



Informatica®  
10.4.0

# Data Quality Integration for PowerCenter **가이드**

## Informatica Data Quality Integration for PowerCenter 가이드

10.4.0

2019년12월

© 저작권 Informatica LLC 2009, 2020

이 소프트웨어와 설명서는 사용 및 공개에 대한 제한 사항이 포함되어 있는 별도의 사용권 계약에 따라서만 제공됩니다. 본 문서의 어떤 부분도 Informatica LLC의 사전 통지 없이 어떠한 형태나 수단(전자적, 사진 복사, 녹음 등)으로 복제되거나 전송될 수 없습니다.

미국 정부 권한. 미국 정부 고객에게 제공되는 프로그램, 소프트웨어, 데이터베이스, 관련 문서 및 기술 데이터는 해당하는 연방 입수 규정 및 기관별 보안 규정에 따라 "상용 컴퓨터 소프트웨어" 또는 "상용 기술 데이터"입니다. 따라서 사용, 복제, 공개, 수정 및 조정은 해당하는 정부 계약에 규정된 제한 사항 및 라이선스 조건을 따르며, 정부 계약 조건에 의해 적용 가능한 한도 내에서, FAR 52.227-19, 상용 소프트웨어 라이선스에 규정된 추가 권한이 적용됩니다.

Informatica, PowerCenter 및 Informatica 로고는 미국과 전 세계 여러 관할 국가에서 Informatica LLC의 상표 또는 등록 상표입니다. Informatica 상표의 현재 목록은 <https://www.informatica.com/trademarks.html> 웹에서 확인할 수 있습니다. 다른 회사 및 제품명은 해당 소유자의 상표 또는 등록 상표일 수 있습니다.

이 소프트웨어 및/또는 설명서의 일부에는 타사의 저작권이 적용될 수 있습니다. 필요한 타사 고지 사항은 제품에 포함되어 있습니다.

이 설명서의 정보는 예고 없이 변경될 수 있습니다. 이 문서에서 문제가 발견되는 경우 [infa\\_documentation@informatica.com](mailto:infa_documentation@informatica.com)으로 보고해 주십시오.

Informatica 제품은 제품이 제공될 당시의 계약 조건에 따라 보증됩니다. Informatica는 상품성과 특정 목적에의 적합성에 대한 보증 그리고 비침해에 대한 보증 또는 조건을 포함하여 어떠한 종류의 명시적이거나 묵시적인 보증 없이 이 문서의 정보를 "있는 그대로" 제공합니다.

발행 날짜: 2020-02-07

# 목차

<b>서문</b>	<b>4</b>
Informatica 리소스	4
Informatica 네트워크	4
Informatica 기술 자료	4
Informatica 설명서	4
Informatica Product Availability Matrix	5
Informatica Velocity	5
Informatica Marketplace	5
Informatica 글로벌 고객 지원 센터	5
<b>장 1: Informatica Data Quality 통합</b>	<b>6</b>
Data Quality 통합 개요	6
데이터 품질 피처 및 기능	7
통합 설치 프로그램	7
PowerCenter로 개체 내보내기	8
참조 데이터 설치	8
제품 통합에 대한 규칙 및 지침	8
PowerCenter에서의 데이터 품질 변환 동작	9
<b>장 2: 설치</b>	<b>10</b>
설치 개요	10
시스템 요구 사항	10
설치하기 전에	10
서버 플러그 인 설치	11
Windows에 설치	11
UNIX에 설치	12
클라이언트 플러그 인 설치	12
PowerCenter를 통해 플러그 인 등록	13
<b>장 3: 참조 데이터 요구 사항</b>	<b>14</b>
참조 데이터 개요	14
참조 테이블 데이터 요구 사항	15
주소 참조 데이터 요구 사항	15
ID 인구집단 데이터 요구 사항	17
확률 모델 및 분류자 모델 요구 사항	18
PowerCenter의 확률 모델 및 분류자 모델 파일 위치	18
PowerCenter 통합 서비스 메모리 할당	19
<b>인덱스</b>	<b>20</b>

# 서문

*Informatica Data Quality Integration for PowerCenter® 사용자 가이드*에서 PowerCenter로 데이터 품질 개체를 설정하고 사용하는 방법에 대한 지침을 확인하십시오. Informatica Data Quality 환경의 참조 데이터 개체 및 파일과 함께 데이터 품질 변환이 포함된 매핑을 PowerCenter 환경으로 내보낼 수 있습니다. 이 가이드에는 데이터 품질 개체에 적용되는 사전 조건과 호환되는 제품 릴리스에 통합 구성 요소를 설치하는 단계가 설명되어 있습니다.

## Informatica 리소스

Informatica는 Informatica Network 및 기타 온라인 포털을 통해 다양한 범위의 제품 리소스를 제공합니다. 리소스를 통해 Informatica 제품 및 솔루션을 최대한 활용하고 다른 Informatica 사용자 및 주제별 전문가로부터 배울 수 있습니다.

### Informatica 네트워크

Informatica Network는 Informatica 기술 자료, Informatica 글로벌 고객 지원 센터 등 여러 리소스로 연결되는 관문입니다. Informatica Network를 시작하려면 <https://network.informatica.com>을 방문하십시오.

Informatica Network 멤버인 경우 다음 옵션이 가능합니다.

- 기술 자료에서 제품 리소스를 검색할 수 있습니다.
- 제품 사용 가능 여부에 대한 정보를 봅니다.
- 지원 사례를 생성하고 검토할 수 있습니다.
- 거주 지역의 Informatica 사용자 그룹 네트워크를 검색하고 동료와 협업 관계 유지

### Informatica 기술 자료

Informatica 기술 자료를 사용하여 사용 방법 문서, 모범 사례, 비디오 자습서, 자주 묻는 질문에 대한 답변 등 제품 리소스를 확인할 수 있습니다.

기술 자료를 검색하려면 <https://search.informatica.com>을 방문하십시오. 기술 자료에 대한 질문, 의견 또는 아이디어가 있는 경우 [KB\\_Feedback@informatica.com](mailto:KB_Feedback@informatica.com)을 통해 Informatica 기술 자료 팀에 문의해 주시기 바랍니다.

### Informatica 설명서

Informatica 설명서 포털에서 확장된 설명서 라이브러리를 탐색하여 현재 및 최근 제품 릴리스를 확인할 수 있습니다. 설명서 포털을 탐색하려면 <https://docs.informatica.com>을 방문하십시오.

제품 설명서에 대한 질문, 의견 또는 아이디어가 있는 경우 [infa\\_documentation@informatica.com](mailto:infa_documentation@informatica.com)에서 Informatica 설명서 팀에 문의해 주시기 바랍니다.

## Informatica Product Availability Matrix

PAM(Product Availability Matrix)은 제품 릴리스에서 지원하는 운영 체제 버전, 데이터베이스 및 데이터 소스 유형과 대상을 나타냅니다.

<https://network.informatica.com/community/informatica-network/product-availability-matrices>에서 Informatica PAM을 찾을 수 있습니다.

## Informatica Velocity

Informatica Velocity는 수백 가지 데이터 관리 프로젝트의 실제 경험을 토대로 Informatica 전문 서비스업에서 개발한 팁과 모범 사례 모음입니다. Informatica Velocity는 전 세계의 조직과 협력하여 성공적인 데이터 관리 솔루션을 계획, 개발, 배포 및 유지 관리하는 Informatica 컨설턴트의 포괄적인 지식을 보여줍니다.

Informatica Velocity 리소스는 <http://velocity.informatica.com>에서 확인할 수 있습니다. Informatica Velocity에 대한 질문, 주석 또는 아이디어가 있으시면 Informatica 전문 서비스업([ips@informatica.com](mailto:ips@informatica.com))에 문의하십시오.

## Informatica Marketplace

Informatica Marketplace는 Informatica 구현을 확대 및 개선하기 위한 솔루션을 찾을 수 있는 포럼입니다.

Marketplace에서 Informatica 개발자와 파트너가 제공하는 수백 개의 솔루션을 활용하여 생산성을 향상시키고 프로젝트의 구현에 걸리는 시간을 줄일 수 있습니다. <https://marketplace.informatica.com>에서 Informatica Marketplace를 찾을 수 있습니다.

## Informatica 글로벌 고객 지원 센터

전화 또는 Informatica Network를 통해 글로벌 지원 센터에 문의할 수 있습니다.

해당 지역의 Informatica 글로벌 고객 지원 전화 번호는 Informatica 웹 사이트 (<https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>)를 방문하여 찾을 수 있습니다.

Informatica Network에서 온라인 지원 리소스를 찾으려면 <https://network.informatica.com>을 방문하고 eSupport 옵션을 선택하십시오.

## 제 1 장

# Informatica Data Quality 통합

이 장에 포함된 항목:

- [Data Quality 통합 개요, 6](#)
- [데이터 품질 피처 및 기능, 7](#)
- [통합 설치 프로그램, 7](#)
- [PowerCenter로 개체 내보내기, 8](#)
- [참조 데이터 설치, 8](#)
- [제품 통합에 대한 규칙 및 지침, 8](#)
- [PowerCenter에서의 데이터 품질 변환 동작, 9](#)

## Data Quality 통합 개요

Informatica 사용자는 Informatica Data Quality에서 매핑을 생성하고 실행할 수 있습니다. 사용자는 매핑을 모델 리포지토리에 저장합니다. 또한 매핑과 모든 관련된 참조 데이터 개체 또는 파일을 내보내 PowerCenter에서 사용할 수 있습니다.

PowerCenter 사용자는 PowerCenter 리포지토리로 매핑을 가져오고 PowerCenter 세션에서 매핑을 실행할 수 있습니다. 매핑을 사용하면 데이터의 콘텐츠와 구조를 분석하고 개선할 수 있습니다.

사용자가 Data Quality 10.4에서 생성한 매핑은 다음 버전의 PowerCenter에서 사용할 수 있습니다.

- PowerCenter 10.4.0.
- 최신 HotFix가 설치된 PowerCenter 10.2.
- 최신 HotFix가 설치된 PowerCenter 10.1.1.
- PowerCenter 10.1.

최신 버전의 Data Quality 및 PowerCenter를 사용하는 경우 추가 설치 없이 모델 리포지토리에서 제공된 개체를 가져올 수 있습니다. 이전 버전의 PowerCenter를 사용하는 경우 최신 버전의 Informatica Data Quality 통합 플러그 인을 설치하고 등록합니다.

# 데이터 품질 피처 및 기능

Informatica Data Quality는 개체를 생성하고 모델 리포지토리에 개체를 저장하고 매핑을 실행하기 위해 Informatica 응용 프로그램을 사용합니다.

현재의 제품 버전을 사용하지 않는 경우 PowerCenter용 Informatica Data Quality 플러그 인을 설치해야 할 수 있습니다.

다음 목적을 달성하는 매핑을 설계하고 실행하려면 Informatica Data Quality를 사용하십시오.

- 프로파일 데이터. 프로파일링은 데이터의 콘텐츠와 구조를 나타냅니다. 프로파일링은 데이터의 강점과 단점을 파악하여 프로젝트 계획을 정립하는 데 도움이 될 수 있으므로 어떤 데이터 프로젝트에서도 핵심 단계로 꼽힙니다.
- 데이터 품질 검토를 위해 성과 기록표를 작성하십시오. 성과 기록표는 프로파일의 품질 측정용 그래픽으로 표현합니다.
- 데이터 값 표준화. 프로파일을 실행할 때 찾는 오류와 불일치를 제거하기 위해 데이터를 표준화하십시오. 구두점, 형식 및 맞춤법에서의 변형을 표준화할 수 있습니다. 예를 들어 도시, 도 및 우편 번호 값이 일관하도록 보장할 수 있습니다.
- 레코드 구문 분석. 레코드 구조를 개선하고 데이터에서 추가적인 정보를 파생하기 위해 데이터 레코드를 구문 분석하십시오. 자유 형식 데이터의 단일 필드를 다양한 정보 유형을 포함한 필드로 분할할 수 있습니다. 또한 레코드에 정보를 추가할 수도 있습니다. 예를 들어 고객 레코드를 개인 또는 비즈니스 고객으로 플래그를 설정할 수 있습니다.
- 우편 주소 유효성 검사. 주소 유효성 검사는 우편 주소 데이터의 정확도와 전달 가능성을 평가 및 향상시킵니다. 주소 유효성 검사는 주소 레코드를 국립 우편 전송 서비스의 참조 데이터와 비교함으로써 주소의 오류를 수정하고 부분 주소를 완성할 수 있게 합니다. 주소 유효성 검사를 통해 메일 전송을 가속화하고 메일 비용을 감소시키는 우편 정보를 추가할 수도 있습니다.
- 중복 레코드 찾기. 중복 레코드 분석은 선택된 데이터 열에서 유사하거나 일치하는 값을 찾기 위해 일련의 레코드를 서로 비교합니다. 필드 값 사이의 높은 일치율을 나타내는 유사성 수준을 설정합니다. 또한 일치 계산에서 각 열에 고정된 상대 가중치도 설정할 수 있습니다. 예를 들어 이름 정보보다 성에 관한 정보가 우선하도록 설정할 수 있습니다.
- 데이터 품질 규칙 생성 및 실행. Informatica는 프로젝트 목표에 맞게 실행하거나 편집할 수 있는 사전 구축 규칙을 제공합니다. Developer tool에서 규칙을 생성할 수 있습니다.
- Informatica 사용자와 공동 작업. 모델 리포지토리에 추가하는 규칙과 참조 데이터 테이블은 개발자 도구와 분석 도구에서 사용자가 사용할 수 있습니다. 사용자는 프로젝트에 대한 공동 작업을 할 수 있고 다른 사용자는 다른 단계의 프로젝트에서 개체 소유권을 보유할 수 있습니다.
- PowerCenter에 매핑 내보내기. 물리적 데이터 통합을 위해 메타데이터를 다시 사용하거나 웹 서비스를 생성하기 위해 PowerCenter에 매핑을 내보낼 수 있습니다.

## 통합 설치 프로그램

개체를 생성하고 저장하는 데 다른 버전의 Data Quality 및 PowerCenter를 사용하는 경우 PowerCenter용 Informatica Data Quality 플러그 인을 설치하십시오. Informatica에서 플러그 인 설치 프로그램을 다운로드하십시오.

사용하고자 하는 PowerCenter의 버전에 대한 플러그 인 버전을 설치하십시오. 플러그 인에는 클라이언트 설치 프로그램과 서버 설치 프로그램이 있습니다.

리포지토리의 맵렛과 매핑을 보고 이를 세션에 추가하려면 **PowerCenter** 클라이언트 시스템에서 클라이언트 설치 프로그램을 실행하십시오. 워크플로우에서 세션을 실행하려면 **PowerCenter** 통합 서비스 시스템에서 서버 설치 프로그램을 실행하십시오.

통합 서비스 시스템에 플러그인을 설치한 후에는 **PowerCenter** 리포지토리를 통해 해당 플러그인을 등록하십시오.

## PowerCenter로 개체 내보내기

**Informatica** 모델 리포지토리에서 맵렛과 매핑 등의 개체를 내보내려면 **Developer tool**을 사용하십시오. 모델 리포지토리에서 **PowerCenter** 리포지토리로 개체를 내보낼 수 있으며 모델 리포지토리에서 파일 시스템으로 개체를 내보낼 수 있습니다. XML 개체를 **PowerCenter** 리포지토리로 가져오려면 **PowerCenter Repository Manager**를 사용하십시오.

- 모델 리포지토리에서 개체 내보내기에 대한 자세한 내용은 *Informatica Developer 매핑 가이드* 및 *Informatica Developer Tool 가이드*를 참조하십시오.
- **PowerCenter** 리포지토리로 개체 가져오기에 관한 정보는 *PowerCenter 리포지토리 가이드*를 참조하십시오.

## 참조 데이터 설치

맵렛 또는 매핑이 참조 데이터를 읽는 경우 **PowerCenter** 통합 서비스 시스템에 참조 데이터 파일을 복사하거나 설치하십시오.

맵렛이나 매핑을 내보낼 때 모델 리포지토리에서 참조 데이터를 내보내기 위해 개발자 도구를 사용할 수 있습니다. **Developer tool**은 참조 데이터를 압축된 파일로 내보냅니다. 파일을 추출하고 통합 서비스 시스템에 해당 참조 데이터를 복사하십시오.

내보내기 프로세스에는 주소 참조 데이터 파일이나 ID 채우기 파일이 포함되지 않습니다. 내보내는 맵렛 또는 매핑이 주소 참조 데이터 파일 또는 ID 채우기 데이터 파일을 읽는 경우 **PowerCenter** 통합 서비스 시스템에 파일을 설치하십시오.

## 제품 통합에 대한 규칙 및 지침

다른 버전의 **Informatica Data Quality**에서 다른 버전의 **PowerCenter**로 데이터 품질 개체를 내보낼 수 있습니다. 개체를 내보내려는 **Informatica Data Quality** 버전에 대한 통합 플러그인을 설치합니다.

**Informatica Data Quality** 및 **PowerCenter**를 통합할 때 다음과 같은 규칙 및 지침을 고려합니다.

- **PowerCenter** 리포지토리 제품 버전이 모델 리포지토리 제품 버전과 일치하는 경우에는 통합 플러그인을 설치하지 마십시오.
- 단일 **PowerCenter** 시스템에서 2개의 **Data Quality** 통합 플러그인을 동시에 실행할 수 없습니다. **PowerCenter**는 가장 최근에 설치한 플러그인을 사용합니다. 예를 들어 **PowerCenter 10.1.1** 시스템에 **Data Quality 10.2** 통합 플러그인을 설치하는 경우 버전 **10.1.1** 모델 리포지토리에서 더 이상 데이터 품질 개체를 가져오고 실행할 수 없습니다. **10.2** 플러그인을 설치하기 전에 **PowerCenter** 리포지토리로 가져온 데이터 품질 개체를 계속 실행할 수 있습니다.



- PowerCenter 리포지토리에서 데이터 품질 맵셋 또는 매핑을 편집할 수 없습니다.

## PowerCenter에서의 데이터 품질 변환 동작

모델 리포지토리에서 PowerCenter 리포지토리로 데이터 품질 맵셋 또는 매핑을 가져올 때 데이터 품질 변환 동작이 변경됩니다.

### 맵셋으로의 변환

다음 변환은 PowerCenter 리포지토리로 가져올 때 맵셋으로 변환할 수 있습니다.

- 통합
- 예외
- 키 생성기
- 일치

이러한 변환을 하나 이상 포함한 매핑을 가져오는 경우 PowerCenter는 각 변환을 확장된 구성을 포함하는 맵셋으로 변환합니다.

이러한 변환을 하나 이상 포함한 맵셋을 가져오는 경우 PowerCenter는 각 변환의 구성을 여러 변환으로 확장합니다.

PowerCenter는 변환 논리를 완료하기 위해 데이터 품질 변환 또는 맵셋 확장에 변환을 추가할 수도 있습니다. 예를 들어 PowerCenter는 추가 변환을 이용해 일치 변환을 확장합니다.

### 결정 변환 고려 사항

변환 스크립트에 시스템 매개 변수가 포함된 경우 결정 변환이 포함된 매핑 또는 맵셋을 PowerCenter로 내보낼 수 없습니다. 내보내기 작업은 시스템 매개 변수를 PowerCenter가 사용할 수 있는 값으로 변환할 수 없습니다. 시스템 매개 변수를 사용하는 결정 변환이 포함된 매핑 또는 맵셋을 내보내기 전에 해당 매개 변수를 적절한 값으로 바꿉니다.

### 일치 변환 고려 사항

매핑이 ID 인덱스 데이터를 읽거나 데이터베이스 테이블에 쓰는 경우 PowerCenter에서 ID 일치 매핑을 실행할 수 없습니다. PowerCenter에서 ID 일치 분석을 수행하려면 ID 인덱스 데이터를 임시 파일에 쓸 수 있도록 Developer tool에서 일치 변환을 구성합니다.

### 다중 전략 변환

Developer tool의 다음 각 변환 내에서 다중 데이터 변환 전략을 정의할 수 있습니다.

- 대/소문자
- 분류자
- 결정
- 키 생성기
- 라벨러
- 일치
- 병합
- 파서
- 표준화

다중 전략 변환을 포함한 맵셋이나 매핑을 가져올 때는 각 전략이 개별 변환으로 변환됩니다.

## 제 2 장

# 설치

이 장에 포함된 항목:

- [설치 개요, 10](#)
- [시스템 요구 사항, 10](#)
- [설치하기 전에, 10](#)
- [서버 플러그 인 설치, 11](#)
- [클라이언트 플러그 인 설치, 12](#)
- [PowerCenter를 통해 플러그 인 등록, 13](#)

## 설치 개요

Data Quality 통합 플러그 인에는 서버 구성 요소 및 클라이언트 구성 요소가 있습니다.

플러그 인을 설치 및 등록하려면 다음 단계를 모두 수행하십시오.

1. PowerCenter 통합 서비스 시스템에서 서버 설치 프로그램을 실행하십시오.
2. PowerCenter 클라이언트 시스템에서 클라이언트 설치 프로그램을 실행하십시오.
3. PowerCenter 리포지토리를 통해 `IDQZIntegration.xml` 파일을 등록하십시오.

## 시스템 요구 사항

Informatica Data Quality 통합 플러그 인의 시스템 요구 사항은 설치하는 PowerCenter 응용 프로그램 및 서비스의 요구 사항과 동일합니다.

PowerCenter 선행 조건 및 시스템 요구 사항에 관한 정보는 PowerCenter 설치 문서를 참조하십시오.

## 설치하기 전에

PowerCenter에서 참조 데이터를 사용하는 경우 다음 설치 프로그램 동작을 고려하십시오.

통합 플러그 인 설치 프로그램은 PowerCenter 통합 서비스 시스템에 속성 파일을 씁니다. 이 속성 파일은 데이터 품질 매핑이 읽을 수 있는 참조 데이터 파일의 위치를 지정합니다.

설치 프로그램은 다음 파일을 씁니다.

- AD50.cfg. 주소 참조 데이터의 구성 속성 저장.
- CLASSIFIER.properties. 분류자 모델 데이터의 구성 속성 저장.
- IDQTx.cfg. ID 인구집단 데이터의 구성 속성 저장.
- NER.properties. 확률 모델 데이터의 구성 속성 저장.

설치 프로그램은 파일을 다음 디렉터리에 추가합니다.

[PowerCenter\_installation]/server/bin

### 주소 참조 데이터 선행 조건

주소 참조 데이터를 설치하기 전에 **PowerCenter** 통합 서비스를 중지하십시오. 데이터를 설치한 후 해당 서비스를 다시 시작하십시오.

## 서버 플러그 인 설치

Informatica Data Quality 통합 서버 플러그 인 설치 프로그램은 **PowerCenter** 통합 서비스 디렉터리 구조에 데이터 품질 변환 응용 프로그램 파일을 추가합니다.

### Windows에 설치

Windows 플랫폼에서 **PowerCenter** 통합 서비스 시스템에 플러그 인을 추가하려면 **Windows** 서버 설치 프로그램을 실행하십시오.

설치 전에 모든 응용 프로그램을 닫으십시오.

1. 시스템 사용자 계정으로 시스템에 로그인하십시오.
2. **Windows**에 대한 통합 서버 파일을 추출하십시오.
3. 파일을 추출했던 디렉터리로 이동하여 **install.exe**를 실행하십시오.
4. 언어를 선택하고 **확인**을 클릭하십시오.
5. **시작** 화면에서 **다음**을 클릭하십시오.
6. **PowerCenter** 서비스가 실행되고 있지 않은지 확인한 후 **확인**을 클릭하십시오.
7. **PowerCenter** 설치 디렉터리에 대한 경로를 지정하십시오. 설치 프로그램은 이 디렉터리에 대한 기본 경로를 제공합니다. 다른 경로를 선택하려면 **선택**을 클릭하십시오.
8. **다음**을 클릭합니다.
9. 설치 전 요약을 검토하십시오. 오류를 수정하려면 **이전**을 클릭하십시오.
10. **설치**를 클릭합니다.
11. 설치 후 요약을 검토하십시오.  
설치 작업에 관한 자세한 정보와 설치된 구성 요소에 대한 구성 속성을 보려면 설치 로그 파일을 확인하십시오.
12. **완료**를 클릭합니다.

## UNIX에 설치

UNIX 플랫폼에서 PowerCenter 통합 서비스 시스템에 플러그 인을 추가하려면 UNIX 서버 설치 프로그램을 실행하십시오.

1. 시스템 사용자 계정으로 시스템에 로그인하십시오.
2. 기타 모든 응용 프로그램을 닫으십시오.
3. UNIX에 대한 통합 서버 파일을 추출하십시오.
4. UNIX 셸을 열고 설치 파일을 추출했던 디렉터리로 이동하십시오.
5. `sh install.bin`을 입력하십시오.
6. 영어를 설치 언어로 수락하려면 **Enter** 키를 누르십시오.
7. 시작 텍스트를 읽고 **Enter** 키를 누르십시오.
8. PowerCenter 설치 디렉터리에 대한 절대 경로를 입력하거나 **Enter** 키를 눌러 기본 경로를 수락하십시오.
9. 설치 전 요약을 검토하십시오. 오류를 수정하려면 **back**을 입력하고 **Enter** 키를 누르십시오.
10. 설치를 시작하려면 **Enter** 키를 누르십시오.
11. 설치 후 요약을 검토하십시오.  
설치 작업에 관한 자세한 정보와 설치된 구성 요소에 대한 구성 속성을 보려면 설치 로그 파일을 확인하십시오.
12. 설치 프로세스를 완료하려면 **Enter** 키를 누르십시오.

## 클라이언트 플러그 인 설치

PowerCenter 클라이언트 시스템에서 Data Quality 통합 클라이언트 설치 프로그램을 실행하십시오.

1. 기타 모든 응용 프로그램을 닫으십시오.
2. 통합 클라이언트 파일을 추출하십시오.
3. 파일을 추출했던 디렉터리를 여십시오.
4. `install.exe`를 실행하십시오.
5. 시작 화면에서 **다음**을 클릭하십시오.  
설치 프로그램은 설치를 계속하기 전에 PowerCenter 클라이언트 응용 프로그램을 닫으라는 프롬프트를 표시합니다.
6. PowerCenter 클라이언트 응용 프로그램이 실행되고 있지 않은지 확인한 후 **확인**을 클릭하십시오.
7. PowerCenter 클라이언트 디렉터리에 대한 경로를 지정하십시오. 설치 프로그램은 이 디렉터리에 대한 기본 경로를 제공합니다. 다른 경로를 선택하려면 **선택**을 클릭하십시오.
8. **다음**을 클릭합니다.
9. 설치 전 요약을 검토하십시오. 오류를 수정하려면 **이전**을 클릭하십시오.
10. **설치**를 클릭합니다.
11. 설치 후 요약을 검토하십시오.  
설치 작업에 관한 자세한 정보와 설치된 구성 요소에 대한 구성 속성을 보려면 설치 로그 파일을 확인하십시오.
12. **완료**를 클릭합니다.

# PowerCenter를 통해 플러그 인 등록

데이터 품질 개체를 포함할 PowerCenter 리포지토리를 통해 플러그 인을 등록하십시오.

통합 설치 프로그램은 IDQZIntegration.xml 플러그 인 파일을 PowerCenter 통합 서비스 시스템의 \$INFA\_HOME\server\bin\native 디렉터리에 씁니다. 이 파일에 액세스할 수 있는지 확인하십시오. 필요한 경우 PowerCenter 통합 서비스 시스템에 이 파일을 복사하거나 FTP로 전송합니다.

- ▶ 플러그인을 등록하려면 *pmrep RegisterPlugin* 명령을 실행합니다. 명령을 실행할 때 -N 옵션을 포함합니다.

*pmrep RegisterPlugin* 명령에 대한 자세한 내용은 *Informatica 명령 참조*를 참조하십시오.

**참고:** 이 플러그 인을 등록하기 위한 적합한 권한이 없는 경우 PowerCenter 리포지토리 서비스 관리자에게 문의하십시오.

## 제 3 장

# 참조 데이터 요구 사항

이 장에 포함된 항목:

- [참조 데이터 개요, 14](#)
- [참조 테이블 데이터 요구 사항, 15](#)
- [주소 참조 데이터 요구 사항, 15](#)
- [ID 인구집단 데이터 요구 사항, 17](#)
- [확률 모델 및 분류자 모델 요구 사항, 18](#)

## 참조 데이터 개요

데이터 품질 변환은 참조 데이터를 읽을 수 있습니다. PowerCenter 리포지토리에 있는 어떤 변환이 참조 데이터를 읽는 경우 해당 참조 데이터 파일이 PowerCenter 환경에 있는지 확인하십시오.

PowerCenter의 데이터 품질 변환은 다음 유형의 참조 데이터를 사용할 수 있습니다.

### 참조 데이터 파일

표준 버전의 비즈니스 용어와 대체 버전의 용어를 포함한 텍스트 파일. 참조 데이터 파일을 편집할 수 있습니다.

### 주소 참조 데이터 파일

한 국가의 모든 우편 주소에 대한 완전한 데이터를 포함하는 데이터 파일. 주소 참조 데이터 파일을 PowerCenter 통합 서비스 시스템에 설치합니다. 주소 참조 데이터 파일은 편집할 수 없습니다.

### ID 인구집단 파일

다른 유형의 ID를 정의하고 일치 변환에서 사용하는 알고리즘을 포함하는 데이터 파일. 인구집단 파일을 PowerCenter 서비스 시스템에 설치합니다. 인구집단 파일은 편집할 수 없습니다.

### 콘텐츠 집합

하나 이상의 변환에 의해 지정된 참조 데이터를 포함하는 데이터 파일. 콘텐츠 집합을 읽는 매핑을 내보낼 때 해당 내보내기 프로세스는 변환 메타데이터가 포함된 콘텐츠 집합 데이터를 내보냅니다. 분류자 및 확률 모델 파일의 경우 해당 프로세스는 파일 데이터를 Data Quality 디렉터리 구조에서 내보냅니다.

## 참조 테이블 데이터 요구 사항

PowerCenter에서 참조 테이블을 읽는 매핑을 실행하기 전에 테이블에 대한 열 데이터를 PowerCenter 서비스 시스템에 복사하십시오.

모델 리포지토리에서 매핑을 내보내면 내보내기 작업이 참조 테이블 열 데이터를 하나 이상의 파일에 기록합니다. 내보내기 작업은 지정된 위치의 압축 파일에 파일을 추가합니다. 압축 파일을 추출하고 열 데이터 파일을 찾으십시오. 파일 이름은 .dic로 끝납니다. 파일을 PowerCenter 서비스 시스템에 복사합니다.

PowerCenter 디렉터리 구조의 참조 데이터 파일 위치는 모델 리포지토리의 참조 테이블 위치와 일치해야 합니다.

다음 디렉터리 형식을 사용하여 참조 테이블 데이터 파일의 올바른 위치를 결정합니다.

<PowerCenter\_Root\_Directory>\services\<Model\_Repository\_Project\_Name>\<Model\_Repository\_Project\_Folder\_Name>

### 참조 테이블 데이터 내보내기에 대한 규칙과 지침

참조 테이블 데이터를 읽는 매핑을 내보내는 경우 다음 규칙 및 지침을 고려하십시오.

- 모델 리포지토리에서 매핑을 내보내려면 내보내기 마법사에서 **Informatica > PowerCenter** 옵션을 선택합니다. **PowerCenter** 옵션을 선택하면 내보내기 작업에서 관리되는 참조 테이블 데이터와 관리되지 않는 참조 테이블 데이터를 파일 시스템에 기록합니다.
- 내보내기 작업은 각 참조 테이블의 첫 번째 열을 유효한 열로 지정합니다.
- PowerCenter는 UTF-8 형식의 참조 데이터 파일을 읽습니다.

### 참조 데이터 파일 위치를 정의하기 위해 환경 변수 사용

통합 서비스가 참조 데이터 파일을 읽는 데 사용하는 경로를 설정하기 위해 PowerCenter 통합 서비스 시스템의 **INFA\_CONTENT** 환경 변수를 정의할 수 있습니다.

PowerCenter 통합 서비스 시스템의 필요한 위치에 파일을 설치할 수 없는 경우에 이 환경 변수를 사용하십시오. 이 파일을 PowerCenter 통합 서비스가 읽을 수 있는 위치에 설치하고 설치된 경로가 포함된 **INFA\_CONTENT** 환경 변수를 생성하십시오.

## 주소 참조 데이터 요구 사항

주소 유효성 검사를 수행하는 매핑은 주소 참조 데이터를 읽습니다. Informatica에서 주소 참조 데이터를 구매하고 다운로드하십시오. 데이터 파일을 PowerCenter 통합 서비스 시스템에 설치하십시오.

다음 테이블에는 설치할 수 있는 주소 참조 데이터의 유형이 설명되어 있습니다.

참조 데이터 유형	사용 시기
일괄	일반적인 주소 유효성 검사 작업에 사용.
CAMEO	인구통계학적 고객 데이터를 주소 레코드에 추가하는 데 사용.

참조 데이터 유형	사용 시기
인증됨	<p>주소 레코드가 메일 전송 서비스에서 정의하는 인증 표준을 충족하는지 확인하는 데 사용.</p> <p>다음 국가는 인증 표준을 정의합니다.</p> <ul style="list-style-type: none"> <li>- 오스트레일리아. AMAS(Address Matching Approval System) 표준에 따라 메일 인증.</li> <li>- 캐나다. SERP(Software Evaluation And Recognition Program) 표준에 따라 메일 인증.</li> <li>- 프랑스. SNA(National Address Management Service) 표준에 따라 메일 인증.</li> <li>- 뉴질랜드. SendRight 표준에 따라 메일 인증.</li> <li>- 미국. CASS(Coding Accuracy Support System) 표준에 따라 메일 인증.</li> </ul>
제안 목록	<p>불완전한 주소 레코드의 유효한 대체 버전을 찾는 데 사용. 실시간으로 단일 주소 레코드를 처리하기 위해 주소 유효성 검사 매핑을 구성할 때 제안 목록 데이터 사용.</p>
Geocode	<p>주소 레코드에 위도 좌표 및 경도 좌표를 추가하는 데 사용.</p>
보조	<p>주소를 포함한 지리적 영역 또는 인구집단 영역을 식별하는 데 사용.</p> <p>주소 유효성 검사기 변환은 다음 국가의 주소 레코드에 보조 데이터를 추가할 수 있습니다.</p> <ul style="list-style-type: none"> <li>- 일본. 주소가 포함된 시가지 구역을 식별.</li> <li>- 세르비아. 주소의 거리 수준 우편 번호를 식별.</li> <li>- 영국. 주소에 배송 지점 식별자를 추가.</li> <li>- 미국. 주소가 포함된 인구집단 센터 식별.</li> </ul>

PowerCenter 통합 서비스는 AD50.cfg 구성 파일에서 주소 참조 데이터에 대한 경로를 읽습니다.

PowerCenter 설치 프로그램은 다음 경로에 AD50.cfg 파일을 씁니다.

[PowerCenter\_Installation]/server/bin

주소 참조 데이터를 설치하는 경우 다음 속성이 AD50.cfg에 설정되어 있는지 확인하십시오.

#### 참조 데이터 위치

ReferenceDataLocation 속성의 주소 참조 데이터에 대한 경로를 입력하십시오.

#### 라이선스 키 데이터

LicenseKey 속성에 라이선스 데이터를 입력하십시오. 둘 이상의 라이선스를 보유한 경우에는 각 라이선스를 쉼표로 구분된 문자열로 입력하십시오.

#### 사전 로드 값

매핑이 읽는 각 참조 데이터 유형에 대한 사전 로드 값으로 최소 하나의 국가 약어를 입력하십시오. 모든 국가에 사전 로드 설정을 적용하려면 ALL을 입력하십시오.

통합 서비스는 각 국가에 대한 데이터를 로드하는 데 서로 다른 방법을 사용할 수 있습니다. 예를 들어 미국 일괄/대화형 데이터에 대한 전체 사전 로드와 영국 일괄/대화형 데이터에 대한 부분 사전 로드를 지정할 수 있습니다. 통합 서비스도 각 데이터 유형에 대해 다른 사전 로드 방법을 사용할 수 있습니다. 예를 들어 미국 일괄/대화형 데이터에 대한 전체 사전 로드와 미국 좌표 부여 데이터에 대한 부분 사전 로드를 지정할 수 있습니다.

전체 사전 로드 설정은 부분 사전 로드 설정보다 우선하며 부분 사전 로드 설정은 데이터 사전 로드를 나타내지 않는 설정보다 우선합니다. 예를 들어 데이터가 없는 사전 로드에는 대해 ALL을 입력하고 전체 사전 로드에는 대해 USA를 입력하는 경우 통합 서비스는 메모리에 모든 미국 데이터를 로드하며 다른 국가에 대한 데이터를 로드하지 않습니다. 사전 로드 요구 사항이 없는 경우 사용하려고 계획하는 모든 참조 데이터 유형의 데이터가 없는 사전 로드에는 대해 ALL을 입력하십시오.



보조 데이터에 대해서는 사전 로드 값을 지정하지 않습니다.  
선택적으로 AD50.cfg에서 기타 속성을 편집할 수 있습니다.

## ID 인구집단 데이터 요구 사항

ID 일치 분석을 수행하는 매핑은 ID 인구집단이라 불리는 참조 데이터 파일을 읽습니다. Informatica에서 ID 인구집단 파일을 구매하고 다운로드할 수 있습니다. PowerCenter 통합 서비스 시스템 또는 PowerCenter 통합 서비스에서 액세스할 수 있는 시스템의 공유 디렉터리에 인구집단 파일을 복사합니다.

이름이 default인 디렉터리에 인구집단 파일을 복사 또는 설치해야 합니다.

PowerCenter 통합 서비스는 IDQTx.cfg 구성 파일 또는 SSAPR 환경 변수에서 인구집단 파일 디렉터리의 위치를 읽을 수 있습니다. PowerCenter 통합 서비스는 IDQTx.cfg 파일에서 위치를 찾습니다. IDQTx.cfg 파일이 없거나 위치를 지정하지 않은 경우 PowerCenter 통합 서비스가 SSAPR 환경 변수에서 위치를 찾습니다.

**참고:** IDQTx.cfg 파일 및 SSAPR 환경 변수는 default 디렉터리의 상위 디렉터리 경로를 지정합니다. 이 경로에는 디렉터리 이름이 포함되지 않습니다.

PowerCenter 설치 프로그램은 다음 경로에 IDQTx.cfg 파일을 씁니다.

[PowerCenter\_Installation]/server/bin

IDQTx.cfg 파일은 인구집단 파일 디렉터리의 위치를 지정하는 데 *IdentityReferenceDataLocation* 속성을 사용합니다. 속성을 디렉터리의 현지 위치로 업데이트합니다.

예를 들어 *IdentityReferenceDataLocation* 속성에 다음 경로를 설정할 수 있습니다.

[PowerCenter\_Installation]/services/DQContent/INFA\_Content/identity/

*IdentityReferenceDataLocation* 속성에는 기본 디렉터리의 절대 경로 또는 상대 경로를 설정할 수 있습니다. PowerCenter 통합 서비스 시스템의 디렉터리 또는 서비스가 액세스할 수 있는 시스템의 공유 디렉터리에 대한 절대 경로를 설정합니다. PowerCenter 통합 서비스 시스템의 디렉터리에 대한 상대 경로를 설정합니다.

PowerCenter 통합 서비스는 다음 디렉터리를 기준으로 상대 경로를 읽습니다.

[PowerCenter\_Installation]/server/bin

예를 들어, 인구집단 파일 디렉터리가 bin 디렉터리에 있는 경우 상대 경로를 지정하려면 다음 상대 경로를 설정합니다.

./

### IDQTx.cfg 속성

IDQTx.cfg 파일에서 IdentityReferenceDataLocation 속성을 설정하는 경우 다음 속성도 설정할 수 있습니다.

#### IdentityCacheDir

매핑을 실행할 때 작성된 임시 데이터 파일을 저장하는 디렉터리의 경로입니다. 이 경로는 상위 디렉터리를 식별합니다. PowerCenter 통합 서비스는 사용자가 지정하는 위치 아래의 디렉터리에 임시 파일을 씁니다.

이 속성은 다음과 같은 기본값을 가지고 있습니다.

./identityCache

#### IdentityIndexDir

매핑을 실행할 때 작성된 임시 인덱스 파일을 포함하는 디렉터리의 경로입니다. ID 일치 분석은 일치 분석 전에 레코드를 그룹으로 정렬하는 데 인덱스를 사용합니다. 이 경로는 상위 디렉터리를 식별합니다.

PowerCenter 통합 서비스는 사용자가 지정하는 위치 아래의 디렉터리에 인덱스 파일을 씁니다.

이 속성은 다음과 같은 기본값을 가지고 있습니다.

`./identityIndex`

**참고:** Developer tool 사용자는 캐시 디렉터리 경로 및 인덱스 디렉터리 경로를 일치 변환에서 설정할 수 있습니다. PowerCenter 통합 서비스는 일치 변환에서 경로를 지정하지 않는 경우 IDQTx.cfg 파일의 경로를 사용합니다.

## 확률 모델 및 분류자 모델 요구 사항

확률 모델 및 분류자 모델은 데이터 문자열의 정보 유형을 식별하는 참조 데이터 파일입니다.

확률 모델 또는 분류자 모델 파일을 읽는 매핑을 실행하기 전에 다음 태스크를 수행하십시오.

- PowerCenter 디렉터리 구조에서 모델 파일 위치를 복사합니다.
- PowerCenter 통합 서비스의 Java 메모리 할당을 업데이트하십시오.

### PowerCenter의 확률 모델 및 분류자 모델 파일 위치

PowerCenter에서 확률 모델 또는 분류자 모델을 읽는 매핑을 생성하기 전에 모델 파일을 PowerCenter 서비스 시스템에 복사합니다.

모델 리포지토리에서 매핑을 내보내면 내보내기 작업이 각 모델을 지정된 위치의 압축 파일에 추가합니다. 압축 파일을 추출하고, 매핑이 읽는 모델 파일을 찾으십시오. 확률 모델 및 분류자 모델 데이터 파일 이름은 각각 .ner 및 .classifier로 끝납니다. PowerCenter 통합 서비스가 읽을 수 있는 디렉터리에 파일을 복사합니다.

기본적으로 PowerCenter 통합 서비스는 다음 디렉터리에서 모델 파일을 읽습니다.

`[PowerCenter_Installation]/server/bin`

파일이 기본 위치에 없는 경우 PowerCenter 통합 서비스는 속성 파일에서 파일 위치를 읽습니다. 모델 파일에 대한 경로로 속성 파일을 업데이트합니다.

다음 테이블에는 속성 파일이 설명되어 있습니다.

파일 이름	설명
CLASSIFIER.properties	<i>ClassifierFileLocation</i> 속성을 포함합니다. 분류자 모델 파일에 대한 경로를 식별하려면 속성을 업데이트하십시오.
NER.properties	<i>NERFileLocation</i> 속성을 포함합니다. 확률 모델 파일에 대한 경로를 식별하려면 속성을 업데이트하십시오.

PowerCenter 통합 서비스는 다음 디렉터리에서 속성 파일을 읽습니다.

`[PowerCenter_Installation]/server/bin`

#### 확률 모델 및 분류자 모델 내보내기에 대한 규칙과 지침

참조 테이블 데이터를 읽는 매핑을 내보내는 경우 다음 규칙 및 지침을 고려하십시오.

- 확률 모델 또는 분류자 모델 데이터를 읽는 매핑을 내보내 PowerCenter에서 사용하려는 경우 내보내기 마법사에서 **Informatica > 개체 메타데이터 파일 내보내기** 옵션을 선택하여 압축 파일을 생성합니다.  
내보내기 작업에서는 매핑을 XML 파일로도 내보냅니다. 이 XML 파일은 무시해도 됩니다.

- 훈련된 모델만 내보내는 옵션을 선택합니다. 훈련된 확률 모델 또는 분류자 모델에는 매핑이 읽는 논리적 데이터가 포함되며 모델의 소스 데이터는 포함되지 않습니다. 훈련된 모델 파일을 **PowerCenter** 서비스 시스템에 복사합니다.
- 개별 작업에서 모델 리포지토리의 매핑을 **PowerCenter** 리포지토리로 내보냅니다. 모델 리포지토리의 매핑을 **PowerCenter** 리포지토리로 내보내려면 **Informatica > PowerCenter** 옵션을 선택합니다.  
**PowerCenter** 옵션을 선택하면 내보내기 작업에서 관리되는 참조 테이블과 관리되지 않는 참조 테이블의 참조 데이터를 복사할 수 있습니다.
- 각 속성 파일의 모델 파일 위치를 식별하기 위해 절대 경로나 상대 경로를 설정할 수 있습니다. 상대 경로를 설정하는 경우 **PowerCenter** 통합 서비스는 해당 경로를 다음 디렉터리에 추가합니다.  
[PowerCenter\_Installation]/server/bin
- 모든 플랫폼의 모델 파일에 대한 디렉터리 경로를 정의하려면 포워드 슬래시를 사용하십시오.

## PowerCenter 통합 서비스 메모리 할당

**PowerCenter** 통합 서비스는 분류자 모델 파일이나 확률 모델 파일을 읽는 데 대량의 메모리를 사용합니다. **PowerCenter** 통합 서비스가 필요한 메모리에 액세스할 수 없는 경우 모델 파일을 읽을 수 없으며 데이터를 처리할 수 없습니다.

**PowerCenter** 통합 서비스의 메모리 할당을 검토하십시오. 분류자 모델 파일이나 확률 모델 파일을 참조하는 세션을 실행하는 경우에는 메모리 할당을 최소한 **512MB**로 설정하십시오.

메모리를 할당하기 위해 **Java SDK 최대 메모리** 프로세스 속성을 사용할 수 있습니다. 속성을 검토하거나 업데이트하려면 **Administrator** 도구에 로그인하고 해당 세션 작업을 실행하는 **PowerCenter** 통합 서비스를 선택하십시오. **프로세스** 보기에서 **Java SDK 최대 메모리** 속성을 찾습니다.

# 인덱스

## D

Data Quality 통합 [6](#)

## I

Informatica Data Quality [6](#)