개입이 어떤 변화를 의도하는지를 확인하고, 사후평가에서는 실제 변화가 그 경로를 통해 발생했는지를 검토해야 한다. 이는 평가결과를 단순한 실적 확인이 아니라 정책효과의 해석으로 연결하기 위한 핵심 조건이라 할 수 있다.

셋째로 평가자료 및 증거기반 평가방식을 고려해야 한다. 기록관리정책은 장기적이고 제도적인 성격을 가지므로 정책성과를 단일 지표만으로 판단하기 어렵다. 평가를 위해 기관의 자체 보고서 뿐만 아니라 점검표, 시스템 로그, 메타데이터, 이관·폐기·보존 기록, 접근권한 관리자료, 교육 이수자료, 이용통계, 감사 및 점검 결과 등 다양한 자료를 결합해야 한다. 따라서 사전평가 단계에서부터 평가항목별로 수집할 자료의 유형, 자료의 신뢰성, 정량자료와 정성자료의 결합 가능성 등을 검토할 필요가 있다.

NARA의 RMSA는 평가자료 확보의 측면에서 중요한 예시라고 할 수 있다. RMSA는 연방기관이 기록관리 요건을 어느 정도 충족하고 있는지를 자체적으로 응답하도록 구성되어 있고, 이를 통해 기관 간 기록관리 수준을 비교하여 취약한 영역을 파악할 수 있다. 다만 자체보고 방식은 응답의 정확성과 객관성에 한계가 있을 수 있어 사후평가에서는 보고자료를 그대로 수용하기보다 이관실적 등 실제 기록관리 산출물들과 대조할 필요가 있다는 점을 고려해야 한다. TNA의 IMA는 기관 방문, 문서 검토, 관계자 면담, 개선계획 확인 등을 통해 기록관리 성숙도를 평가한다는 점에서 증거기반 평가의 사례로 볼 수 있다. 이 방식은 단순한 수치 보고만으로는 파악하기 어려운 조직문화, 책임체계, 위험관리 수준, 실무역량을 평가할 수 있게 한다. 공공기록관리정책 평가에서도 정량지표뿐 아니라 면담, 사례분석, 전문가 검토, 기관별 개선계획 분석 등을 결합할 필요가 있다. NAA의 디지털 역량진단은 디지털 기록관리 수준을 진단하기 위해 기관의 디지털 기록 포착, 메타데이터 관리, 시스템 연계, 접근권한, 보존가능성 등을 점검한다. 이는 디지털 기록환경에서 평가자료가 단순한 문서목록에 그치지 않고, 시스템 내부의 로그, 메타데이터 구조, 업무시스템과 기록관리시스템 간 연계상태, 장기접근 가능성 등 기술적 증거를 포함해야 함을 보여준다.

따라서 공공기록관리정책 평가모형은 평가항목을 설정하는 데 그치지 않고 각 항목을 판단할 수 있는 증거자료를 함께 설계해야 한다. 특히 사후평가에서는 정책 담당기관의 설명이나 자체보고만으로 성과를 판단하기보다는 객관적 자료와 현장 증거를 통해 성과를 확인하도록 하여 평가의 신뢰성을 높일 수 있다.

넷째로 평가결과의 타당성뿐만 아니라 평가과정의 정당성과 신뢰성을 확보할 수 있도록 설계되어야 한다. 기록관리정책 평가는 행정기관 내부의 단순한 실적점검이 아니라 공공기록관리정책을 통해 행정의 설명책임과 국민의 알 권리가 보장되는지를 확인하는 과정이다. 따라서 평가가 신뢰성을 갖기 위해서는 평가주체의 독립성, 평가기준의 공개성, 평가방법의 명확성, 이해관계자의 참여, 평가결과의 설명 가능성이 확보되어야 한다.

이와 관련하여 UNEG의 평가 규범과 기준은 평가의 독립성, 공정성, 투명성, 윤리성, 이해관계자 참여를 강조한다는 점에서 중요한 시사점을 제공한다. 평가절차의 객관성을 확보하기 위해 평가기준과 방법을 사전에 공개하고,

평가결과뿐 아니라 자료의 범위, 분석방법, 한계를 함께 제시해야 한다. 특히 공공기록관리정책은 정보의 접근권, 개인정보 보호, 비공개기록 관리, 민감정보 처리와도 관련될 수 있으므로 평가과정에서 윤리성과 권리보호의 관점도 함께 고려되어야 한다.

NARA의 SAORM 보고체계는 고위책임자가 기관의 기록관리 현황을 보고하도록 함으로써 기록관리 책임을 실무부서 차원에만 두지 않고 기관의 책임으로 확장하여 책임성을 확보한 사례로 볼 수 있다. 이는 공공기록관리 정책 평가에서도 기관장 또는 고위관리자의 책임, 조직 차원의 지원, 예산과 인력 확보 여부를 함께 평가해야 함을 시사한다. TNA의 IMA는 평가결과를 통해 기관의 정보관리 위협과 개선 필요성을 제시하고 후속 개선계획의 이행을 확인도록 한다. 이는 평가절차가 단순히 문제를 지적하는 데서 끝나는 것이 아니라 기관이 그 결과에 대해 설명하고 개선하도록 요구하는 책임성 구조를 내포한다는 점에서 의미가 있다. NAA의 디지털 기록관리 점검 역시 기관이 디지털 기록을 적절히 관리하고 있는지를 확인함으로써 디지털 환경에서의 기록관리 책임을 구체화한다.

따라서 평가모형은 평가결과의 산출뿐 아니라 평가과정의 정당성을 함께 고려해야 한다. 평가가 ‘어떤 기준에 따라 어떤 자료를 사용하여 누구의 책임 아래 수행되었는지’를 명확히 제시함으로써 평가결과는 정책개선과 공공 책임성 확보의 근거로 활용될 수 있다.

다섯째로 디지털 기록환경의 특성과 기록관리의 장기 지속가능성을 종합적으로 반영하도록 설계되어야 한다. 오늘날 공공기록관리정책은 종이기록 중심의 보존관리에서 벗어나 전자기록, 행정정보 데이터세트, 업무시스템 기록, 웹기록, 메타데이터, 장기보존 포맷, 접근권한 관리 뿐만 아니라 인공지능 적용 등 복합적인 디지털 환경을 대상으로 한다. 따라서 정책평가도 물리적 보존 여부나 절차 준수 여부만이 아니라 디지털 기록이 생산부터 보존·활용까지 지속적으로 신뢰성을 유지할 수 있는지를 평가하는 체제를 갖추는 것이 필요하다.

이와 관련하여 OECD DAC의 지속가능성 기준을 참고할 수 있다. 지속가능성은 정책성고가 일시적 산출에 그치지 않고 장기적으로 유지될 수 있는지를 판단하는 기준이다. 기록관리정책에서 지속가능성은 시스템 구축 이후에도 기록의 무결성, 진본성, 이용가능성, 접근가능성이 유지되는지와 관련된다. 예를 들어 전자기록 보존정책은 시스템을 도입했다는 사실만으로 지속가능성이 확보되었다고 보기 어려우며 포맷 변환전략, 메타데이터 보존, 장기보존을 위한 인프라, 전문인력, 예산, 법·제도적 지원이 함께 유지되어야 한다.

NAA의 디지털 연속성 정책과 디지털 역량진단 역시 디지털 지속가능성을 평가하는 대표적 사례이다. 두 사례는 디지털 기록이 업무시스템에서 적절히 포착되고 필요한 메타데이터와 함께 관리되며 시간이 지나도 검색·접근·활용될 수 있어야 함을 강조한다. 이는 공공기록관리정책 평가에서 기록의 존재 여부뿐 아니라 기록이 업무 맥락과 함께 유지되는지, 시스템 변화 이후에도 접근 가능한지, 장기보존을 위한 기술적·관리적 조치가 마련되어 있는지를 평가해야 함을 보여준다. TNA의 IMA에서도 디지털 업무환경에 대한 대응은 중요한 평가요소로 고려된다. 기관의 정보관리 성숙도는 종이기록 관리역량만으로 판단될 수 없으며 디지털 정보의 분류, 저장, 보존, 접근, 위협관리 체계가 함께 평가되어야 한다. NARA 역시 전자기록관리와 디지털 이관 역량을 지속적으로 강조해 왔으며 이는 연방기관의 기록관리 책임이 디지털 환경에서 어떻게 재구성되는지를 보여준다.

디지털 환경은 정책목표, 성과경로, 평가자료, 책임성, 지속가능성 전반에 영향을 미치는 조건이다. 따라서 공공 기록관리정책 평가모형은 디지털 기록환경을 별도의 기술적 부문으로만 다루어서는 안 된다. 평가모형은 디지털 기록의 생산·획득·보존·이용 전 과정을 포괄하여 장기적으로 기록의 신뢰성과 접근성을 유지할 수 있는지를 판단할 수 있어야 한다.

마지막으로, 평가결과가 정책개선으로 환류될 수 있는지를 고려해야 한다. 정책평가는 성공과 실패를 판정하는 절차에 그치지 않고 정책의 설계와 집행을 개선하기 위한 학습 과정이어야 한다. 특히 공공기록관리정책은 법·제도, 표준, 시스템, 조직역량, 예산, 교육, 기관 간 협력체계가 복합적으로 작동하므로 평가결과가 후속 정책설계와 자원배분, 지침 개정, 기관 컨설팅 및 시스템 개선으로 연결될 필요가 있다.

세계은행 IEG의 평가는 사업의 성과를 판단하는 데 그치지 않고, 왜 그러한 결과가 나타났는지, 향후 유사 정책이나 사업에서 무엇을 개선해야 하는지를 분석한다는 점에서 평가결과의 활용과 학습을 강조한다는 것을 알 수 있다. 특히 사후평가를 통해 특정 정책이 목표를 달성했는지 여부만을 확인하는 것이 아니라 성과가 미흡했다면 그 원인이 제도 설계, 예산 부족, 기관 역량, 시스템 한계, 표준 부재, 교육 부족 중 어디에 있는지를 분석해야 한다.

TNA의 IMA는 개선계획과 후속조치를 통해 환류의 중요성을 보여준다. 평가결과가 기관의 정보관리 위험을 진단하는 데 그치지 않고 개선계획의 수립과 이행으로 이어질 때 평가의 실효성이 확보된다. NARA의 RMSA 역시 기관별 취약 영역을 확인하고 향후 기록관리 개선과 연계할 수 있다는 점에서 환류 기능을 가지며, NAA의 디지털 역량점검도 기관이 자신의 디지털 기록관리 수준을 진단하고 개선방향을 설정하도록 지원한다는 점에서 학습지향적 평가로 볼 수 있다.

OECD DAC 기준 중 효과성, 영향, 지속가능성 역시 평가결과의 환류와 연결된다. 정책이 단기적 산출을 달성했다라도 장기적 영향이나 지속가능성이 낮다면 다음에 이어질 후속정책 설계시 보완이 필요하다. 예를 들어 기록관리 교육사업이 단기간에 많은 교육 이수자를 배출했다라도 실제 기관의 기록관리 실무 개선이나 디지털 기록관리 역량 향상으로 이어지지 않았다고 판단된다면 교육내용, 대상, 방식, 사후지원 체계를 재설계해야 한다.

따라서 공공기록관리정책 평가모형은 평가결과의 활용할 수 있는 수단을 포함해야 한다. 평가결과가 법령 개정, 표준 개선, 예산 조정, 시스템 고도화, 기관 컨설팅, 교육 프로그램 개편, 후속 평가항목 보완으로 연결된다면 공공기록관리정책에 대한 평가는 행정적 의례가 아니라 정책학습의 장치로 기능할 수 있다.

4.2 공공기록관리정책 평가모형 설계원칙 종합

이상의 내용을 종합하면 공공기록관리정책의 평가모형은 해외 기준과 국가별 사례를 단순히 병렬적으로 수용하는 방식이 아니라 평가모형 개발에 필요한 핵심적인 원칙을 중심으로 재구성될 필요가 있다. OECD DAC는 정책목표와 평가기준의 정합성, 효과성, 지속가능성을 검토하는 데 유용하며, UNEG는 평가절차의 독립성, 투명성, 윤리성, 이해관계자 참여를 강조한다. 세계은행 IEG는 정책논리, 결과사슬, 증거기반 평가, 평가결과의 학습과 환류 측면에서 시사점을 제공한다. 각국의 사례는 국제적 평가원칙이 기록관리 분야에서 어떻게 구체화될 수 있는지를 보여준다.

앞서 도출한 여섯 개의 설계원칙은 평가시점에 따라 다음과 같이 적용될 수 있다. 사전평가에서는 각 고려사항별로 정책설계과정에서 타당성을 가졌는지, 정책을 추진하기 전 실행가능성이 있는지, 의도한대로 변화를 유도할 수 있는지 등을 검토해야 한다. 사후평가에서는 당초 설정한 성과경로가 실제로 작동했는지, 정책성과가 객관적 증거로 확인되는지, 평가절차가 공정하고 투명하게 수행되었는지, 디지털 기록관리 역량과 장기 지속가능성이 개선되었는지, 평가결과가 제도개선으로 연결되었는지를 판단해야 한다. <표 4>는 여섯 가지 설계원칙별로 핵심 평가질문과 평가시점별 고려사항을 정리하고 평가모형으로 구체화할 때 참조할 수 있는 국제기준과 해외 사례를 함께 제시한 것이다. 이 표는 향후 평가모형의 평가차원과 지표를 개발할 때 참조사항으로 활용될 수 있다.

〈표 4〉 기록관리정책 평가모형 개발을 위한 설계원칙 및 평가시점별 고려사항

설계원칙	핵심 평가질문	평가시점별 고려사항		참조사례
		사전평가	사후평가	
정책목표와 평가기준의 정합성	정책이 해결하려는 기록 관리 문제가 명확하며 평가기준이 공공가치와 연결되어 있는가?	정책문제, 정책목표, 추진수단, 평가기준이 행정 책임성, 알 권리, 기록정보 신뢰성, 디지털 장기보존 등과 부합하는지 검토한다.	정책성과가 당초 목표와 공공 기록관리의 핵심 가치에 비추어 적절하게 달성되었는지 확인한다.	OECD DAC(적절성·일관성), UNEG(인권 기반 접근), NARA, TNA, NAA
정책수단과 정책성과간 논리관계 명확화	정책수단이 어떤 과정을 거쳐 산출, 성과, 영향으로 이어지는가?	투입, 활동, 산출, 결과, 영향의 연결관계를 설정하고, 각 정책수단이 의도한 변화를 유발할 수 있는지 검토한다.	당초 설정한 성과경로가 실제로 작동했는지, 정책개입이 성과에 어느 정도 기여했는지 분석한다.	OECD DAC(효율성), 세계은행 IEG(결과사슬), NARA, TNA, NAA
평가자료와 증거기반의 확보 가능성	평가항목을 판단할 수 있는 신뢰성 있는 자료와 증거가 확보되는가?	자체보고, 점검표, 메타데이터, 시스템 로그, 이관·폐기·보존 기록, 접근권한 자료 등 필요 데이터를 사전에 설계한다.	자체보고에만 의존하지 않고 객관적 증거와 현장자료를 결합하여 성과를 검증한다.	NARA(RMSA), TNA(IMA), NAA
평가과정의 정당성·신뢰성 확보	평가가 공정하고 투명하게 수행되며 결과에 대한 설명책임이 확보되는가?	평가주체, 기준, 방법, 자료범위, 이해관계자 참여절차를 사전에 명확히 하고 공개 가능성을 검토한다.	평가결과뿐 아니라 방법, 근거자료, 한계, 해석과정을 공개하여 평가의 정당성과 책임성을 확보한다.	UNEG(신뢰성), NARA(SAORM), TNA
디지털 기록환경과 장기 지속가능성	정책이 디지털 기록의 생산, 획득, 보존, 이용 전 과정과 장기 지속가능성을 반영하는가?	전자기록, 행정정보 데이터 세트, 업무시스템 기록, 메타데이터, 접근권한, 보존포맷, 시스템 연계를 사전에 고려한다.	디지털 기록의 무결성, 진본성, 신뢰성, 이용가능성, 장기 접근성이 실제로 유지되고 개선되었는지 확인한다.	OECD DAC(지속가능성), NAA, TNA, NARA
평가결과의 환류와 제도개선 가능성	평가결과가 후속 정책설계와 제도개선으로 연결되는가?	평가결과가 법령, 표준, 예산, 시스템, 교육, 기관지원, 후속 평가항목 개선으로 이어질 수 있는 구조를 설계한다.	성과 미흡의 원인을 분석하고, 개선계획의 수립·이행·재평가로 연결되는 환류체계를 확인한다.	세계은행 IEG, UNEG(유용성), TNA, NARA

5. 결 론

이 연구는 공공기록관리정책 평가가 기관별 실적 확인이나 법령 준수 점검에 한정되어서는 안 된다는 문제의식에서 출발하여 향후 공공기록관리정책 평가모형 개발시 필요한 참조 준거를 도출하기 위해 정책평가 국제기준과 주요국의 기록관리 평가 사례를 분석하는데 중점을 두었다. 평가모형을 개발하기 위해서는 평가의 원칙 혹은 규범, 세부 평가기준, 성과관리와 환류를 포괄하는 평가체계가 필요하다는 관점에서 이 연구는 평가기준인 OECD DAC 평가기준, 평가규범인 UNEG 평가 규범 및 표준, 평가체제로써의 세계은행 IEG 평가 원칙과 미국 NARA, 영국 TNA, 호주 NAA의 평가제도를 평가의 목적과 성격, 구성요소, 평가방식, 활용과 한계라는 공통 분석기준에 따라 비교분석하였다. 분석 결과를 종합하여 정책목표와 평가기준의 정합성, 정책논리와 성과경로의 명확성, 증거 기반 평가자료의 확보, 평가절차의 독립성·투명성·책임성, 디지털 기록환경과 장기 지속가능성, 평가결과의 환류라는 여섯 개의 평가모형 설계시 고려사항을 사전평가와 사후평가의 관점에서 도출하고 각 고려사항별로 참조 가능한 국제기준과 사례를 제시하였다.

이 연구는 몇 가지 의의를 가진다. 첫째로 공공기록관리정책 평가를 법령 준수 중심의 행정점검에서 정책의 설계, 집행, 산출, 성과, 디지털 대응, 환류를 함께 검토하는 통합적 정책평가로 확장하기 위한 논거를 제시하였다. 둘째로 성격이 서로 다른 국제기준과 해외 사례를 공통의 분석기준에 따라 체계적으로 비교함으로써 각 기준과

사례가 평가모형의 어떤 요소를 채워줄 수 있는지를 명확하게 하였다. 셋째로 기록관리정책의 고유한 특성인 기록 생애주기, 기록 품질, 공공책무성, 이용가능성, 디지털 연속성을 반영한 고려사항을 도출함으로써 일반 정책 평가기준을 기록관리 분야에 맞게 구체화하고자 시도하였다.

다만 이 연구는 국제기준과 해외 평가 사례를 문헌자료 중심으로 분석하였으므로 각 제도가 실제 기관 현장에서 어떻게 운영되고 평가결과가 어느 정도 정책개선으로 이어지는지에 대한 경험적 검증에는 한계가 있다. 또한 이 연구는 평가모형 자체를 제안하기보다 그 설계에 필요한 설계원칙을 도출하는 데 초점을 두었으므로 각 사례별 특징을 개괄적으로 제시하는 데에 그쳤다. 사례별 교차비교를 통한 평가차원과 세부 지표, 측정척도, 평가방법, 평가주체를 포괄하는 평가모형의 구축은 후속 과제로 남는다.

이어질 후속 연구에서는 이 연구에서 도출한 여섯 가지 설계원칙을 토대로 평가기준, 평가항목, 지표, 평가방법, 평가주체를 포괄하는 공공기록관리정책 평가모형을 구체적으로 개발할 필요가 있다. 아울러 개발된 평가모형에 대해 기록관리 전문가, 정책평가 전문가, 공공기관 실무자 등을 대상으로 한 전문가 검토를 통해 타당성을 검증하고 중앙행정기관, 지방자치단체, 교육청, 공공기관 등 다양한 기관 유형에 시범 적용하여 자료 확보 가능성, 평가 부담, 결과 활용 가능성을 분석해야 한다. 궁극적으로 공공기록관리정책 평가는 기록관리 실적을 확인하는 절차를 넘어 공공기록이 행정책무성, 시민의 권리, 공공의 기억, 민주적 신뢰를 뒷받침하는 정책적 기반으로 기능하고 있는지를 판단하는 방향으로 발전해야 할 것이다.

참고문헌

공공기록물 관리에 관한 법률. 법률 제21066호

곽동철 (2001). 우리나라 기록관리정책의 발전전략에 관한 고찰. 한국기록관리학회지, 1(2), 7-30.

<https://doi.org/10.14404/JKSARM.2001.1.2.007>

국가기록원 (2024. 12). 2025년 기록관리 평가 계획.

출처: https://www.archives.go.kr/next/common/downloadBoardFile.do?board_seq=101792&board_file_seq=3

박홍윤, 박규동 (2023). 정책평가론: 실제와 사례. 고양: 대영문화사.

서혜란 (2009). 한국 공공기록관리 정책의 연대기적 검토. 한국기록관리학회지, 9(2), 189-214.

<https://doi.org/10.14404/JKSARM.2009.9.2.189>

설문원 (2020). 디지털 전환 시대의 공공기록정책: 기록자산으로서 정보의 관리. 기록학연구, 63, 5-36.

<http://doi.org/10.20923/kjas.2020.63.005>

이승일 (2014). 한국 국가기록관리체제의 개혁과 국가기록원 개편. 기록학연구, 41, 39-73.

<https://doi.org/10.20923/KJAS.2020.63.005>

이영학 (2009). 국가기록관리정책의 미래. 한국기록관리학회지, 9(2), 215-231. <https://doi.org/10.14404/JKSARM.2009.9.2.215>

이영학 (2012). 참여정부 기록관리정책의 특징. 기록학연구, 33, 113-153. <https://doi.org/10.20923/KJAS.2012.33.113>

이원규 (2010). 지방 공공기록관리의 정책과제. 기록학연구, 26, 95-102. <https://doi.org/10.20923/KJAS.2010.26.095>

정부업무평가 기본법. 법률 제21065호

ERA-LEARN (n.d.). Evaluation Types and Timing. Available:

https://www.era-learn.eu/support-for-partnerships/governance-administration-legal-base/monitoring-and-evaluation-of-european-r-and-i-partnerships_the-ripe-toolkit/monitoring-and-evaluation-basics/selecting-the-evaluation-type-timing-and-evaluation-questions

National Archives and Records Administration (n.d.a). Records Management Self-Assessment (RMSA). Available:

<https://www.archives.gov/records-mgmt/resources/rm-assessments>

- National Archives and Records Administration (n.d.b). Senior Agency Official for Records Management Annual Reports. Available: <https://www.archives.gov/records-mgmt/resources/saorm-reports>
- National Archives of Australia (2016). Check-up digital: Analysis of 2016 survey data. Available: <https://www.naa.gov.au/information-management/check-up-survey#previous>
- National Archives of Australia (2021). Digital Continuity 2020 Policy Report on implementation by the Australian Government agencies Available: <https://www.naa.gov.au/information-management/check-up-survey#previous>
- OECD (2010). Quality Standards for Development Evaluation, DAC Guidelines and Reference Series, OECD Publishing, Paris, <https://doi.org/10.1787/9789264083905-en>.
- OECD (2019). Better Criteria for Better Evaluation: Revised Evaluation Criteria Definitions and Principles for Use, OECD Publishing, Paris, <https://doi.org/10.1787/15a9c26b-en>.
- OECD (2021). Applying Evaluation Criteria Thoughtfully, OECD Publishing, Paris, <https://doi.org/10.1787/543e84ed-en>.
- OECD Development Assistance Committee (1991). Principles for evaluation of development assistance. Available: <http://www.oecd.org/dataoecd/secure/9/11/31779367.pdf>.
- Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2019). Evaluation: A Systematic Approach (8th ed.). California: SAGE Publications.
- Shepherd, E. & Yeo, G. (2003). Managing records: a handbook of principles and practice. London: Facet publishing.
- The National Archives (2021, July 15). Code of Practice on the management of records issued under section 46 of the Freedom of Information Act 2000.
- The National Archives (n.d.). Information Management Assessment. Available: <https://www.nationalarchives.gov.uk/information-management/manage-information/ima/>
- United Nations Evaluation Group (2016). Norms and Standards for Evaluation. United Nations Evaluation Group.
- United Nations Evaluation Group (2020). Ethical Guidelines for Evaluation. United Nations Evaluation Group.
- Vedung, E. (1997). Public Policy and Program Evaluation. New York: Routledge.
- Weiss, C. H. (1998). Evaluation: Methods for studying programs and policies. N.J.: Prentice Hall.
- World Bank Group (2019). World Bank Group Evaluation Principles. Independent Evaluation Group, World Bank Group.
- World Bank Independent Evaluation Group (2015). World Bank Group Evaluation Principles. World Bank.
- World Bank Independent Evaluation Group (n.d.). Evaluation Manual and Methods Resources. Available: <https://ieg.worldbankgroup.org/methods-resources>

• 국문 참고자료의 영어 표기

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

- Framework Act on Government Performance Evaluation. Act No.21065
- Kwack, Dong-Chul (2001). A study on developing strategies of the policy for archives and records management in Korea. Journal of Korean Society of Archives and Records Management, 1(2), 7-30. <https://doi.org/10.14404/JKSARM.2001.1.2.007>
- Lee, Seung-Il (2014). Reformation of the Korea's national archiving system. The Korean Journal of Archival Studies, 41, 39-73. <https://doi.org/10.20923/KJAS.2020.63.005>
- Lee, Won-Gyu (2010). Policy tasks for local public records management. The Korean Journal of Archival Studies, 26, 95-102. <https://doi.org/10.20923/KJAS.2010.26.095>
- Lee, Young-Hak (2009). The policy of national records management in the future. Journal of Korean Society of Archives and Records Management, 9(2), 215-231. <https://doi.org/10.14404/JKSARM.2009.9.2.215>
- Lee, Young-Hak (2012). The characteristics of records management policy during participation government (2003~2008). The Korean Journal of Archival Studies, 33, 113-153. <https://doi.org/10.20923/KJAS.2012.33.113>

Park, Hong-Yun & Park, Kyu-Dong (2023). Policy Evaluation. Goyang: Daeyeong Publishing.

Public Records Management Act. Act No.21066

Seol, Moon-won (2020). Managing information as records asset: Public records policies in the digital transformation era. The Korean Journal of Archival Studies, 63, 5-36. <http://doi.org/10.20923/kjas.2020.63.005>

Suh, Hye Ran (2009). A chronological review of the public records management policies in Korea from 1948 to present. Journal of Korean Society of Archives and Records Management, 9(2), 189-214.

<https://doi.org/10.14404/JKSARM.2009.9.2.189>

The National Archives of Korea (2024, December). 2025 Records Management Evaluation Plan. Available:

https://www.archives.go.kr/next/common/downloadBoardFile.do?board_seq=101792&board_file_seq=3