



Informatica®  
10.4.0

# 새로운 기능 가이드

Informatica 새로운 기능 가이드

10.4.0

2019년 12월

© 저작권 Informatica LLC 2009, 2019

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.

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation is subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License.

Informatica, the Informatica logo, Big Data Management, PowerCenter, and PowerExchange are trademarks or registered trademarks of Informatica LLC in the United States and many jurisdictions throughout the world. A current list of Informatica trademarks is available on the web at <https://www.informatica.com/trademarks.html>. Other company and product names may be trade names or trademarks of their respective owners.

Portions of this software and/or documentation are subject to copyright held by third parties. Required third party notices are included with the product.

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: 2019-12-19

# 목차

<b>서문</b>	<b>8</b>
Informatica 리소스	8
Informatica 네트워크	8
Informatica 기술 자료	8
Informatica 설명서	8
Informatica Product Availability Matrix	9
Informatica Velocity	9
Informatica Marketplace	9
Informatica 글로벌 고객 지원 센터	9
<b>장 1: 새로운 기능(10.4.0)</b>	<b>10</b>
CI/CD	10
데이터 통합 서비스 REST API	10
infacmd dis 명령	11
역방향 프록시 서버	12
infacmd roh 명령	12
응용 프로그램 패치 배포	13
런타임 응용 프로그램에 연결	13
Object Explorer 보기	14
태그	14
명령줄 프로그램	15
infacmd isp 명령(새로운 기능 10.4.0)	15
Data Engineering Integration	15
새로운 데이터 유형 지원	15
AWS Databricks 통합	16
ALDS Gen2 리소스에 대한 HDInsight 액세스를 위한 클러스터 워크플로우	16
Databricks Delta Lake 저장소 액세스	16
매핑에 사용되는 노드 표시	17
로그 집계	17
Spark 엔진에서 계층적 데이터 구문 분석	17
Spark 엔진의 프로필 및 샘플링 옵션	17
Python 변환	18
Sqoop	18
Data Engineering Streaming	18
스트리밍 매핑의 Confluent 스키마 레지스트리	19
스트리밍 매핑의 데이터 품질 변환	19
스트리밍 매핑의 임시 클러스터	19
Amazon S3의 파일 이름 포트	19
Microsoft Azure Data Lake Storage Gen2	19

Azure Databricks의 스트리밍 매핑. . . . .	20
Data Engineering Streaming의 동적 매핑. . . . .	21
Enterprise Data Catalog. . . . .	21
리소스 및 클래스에 사용자 지정 특성 할당. . . . .	21
새 리소스. . . . .	21
참조 리소스 및 참조 자산. . . . .	21
검색 결과에서 자산 내보내기. . . . .	22
연계 및 영향 필터. . . . .	22
자산 제어 요약. . . . .	22
규칙 및 성과 기록표. . . . .	22
고유 키 유추. . . . .	22
CLOB 파일 유형의 데이터 도메인 검색. . . . .	23
Spark 엔진의 데이터 검색 및 샘플링 옵션. . . . .	23
추적 기술 미리보기. . . . .	23
데이터 미리보기 및 프로비저닝. . . . .	24
독립 실행형 스캐너 유틸리티의 지원되는 리소스 유형. . . . .	24
REST API. . . . .	25
Enterprise Data Preparation. . . . .	25
Data Lake 액세스 관리. . . . .	25
데이터 소스로 Microsoft Azure Data Lake Storage 사용. . . . .	25
Data Lake에 파일 게시. . . . .	25
Data Lake에 파일 업로드. . . . .	25
Informatica 매핑. . . . .	26
매핑 매개 변수에 매핑 출력 바인딩. . . . .	26
CLAIRE 권장 사항 및 인사이트. . . . .	27
매핑 최적화 프로그램 수준 업데이트. . . . .	27
Informatica 변환. . . . .	27
주소 유효성 검사기 변환. . . . .	27
Informatica 워크플로우. . . . .	28
Amazon EMR 클러스터 생성 태스크 고급 속성. . . . .	28
Informatica 설치. . . . .	28
PostgreSQL. . . . .	28
자동 모드의 설치 전(i10Pi) 시스템 검사 도구. . . . .	28
자동 설치 속성 파일의 암호 암호화. . . . .	29
지능형 구조 모델. . . . .	29
추가 입력 유형. . . . .	29
디자인 타임에 샘플에서 모델 생성. . . . .	29
식별되지 않은 데이터 처리. . . . .	29
PowerCenter. . . . .	29
연결. . . . .	30
PowerExchange 어댑터. . . . .	30
PowerExchange Adapters for Informatica. . . . .	30

PowerExchange Adapters for PowerCenter. . . . .	33
보안. . . . .	35
서로 다른 SAML ID 공급자를 사용하도록 웹 응용 프로그램 구성. . . . .	35
<b>장 2: 새로운 기능(10.2.2 HotFix 1).....</b>	<b>36</b>
명령줄 프로그램 . . . . .	36
infacmd ldm 명령. . . . .	36
Enterprise Data Catalog. . . . .	36
Azure Data Lake Storage Gen2. . . . .	36
대/소문자 구분 없는 연결 . . . . .	37
Enterprise Data Catalog Tableau 확장 . . . . .	37
새 리소스. . . . .	37
프로파일링 웨어하우스 스캐너. . . . .	37
REST API. . . . .	37
검색 개선 사항 . . . . .	38
검색 탭. . . . .	38
독립 실행형 스캐너 유틸리티의 지원되는 리소스 유형. . . . .	38
기술 미리 보기. . . . .	39
<b>장 3: 새로운 기능(10.2.2 서비스 팩 1).....</b>	<b>40</b>
Big Data Management. . . . .	40
Sqoop. . . . .	40
Big Data 스트리밍. . . . .	41
Amazon S3 대상. . . . .	41
TIME_RANGE 함수. . . . .	41
여러 HDFS 파일에 데이터 기록 가능. . . . .	41
Enterprise Data Catalog. . . . .	41
독립 실행형 스캐너 유틸리티의 지원되는 리소스 유형. . . . .	41
Enterprise Data Preparation. . . . .	42
모든 유추된 데이터 유형 되돌리기. . . . .	42
Avro 및 Parquet 파일 준비. . . . .	42
PowerExchange Adapters for Informatica. . . . .	42
PowerExchange for Hive. . . . .	42
<b>장 4: 새로운 기능(10.2.2).....</b>	<b>43</b>
응용 프로그램 서비스. . . . .	43
대량 수집 서비스. . . . .	43
메타데이터 액세스 서비스. . . . .	43
REST 작업 힙 서비스. . . . .	44
Big Data Management. . . . .	44
Azure Databricks 통합. . . . .	44
계층 데이터. . . . .	45

고가용성. . . . .	46
지능형 구조 모델. . . . .	46
대량 수집. . . . .	47
모니터링. . . . .	47
보안. . . . .	48
대상. . . . .	48
<b>Big Data 스트리밍 . . . . .</b>	<b>48</b>
Azure 이벤트 헵 데이터 개체. . . . .	48
Amazon Kinesis 연결의 교차 계정 IAM 역할. . . . .	48
지능형 구조 모델. . . . .	49
Big Data Streaming 데이터 개체에 대한 헤더 포트. . . . .	49
Amazon Kinesis 연결의 AWS 자격 증명 프로필. . . . .	49
Spark 구조화된 스트리밍. . . . .	49
창 변환. . . . .	49
명령줄 프로그램. . . . .	50
infacmd dis 명령. . . . .	50
infacmd ihs 명령. . . . .	51
infacmd ipc 명령. . . . .	51
infacmd ldm 명령. . . . .	51
infacmd mi 명령. . . . .	52
infacmd ms 명령. . . . .	53
infacmd oie 명령. . . . .	53
infacmd tools 명령. . . . .	53
infasetup 명령. . . . .	54
<b>Enterprise Data Catalog. . . . .</b>	<b>54</b>
열에 비즈니스 제목 자동 할당. . . . .	54
자산에 대한 사용자 공동 작업. . . . .	54
설치 프로그램을 사용하여 Enterprise Data Catalog 응용 프로그램 서비스 생성. . . . .	55
사용자 지정 메타데이터 유효성 검사 유틸리티. . . . .	55
변경 알림. . . . .	55
Business Glossary 할당 보고서. . . . .	55
운영 체제 프로필. . . . .	56
REST API. . . . .	56
소스 메타데이터 및 데이터 프로필 필터. . . . .	56
스캐너 유틸리티. . . . .	56
리소스 유형. . . . .	56
<b>Enterprise Data Lake. . . . .</b>	<b>57</b>
활성 규칙 적용. . . . .	57
중복 행 삭제. . . . .	57
열 데이터 클러스터 및 범주화. . . . .	57
CLAIRE 기반 권장 사항. . . . .	58
조건부 집계. . . . .	58

데이터 마스킹.....	58
지역화.....	58
분할된 소스 및 대상.....	58
레시피 단계에 설명 추가.....	59
레시피를 매핑으로 저장.....	59
Amazon S3, ADLS, WASB, MapR-FS를 데이터 소스로.....	59
통계 함수.....	59
날짜 및 시간 함수.....	60
수학 함수.....	60
텍스트 함수.....	60
창 함수.....	61
감사 이벤트 제거.....	61
Spark 실행 엔진.....	61
Informatica Developer.....	62
응용 프로그램.....	62
Informatica 매핑.....	62
데이터 유형.....	62
매핑 출력.....	62
매핑 매개 변수.....	62
최적화 수준.....	62
Sqoop.....	63
Informatica 변환.....	63
주소 유효성 검사기 변환.....	63
업데이트 전략 변환.....	66
PowerExchange Adapters for Informatica.....	66
PowerExchange for Amazon Redshift.....	67
PowerExchange for Amazon S3.....	67
PowerExchange for Google BigQuery.....	68
PowerExchange for HBase.....	68
PowerExchange for HDFS.....	68
PowerExchange for Hive.....	69
PowerExchange for MapR-DB.....	70
PowerExchange for Microsoft Azure Blob Storage.....	70
PowerExchange for Microsoft Azure Cosmos DB SQL API.....	70
PowerExchange for Microsoft Azure Data Lake Store.....	70
PowerExchange for Microsoft Azure SQL Data Warehouse.....	71
PowerExchange for Salesforce.....	71
PowerExchange for Snowflake.....	72
PowerExchange for Teradata Parallel Transporter API.....	72

# 서문

이 릴리스의 Informatica 제품에 포함된 새로운 기능과 향상된 기능에 대해 알아보려면 *Informatica® 새로운 기능 및 향상된 기능 가이드*를 참조하십시오.

## 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

## 새로운 기능(10.4.0)

이 장에 포함된 항목:

- [CI/CD, 10](#)
- [명령줄 프로그램, 15](#)
- [Data Engineering Integration, 15](#)
- [Data Engineering Streaming , 18](#)
- [Enterprise Data Catalog, 21](#)
- [Enterprise Data Preparation, 25](#)
- [Informatica 매핑 , 26](#)
- [Informatica 변환, 27](#)
- [Informatica 워크플로우, 28](#)
- [Informatica 설치, 28](#)
- [지능형 구조 모델, 29](#)
- [PowerCenter, 29](#)
- [PowerExchange 어댑터, 30](#)
- [보안, 35](#)

## CI/CD

이 섹션에서는 버전 10.4.0에서 향상된 CI/CD 기능에 대해 설명합니다.

CI/CD(연속 통합/연속 배포)는 CI/CD 파이프라인에서 통합 및 배포 작업을 자동화하는 방법입니다. 버전 10.4.0에서는 이러한 향상된 기능을 CI/CD 파이프라인에 통합하여 개체를 배포하고 테스트한 후 프로덕션 환경에 배포하는 방법을 개선할 수 있습니다.

## 데이터 통합 서비스 REST API

버전 10.4.0에서는 데이터 통합 서비스 REST API를 사용하여 CI/CD 파이프라인의 태스크를 자동화할 수 있습니다.

REST API가 자동화할 수 있는 일부 태스크에는 다음과 같은 태스크가 포함됩니다.

**개체 쿼리.**

모델 리포지토리의 디자인 타임 개체와 데이터 통합 서비스에 배포된 런타임 개체를 포함한 개체를 쿼리합니다.

이 쿼리를 다른 REST API 요청에 전달할 수 있습니다. 예를 들어 버전 제어 작업에 쿼리를 전달하여 특정 개체 집합에 대한 버전 제어를 수행할 수 있습니다. 또한 쿼리를 전달하여 특정 디자인 타임 개체를 응용 프로그램 패치 보관 파일에 배포할 수 있습니다.

#### 버전 제어 작업 수행.

체크 인, 체크 아웃, 체크 아웃 실행 취소 또는 체크 아웃된 디자인 타임 개체를 다른 개발자에게 재할당하는 버전 제어 작업을 수행합니다.

#### 태그 관리.

디자인 타임 개체에 할당된 태그를 관리합니다. 새 태그를 할당하거나 개체의 태그를 교체할 수 있습니다. 또한 개체에 할당된 태그를 취소할 수도 있습니다.

#### 응용 프로그램 업데이트.

디자인 타임 개체를 응용 프로그램 패치 보관 파일에 배포하고 파일을 데이터 통합 서비스에 배포하여 배포된 증분 응용 프로그램을 업데이트합니다.

#### 응용 프로그램 관리.

프로젝트 또는 조직 요구 사항에 기반하여 응용 프로그램을 시작하거나, 배포 해제하거나, 중지합니다.

#### 배포된 매핑 실행.

배포된 매핑을 실행하여 응용 프로그램 출력을 테스트합니다.

#### 매핑 비교.

동일한 도메인에서 두 매핑을 비교합니다.

예를 들어 두 디자인 타임 매핑을 비교하거나 디자인 타임 매핑과 런타임 매핑을 비교할 수 있습니다.

사용 가능한 REST API 요청과 각 요청의 매개 변수를 보려면 **Administrator** 도구에서 데이터 통합 서비스 프로세스 속성이나 REST 작업 협 서비스 속성을 통해 REST API 설명서에 액세스하십시오.

**infacmd** 명령줄 프로그램과 달리, REST API에는 설치 요구 사항이 없으며 클라이언트 시스템에 **Informatica** 서비스가 설치되어 있지 않은 환경에서도 REST API를 실행할 수 있습니다.

REST API에 대한 자세한 내용은 *Informatica 10.4.0 응용 프로그램 서비스 가이드*에서 "데이터 통합 서비스 REST API" 장을 참조하십시오.

## infacmd dis 명령

버전 10.4.0에서는 **infacmd dis** 명령을 사용하여 디자인 타임 개체, 응용 프로그램의 런타임 개체 및 배포된 매핑에 대한 작업을 수행할 수 있습니다. 또한 **infacmd dis** 명령을 사용하여 데이터 통합 서비스 유틸리티를 실행할 수 있습니다.

다음 테이블에는 새로운 **infacmd dis** 명령이 설명되어 있습니다.

명령	설명
compareMapping	쿼리된 두 매핑을 비교합니다. 매핑을 쿼리하여 매핑 속성, 변환 속성 및 변환 내 포트를 비교하십시오. 디자인 타임 매핑을 쿼리하려면 디자인 타임 모델 리포지토리를 지정하십시오. 런타임 매핑을 쿼리하려면 모델 리포지토리를 지정하지 마십시오. 이 쿼리에서는 명령을 실행하도록 지정한 데이터 통합 서비스를 사용합니다.
deployObjectsToFile	디자인 타임 개체를 응용 프로그램 패치 보관 파일에 배포합니다.
queryDesignTimeObjects	모델 리포지토리 서비스에서 디자인 타임 개체를 쿼리합니다.

명령	설명
queryRunTimeObjects	데이터 통합 서비스에 배포된 런타임 개체를 쿼리하고 개체 목록을 반환합니다.
replaceAllTag	모델 리포지토리 서비스의 쿼리된 개체에서 한 태그를 지정된 태그로 바꿉니다.
tag	모델 리포지토리 서비스의 쿼리된 개체에 태그를 할당합니다.
untag	모델 리포지토리 서비스의 쿼리된 개체에서 태그를 제거합니다.
listPatchNames	중분 응용 프로그램에 적용된 모든 패치를 나열합니다.

자세한 내용은 *Informatica 10.4.0 명령 참조*에서 "infacmd dis 명령 참조" 장을 참조하십시오.

## 역방향 프록시 서버

버전 10.4.0에서는 REST 작업 협을 사용하여 데이터 통합 서비스 그리드의 노드 간에서 데이터 통합 서비스 API 요청의 부하 분산 처리를 수행하는 역방향 프록시 서버를 시작할 수 있습니다.

REST 작업 협에서는 그리드가 활성화됩니다.

역방향 프록시 서버에 대한 자세한 내용은 *Informatica 10.4.0 응용 프로그램 서비스 가이드*에서 "시스템 서비스" 장을 참조하십시오.

## infacmd roh 명령

버전 10.4.0에서는 다음과 같은 infacmd roh 명령을 사용하여 REST 작업 협 서비스 프로세스를 업데이트할 수 있습니다.

다음 테이블에는 새로운 infacmd roh 명령이 설명되어 있습니다.

명령	설명
listReverseProxyServerOptions	역방향 프록시 서버 속성을 나열합니다.
listServiceProcessOptions	REST 작업 협 서비스 프로세스 속성을 나열합니다.
updateServiceOptions	REST 작업 협 서비스 속성을 업데이트합니다.
updateReverseProxyServerOptions	도메인의 역방향 프록시 서버 속성을 업데이트합니다.

버전 10.4.0에서는 다음과 같은 infacmd roh 명령의 이름이 바뀌었습니다.

- listROHProperties가 listProcessProperties로 바뀌었습니다.
- updateROHService가 updateServiceProcessOptions로 바뀌었습니다.

**참고:** 이전 명령 이름을 사용하는 모든 스크립트를 업데이트하십시오.

자세한 내용은 *Informatica 10.4.0 명령 참조*에서 "infacmd roh 명령 참조" 장을 참조하십시오.

## 응용 프로그램 패치 배포

버전 10.4.0에서는 응용 프로그램 패치 배포의 다음과 같은 영역에 새로운 기능이 추가되었습니다.

### 증분 배포 마법사

버전 10.4.0에서는 증분 배포 마법사에서 다음 태스크를 수행할 수 있습니다.

- 필요에 따라 패치 설명을 입력할 수 있습니다.
- 필요에 따라 상태 정보를 유지하거나 무시하도록 선택할 수 있습니다.  
상태 정보란 매핑 출력이나 시퀀스 생성기 변환과 같은 런타임 개체의 속성 및 매핑 속성을 가리킵니다.

증분 배포 마법사에 대한 자세한 내용은 *Informatica 10.4.0 Developer Tool 가이드*에서 "응용 프로그램 패치 배포" 장을 참조하십시오.

상태 정보에 대한 자세한 내용은 *Informatica 10.4.0 Developer Tool 가이드*에서 "응용 프로그램 배포" 장을 참조하십시오.

### 패치 기록

버전 10.4.0에서는 증분 배포 마법사의 패치 기록에 증분 응용 프로그램을 업데이트하기 위해 배포된 패치의 패치 이름 및 패치 설명이 모두 표시됩니다. 패치가 생성된 시간은 패치 설명 시작 부분에 추가되어 있습니다.

또한 Administrator 도구를 사용하여 배포된 증분 응용 프로그램의 패치 기록을 볼 수 있습니다.

패치 기록에 대한 자세한 내용은 *Informatica 10.4.0 Developer Tool 가이드*에서 "응용 프로그램 패치 배포" 장을 참조하십시오.

배포된 응용 프로그램에 대한 자세한 내용은 *Informatica 10.4.0 응용 프로그램 서비스 가이드*에서 "데이터 통합 서비스 응용 프로그램" 장을 참조하십시오.

### 응용 프로그램 패치 보관 파일

버전 10.4.0에서는 응용 프로그램 패치 보관 파일에 추가적으로 패치 설명이 저장되고 상태 정보가 유지되는지 여부가 나타납니다.

패치 기록에 대한 자세한 내용은 *Informatica 10.4.0 Developer Tool 가이드*에서 "응용 프로그램 패치 배포" 장을 참조하십시오.

### infacmd tools 명령

버전 10.4.0에서는 infacmd tools patchApplication 명령에 다음과 같은 새로운 옵션이 추가되었습니다.

옵션	인수	설명
-RetainStateInformation -rsi	True  False	선택 사항입니다. 상태 정보가 유지되는지, 아니면 삭제되는지를 나타냅니다. 상태 정보란 매핑 출력이나 시퀀스 생성기 변환과 같은 런타임 개체의 속성 및 매핑 속성을 가리킵니다.

자세한 내용은 *Informatica 10.4.0 명령 참조*에서 "infacmd tools 명령 참조" 장을 참조하십시오.

## 런타임 응용 프로그램에 연결

버전 10.4.0에서는 Developer tool에서 런타임 응용 프로그램에 연결할 수 있습니다. 런타임 응용 프로그램에 연결한 후 **Object Explorer** 보기에서 응용 프로그램을 확장하고 **편집기**에서 런타임 개체의 읽기 전용 복사본을 열 수 있습니다.

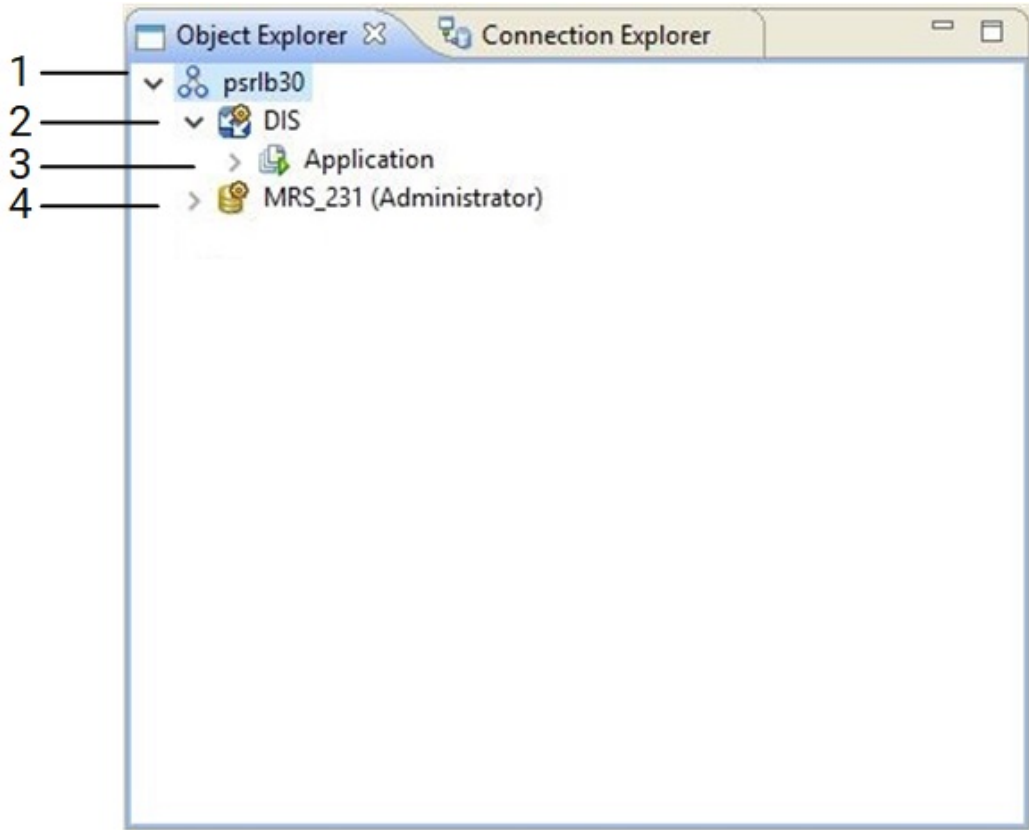
런타임 응용 프로그램에 연결한 후에는 Developer tool에서 수행하는 검색이 응용 프로그램의 런타임 개체를 찾을 수 있습니다.

런타임 응용 프로그램 연결 및 런타임 개체 보기에 대한 자세한 내용은 *Informatica 10.4.0 Developer Tool 가이드*에서 "응용 프로그램 배포" 장을 참조하십시오.

## Object Explorer 보기

버전 10.4.0에서는 모델 리포지토리 또는 런타임 응용 프로그램에 연결한 후 Developer tool의 **Object Explorer** 보기에 도메인이 표시됩니다. 도메인을 확장하여 모델 리포지토리의 디자인 타임 개체 또는 런타임 응용 프로그램의 런타임 개체를 확인할 수 있습니다.

다음 이미지는 **Object Explorer** 보기를 보여 줍니다.



1. 도메인
2. 데이터 통합 서비스
3. 런타임 응용 프로그램
4. 모델 리포지토리

Developer tool의 사용자 인터페이스에 대한 자세한 내용은 *Informatica 10.4.0 Developer Tool 가이드*에서 "Informatica Developer" 장을 참조하십시오.

## 태그

버전 10.4.0에서는 태그에 다음과 같은 기능이 있습니다.

- 태그와 연결된 매핑을 배포할 경우 데이터 통합 서비스에서 매핑의 런타임 버전에 태그가 전파됩니다.
- 응용 프로그램 패치를 사용하여 배포된 매핑을 업데이트하면 패치의 이름이 매핑의 런타임 버전에 태그로 연결됩니다.

태그에 대한 자세한 내용은 *Informatica 10.4.0 Developer tool 가이드*에서 "Informatica Developer" 장을 참조하십시오.

## 명령줄 프로그램

이 섹션에서는 버전 10.4.0의 새로운 명령에 대해 설명합니다.

### infacmd isp 명령(새로운 기능 10.4.0)

다음 테이블에는 새로운 infacmd isp 명령이 설명되어 있습니다.

명령	설명
addCustomLDAPType	LDAP 디렉터리 서비스를 정의하는 사용자 지정 LDAP 유형을 추가합니다.
listAllCustomLDAPTypes	지정된 도메인에서 사용되는 모든 사용자 지정 LDAP 유형에 대한 구성 정보를 나열합니다.
listAllLDAPConnectivity	지정된 도메인에서 사용되는 모든 LDAP 구성에 대한 구성 정보를 나열합니다.
listCustomLDAPType	사용자 지정 LDAP 유형에 대한 구성 정보를 나열합니다.
removeCustomLDAPType	지정된 도메인에서 지정된 사용자 지정 LDAP 유형을 제거합니다.
removeLDAPConnectivity	지정된 도메인에서 지정된 LDAP 구성을 제거합니다.
updateCustomLDAPType	지정된 사용자 지정 LDAP 유형을 업데이트합니다.
updateLDAPConnectivity	지정된 LDAP 구성을 업데이트합니다.

자세한 내용은 *Informatica 10.4.0 명령 참조*를 참조하십시오.

## Data Engineering Integration

이 섹션에서는 버전 10.4.0의 새로운 Data Engineering Integration 기능에 대해 설명합니다.

### 새로운 데이터 유형 지원

버전 10.4.0에서는 복합 파일에 대한 다음과 같은 새로운 데이터 유형을 사용할 수 있습니다.

- 원시 환경 또는 Hadoop 환경에서 Avro 및 Parquet 복합 파일 개체를 읽거나 쓰는 매핑을 실행하는 경우 다음과 같은 데이터 유형을 사용할 수 있습니다.
  - 날짜
  - 10진수
  - 타임스탬프

- 시간 데이터 유형을 사용하여 원시 환경 또는 **Blaze** 엔진에서 **Avro** 또는 **Parquet** 복합 파일 개체를 읽고 쓸 수 있습니다.
- **Databricks Spark** 엔진에서 매핑을 실행할 때 날짜, 시간, 타임스탬프 및 10진수 데이터 유형을 사용할 수 있습니다.

이러한 새 데이터 유형은 다음과 같은 어댑터에 적용할 수 있습니다.

- PowerExchange for HDFS
- PowerExchange for Amazon S3
- PowerExchange for Google Cloud Storage
- PowerExchange for Microsoft Azure Blob Storage
- PowerExchange for Microsoft Azure Data Lake Storage Gen1
- PowerExchange for Microsoft Azure Data Lake Storage Gen2

데이터 유형에 대한 자세한 내용은 *Data Engineering Integration 10.4.0 사용자 가이드*에서 "데이터 유형 참조" 장을 참조하십시오.

## AWS Databricks 통합

버전 10.4.0에서는 Informatica 도메인과 AWS의 Databricks를 통합할 수 있습니다.

AWS Databricks를 사용하여 다음과 같은 기능을 포함하는 매핑을 실행할 수 있습니다.

- Databricks 환경 내에서 Amazon S3(Simple Storage Service) 및 Amazon Redshift의 소스 및 대상에 대해 매핑을 실행할 수 있습니다.
- 클러스터 워크플로우를 개발하여 AWS의 Databricks를 사용하는 임시 클러스터를 생성할 수 있습니다.
- Databricks Spark 엔진에서 실행되도록 구성된 매핑에 Python 변환을 추가할 수 있습니다. Python 변환은 기술 미리보기에서만 지원됩니다.

AWS Databricks는 Azure Databricks와 동일한 데이터 유형을 지원합니다.

자세한 내용은 다음 가이드를 참조하십시오.

*Data Engineering 10.4.0 통합 가이드*

*Data Engineering 10.4.0 관리자 가이드*

*Data Engineering Integration 10.4.0 사용자 가이드*

*Informatica 10.4.0 Developer 워크플로우 가이드*

## ALDS Gen2 리소스에 대한 HDInsight 액세스를 위한 클러스터 워크플로우

버전 10.4.0에서는 Azure HDInsight 클러스터에서 실행되어 ADLS Gen2 리소스에 액세스하는 클러스터 워크플로우를 생성할 수 있습니다.

클러스터 워크플로우에 대한 자세한 내용은 *Informatica Data Engineering Integration 10.4.0 사용자 가이드*를 참조하십시오.

## Databricks Delta Lake 저장소 액세스

버전 10.4.0에서는 Databricks Delta Lake 저장소를 소스 및 대상으로 액세스할 수 있습니다.

매핑에서 AWS 및 Azure 플랫폼의 Delta Lake 리소스에 액세스할 수 있습니다.



Delta Lake 테이블의 액세스 구성에 대한 자세한 내용은 *Data Engineering Integration Guide*에서 참조하십시오. Delta Lake 테이블에 액세스하는 매핑 생성에 대한 자세한 내용은 *Data Engineering Integration User Guide*에서 참조하십시오.

## 매핑에 사용되는 노드 표시

버전 10.4.0에서는 지정된 기간 동안 매핑에 사용되는 클러스터 노드의 최대 수를 확인할 수 있습니다.

REST 작업 헵 API `ClusterStats(startTimeInmillis=[value], endTimeInmillis=[value])`를 사용하여 지정된 기간 동안 매핑에 사용되는 클러스터 구성의 Hadoop 노드 최대 수를 확인할 수 있습니다.

REST API에 대한 자세한 내용은 *Data Engineering 10.4.0 관리자 가이드*에서 "REST API 모니터링 참조" 장을 참조하십시오.

## 로그 집계

버전 10.4.0에서는 Hadoop 환경에서 실행되는 배포된 매핑에 대한 집계된 로그를 가져올 수 있습니다.

모니터링 도구의 작업 ID를 기반으로 또는 `infacmd ms fetchAggregatedClusterLogs` 명령을 사용하여 매핑에 대한 집계된 클러스터 로그를 수집할 수 있습니다. 작업 ID를 기반으로 매핑에 대한 집계된 클러스터 로그의 .zip 또는 tar.gz 파일을 가져오고 집계된 로그 파일을 압축하여 대상 디렉터리에 쓸 수 있습니다.

자세한 내용은 *Informatica 10.4.0 관리자 가이드*를 참조하십시오.

## Spark 엔진에서 계층적 데이터 구문 분석

버전 10.4.0에서는 복합 함수를 사용하여 매핑에서 최대 5MB의 데이터 미드스트림을 구문 분석할 수 있습니다.

Spark 엔진은 다음과 같은 복합 함수를 사용하여 원시 문자열 소스 데이터를 구문 분석할 수 있습니다.

- PARSE\_JSON
- PARSE\_XML

이러한 복합 함수는 소스 문자열의 JSON 또는 XML 데이터를 구문 분석하고 구조 대상 데이터를 생성합니다.

자세한 내용은 *Informatica Data Engineering Integration 10.4.0 사용자 가이드*에서 "계층적 데이터 처리" 장을 참조하십시오.

복합 함수에 대한 자세한 내용은 *Informatica 10.4.0 Developer 변환 언어 참조*에서 "함수" 장을 참조하십시오.

## Spark 엔진의 프로필 및 샘플링 옵션

버전 10.4.0에서는 Spark에서 프로필을 실행하고 샘플링 옵션을 선택할 수 있습니다.

### Spark 엔진에서 프로필 실행

Informatica Developer 및 Informatica Analyst 도구에서 Spark 엔진에 대한 프로필을 생성하고 실행할 수 있습니다. Spark 엔진에서 데이터 도메인 검색을 수행하고 성과 기록표를 생성할 수 있습니다.

### Spark 엔진의 샘플링 옵션

다음과 같은 샘플링 옵션을 선택하여 Spark 엔진에서 프로필을 실행할 수 있습니다.

- **n개 제한** 샘플링 옵션은 데이터 개체의 행 수를 기반으로 프로필을 실행합니다. Hadoop 환경에서 프로필을 실행하도록 선택하면 Spark 엔진이 데이터 개체의 여러 파티션에서 샘플을 수집하고 샘플 크기를 계산하도록 단일 노드에 샘플을 푸시합니다. 고급 필터를 사용하는 프로필에는 n개 제한 샘플링 옵션을 적용할 수 없습니다.

Oracle, SQL Server 및 DB2 데이터베이스에서 지원됩니다.

- **무작위 백분율** 샘플링 옵션은 데이터 개체의 행에서 백분율만큼 프로필을 실행합니다.

Spark 엔진의 프로필 및 샘플링 옵션에 대한 자세한 내용은 *Informatica 10.4.0 Data Discovery 가이드*를 참조하십시오.

## Python 변환

버전 10.4.0에서는 Python 변환에 다음과 같은 기능이 있습니다.

### 활성 모드

활성 Python 변환을 생성할 수 있습니다. 활성 변환인 Python 변환에서는 변환을 통과하는 행 수를 변경할 수 있습니다. 예를 들어 Python 변환이 단일 입력 행에서 여러 출력 행을 생성하거나 여러 입력 행에서 단일 출력 행을 생성할 수 있습니다.

자세한 내용은 *Informatica Data Engineering Integration 10.4.0 사용자 가이드*에서 "Python 변환" 장을 참조하십시오.

### 분할된 데이터

Python 코드를 실행하여 데이터의 기본 분할 스키마를 기반으로 들어오는 데이터를 처리하거나 Python 코드를 실행하기 전에 데이터를 재분할할 수 있습니다. Python 코드를 실행하기 전에 데이터를 재분할하려면 하나 이상의 입력 포트를 파티션 키로 선택합니다.

자세한 내용은 *Informatica Data Engineering Integration 10.4.0 사용자 가이드*에서 "Python 변환" 장을 참조하십시오.

## Sqoop

버전 10.4.0에서는 JDBC 연결에서 다음과 같은 Sqoop 인수를 구성할 수 있습니다.

- --update-key
- --update-mode
- --validate
- --validation-failurehandler
- --validation-threshold
- --validator
- --mapreduce-job-name
- --bindir
- --class-name
- --jar-file
- --outdir
- --package-name

이러한 Sqoop 인수 구성에 대한 자세한 내용은 Sqoop 설명서를 참조하십시오.

# Data Engineering Streaming

이 섹션에서는 버전 10.4.0의 새로운 Data Engineering Streaming 기능에 대해 설명합니다.

## 스트리밍 매핑의 Confluent 스키마 레지스트리

버전 10.4.0에서는 스키마 레지스트리를 사용하는 스트리밍 매핑에서 **Confluent Kafka**를 소스 및 대상으로 사용할 수 있습니다.

**Confluent Kafka**를 사용하여 스트리밍 매핑에서 **Apache Avro** 스키마를 저장하고 검색할 수 있습니다. 스키마 레지스트리는 **Kafka**를 기본 저장 메커니즘으로 사용합니다.

자세한 내용은 *Data Engineering Streaming 10.4.0 사용자 가이드*를 참조하십시오.

## 스트리밍 매핑의 데이터 품질 변환

버전 10.4.0에서는 스트리밍 매핑에서 데이터 품질 변환을 사용할 수 있습니다.

스트리밍 매핑에서 다음과 같은 데이터 품질 변환을 사용하여 스트리밍 데이터에 데이터 품질 프로세스를 적용할 수 있습니다.

- 주소 유효성 검사기 변환
- 분류자 변환
- 파서 변환
- 표준화 변환

자세한 내용은 *Data Engineering Streaming 10.4.0 사용자 가이드*를 참조하십시오.

## 스트리밍 매핑의 임시 클러스터

버전 10.4.0에서는 클라우드 플랫폼 클러스터에서 매핑 및 기타 태스크를 실행하는 임시 클러스터를 생성하는 워크플로우를 실행할 수 있습니다.

클러스터가 삭제된 지점부터 데이터 프로세스를 다시 시작하려면 외부 저장소 및 검사점 디렉터리를 지정하여 임시 클러스터에서 스트리밍 매핑을 실행할 수 있습니다.

자세한 내용은 *Data Engineering Streaming 10.4.0 사용자 가이드*를 참조하십시오.

## Amazon S3의 파일 이름 포트

버전 10.4.0에서는 Amazon S3 파일에 대한 데이터 개체 쓰기 작업을 생성할 때 기본적으로 파일 이름 포트가 나타납니다.

런타임 시 데이터 통합 서비스에서는 파일 이름 포트의 각 값에 대한 별도의 디렉터리를 생성하고 대상 파일을 해당 디렉터리에 추가합니다.

자세한 내용은 *Data Engineering Streaming 10.4.0 사용자 가이드*를 참조하십시오.

## Microsoft Azure Data Lake Storage Gen2

버전 10.4.0에서는 Microsoft Azure Data Lake Storage Gen2를 스트리밍 매핑의 대상으로 사용할 수 있습니다.

Azure Data Lake Storage Gen2는 Azure Blob Storage 기반으로 구축됩니다. Azure Data Lake Storage Gen2에는 Azure Data Lake Storage Gen1의 기능과 Azure Blob Storage 기능이 모두 있습니다. Azure Databricks 버전 5.4 또는 Azure HDInsight 버전 4.0을 사용하여 Azure Data Lake Storage Gen2에 저장된 데이터에 액세스할 수 있습니다.

자세한 내용은 *Data Engineering Streaming 10.4.0 사용자 가이드*를 참조하십시오.

## Azure Databricks의 스트리밍 매핑

버전 10.4.0에서는 Microsoft Azure 클라우드 서비스의 Azure Databricks 서비스에서 스트리밍 매핑을 실행할 수 있습니다.

### 소스 및 대상

Databricks 환경에 포함된 다음과 같은 소스 및 대상에 대해 스트리밍 매핑을 실행할 수 있습니다.

Microsoft Azure Event Hubs

ADLS Gen2(Azure Data Lake Storage Gen2)

### 변환

다음 변환을 Databricks 스트리밍 매핑에 추가할 수 있습니다.

집계

식

필터

조이너

노멀라이저

순위

라우터

합집합

창

### 데이터 유형

다음과 같은 데이터 유형이 지원됩니다.

배열

Bigint

날짜/시간

10진수

배정밀도

정수

맵

구조

텍스트

문자열

### 워크플로우

클러스터 워크플로우를 개발하여 Databricks 환경에서 임시 클러스터를 생성할 수 있습니다. ADLS Gen1(Azure Data Lake Storage Gen1) 및 ADLS Gen2(Azure Data Lake Storage Gen2)를 사용하여 Databricks 환경에 임시 클러스터를 생성합니다.

Azure Databricks의 스트리밍 매핑에 대한 자세한 내용은 *Data Engineering Streaming 10.4.0 사용자 가이드*를 참조하십시오.

## Data Engineering Streaming의 동적 매핑

버전 10.4.0에서는 기술 미리보기를 통해 Data Engineering Streaming의 동적 매핑 지원을 사용할 수 있습니다.

Confluent Kafka 데이터 개체를 스트리밍 매핑에서 동적 소스 및 대상으로 사용할 수 있습니다.

기술 미리 보기 기능은 평가 목적으로 지원되지만 기능이 보장되지 않으며 프로덕션용으로 준비되지 않았습니다. 비프로덕션 환경에서만 사용할 것을 권장합니다. Informatica는 향후 프로덕션용 릴리스에 미리 보기 기능을 포함할 계획이지만 시장 또는 기술적 상황의 변화에 따라 여의치 않을 경우 이러한 기능을 포함하지 않을 수 있습니다. 자세한 내용은 Informatica 글로벌 고객 지원 센터에 문의하십시오.

## Enterprise Data Catalog

이 섹션에서는 버전 10.4.0의 새로운 Enterprise Data Catalog 기능에 대해 설명합니다.

### 리소스 및 클래스에 사용자 지정 특성 할당

버전 10.4.0에서는 카탈로그의 특정 리소스 및 클래스에 또는 카탈로그의 모든 리소스 및 클래스에 사용자 지정 특성을 할당할 수 있습니다. 사용자 지정 특성을 생성하거나 수정할 때 이 작업을 수행할 수 있습니다.

자세한 내용은 *Informatica 10.4.0 Catalog 관리자 가이드*를 참조하십시오.

### 새 리소스

버전 10.4.0에서는 다음과 같은 새로운 리소스가 Enterprise Data Catalog에 추가됩니다.

- AWS Glue
- Microsoft Power BI
- Apache Cassandra
- Snowflake
- Google Cloud Storage

위의 모든 리소스에서 메타데이터, 관계 및 연계 정보를 추출할 수 있습니다. 자세한 내용은 *Informatica 10.4.0 Enterprise Data Catalog 스캐너 구성 가이드*를 참조하십시오.

### 참조 리소스 및 참조 자산

버전 10.4.0에서는 리소스를 구성하여 리소스에서 참조하는 카탈로그의 데이터 소스 또는 다른 리소스에 대한 메타데이터를 추출할 수 있습니다. 예를 들어 PowerCenter 매핑의 소스 및 대상 테이블, Tableau 보고서의 소스 테이블 및 파일이 포함됩니다. 이러한 참조된 데이터 소스는 참조 리소스라고 하며, 데이터 소스에 포함된 자산은 참조 자산이라고 합니다. Enterprise Data Catalog 응용 프로그램에서 이러한 참조 리소스 및 참조 자산을 확인할 수 있습니다. Enterprise Data Catalog에서 전체 연계를 보려면 참조 리소스에 대한 연결 할당을 수행할 수 있습니다. 참조 리소스와 참조 리소스 데이터 소스에 대해 Enterprise Data Catalog에서 구성한 리소스 간에 연결 할당을 수행할 수 있습니다.

다음과 같은 리소스를 구성하여 리소스에서 참조하는 카탈로그의 데이터 소스 또는 다른 리소스에 대한 메타데이터를 추출할 수 있습니다.

- PowerCenter

- AWS Glue
- Tableau Server
- Coudera Navigator
- Apache Atlas
- Informatica Intelligent Cloud Services
- Informatica Platform
- SQL Server 통합 서비스

자세한 내용은 *Informatica 10.4.0 Catalog 관리자 가이드* 및 *Informatica 10.4.0 Enterprise Data Catalog 사용자 가이드*를 참조하십시오.

## 검색 결과에서 자산 내보내기

버전 10.4.0에서는 검색 결과에 나타나는 모든 자산을 CSV(쉼표로 구분된 값) 파일로 내보내고 이 CSV 파일을 Enterprise Data Catalog로 다시 가져올 수 있습니다. 원하는 검색 탭에서 검색 결과를 구체적으로 지정한 후 자산을 내보내십시오.

자세한 내용은 *Informatica 10.4.0 Enterprise Data Catalog 사용자 가이드*에서 *자산 테스트* 장을 참조하십시오.

## 연계 및 영향 필터

버전 10.4.0에서는 필터를 생성하여 연계 및 영향 보기에 적용할 수 있습니다. 필터를 사용하여 연계 및 영향 보기에서 결과 범위를 좁히십시오. 연계 및 영향 보기에서 사용할 수 있는 자산 유형을 기반으로 다양한 필터 옵션 조합을 사용하여 사용자 지정된 필터를 생성합니다.

자세한 내용은 *Informatica 10.4.0 Enterprise Data Catalog 사용자 가이드*에서 *연계 및 영향 보기* 장을 참조하십시오.

## 자산 제어 요약

버전 10.4.0에서는 테이블 및 파일 자산 유형에 대한 **자산 제어 요약** 탭에서 제어 흐름 자산을 볼 수 있습니다. 제어 흐름은 데이터 흐름에 제약 또는 조건을 부과합니다. 예를 들어 SQL 쿼리가 WHERE 절에 제약 조건을 포함하거나 매핑이 변환을 포함할 수 있습니다.

자세한 내용은 *Informatica 10.4.0 Enterprise Data Catalog 사용자 가이드*에서 *연계 및 영향 보기* 장을 참조하십시오.

## 규칙 및 성과 기록표

버전 10.4.0에서는 프로파일링 웨어하우스에서 규칙 및 성과 기록표를 추출하여 Enterprise Data Catalog에서 볼 수 있습니다. 규칙 및 성과 기록표에 대한 열 결과를 보고 데이터 소스의 데이터 품질을 측정할 수 있습니다.

자세한 내용은 *Informatica 10.4.0 Enterprise Data Catalog 사용자 가이드*에서 *자산 보기* 장을 참조하십시오.

## 고유 키 유추

버전 10.4.0에서는 소스 데이터 개체의 열에서 고유 키 후보를 생성할 수 있습니다. 고유 키 유추 결과를 보고 중복 값이 있는 열을 확인하여 데이터 품질 문제를 파악할 수 있습니다.

유추된 고유 키 유추 결과를 수락하거나 거부할 수 있습니다. 유추된 고유 키 유추를 수락하거나 거부한 후 고유 키 유추를 재설정하여 유추된 상태를 복원할 수 있습니다.

자세한 내용은 *Informatica 10.4.0 Enterprise Data Catalog 사용자 가이드*에서 **자산 보기** 장을 참조하십시오.

## CLOB 파일 유형의 데이터 도메인 검색

버전 10.4.0에서는 CLOB 파일 유형에서 데이터 도메인 검색을 수행할 수 있습니다.

자세한 내용은 *Informatica 10.4.0 Enterprise Catalog 관리자 가이드*에서 **Enterprise Data Catalog 개념** 장을 참조하십시오.

## Spark 엔진의 데이터 검색 및 샘플링 옵션

버전 10.4.0에서는 프로필을 실행하여 Spark 엔진에서 데이터 도메인을 검색하고 샘플링 옵션을 선택할 수 있습니다.

### Spark 엔진의 데이터 도메인 검색

Spark 엔진에서 데이터 도메인 검색을 수행할 수 있습니다.

### Spark 엔진의 샘플링 옵션

다음과 같은 샘플링 옵션을 선택하여 Spark 엔진에서 데이터 도메인을 검색할 수 있습니다.

- **n개 제한** 샘플링 옵션은 데이터 개체의 행 수를 기반으로 프로필을 실행합니다. Hadoop 환경에서 데이터 도메인을 검색하도록 선택하면 Spark 엔진이 데이터 개체의 여러 파티션에서 샘플을 수집하고 샘플 크기를 계산하도록 단일 노드에 샘플을 푸시합니다.
- **무작위 백분율** 샘플링 옵션은 데이터 개체의 행에서 백분율만큼 프로필을 실행합니다.

자세한 내용은 *Informatica 10.4.0 Enterprise Catalog 관리자 가이드*에서 **Enterprise Data Catalog 개념** 장을 참조하십시오.

## 추적 기술 미리보기

Enterprise Data Catalog 버전 10.4.0에는 기술 미리보기로 제공되는 기능이 포함되어 있습니다.

기술 미리보기 기능은 지원되지만 보장되지 않으며 프로덕션용으로 준비되지 않았습니다. 비프로덕션 환경에서만 사용할 것을 권장합니다. Informatica는 향후 프로덕션용 GA 릴리스에 미리보기 기능을 포함할 계획이지만 시장 또는 기술적 상황의 변화에 따라 여의치 않을 경우 이러한 기능을 포함하지 않도록 선택할 수 있습니다. 자세한 내용은 Informatica 글로벌 고객 지원 센터에 문의하십시오.

- 버전 10.4.0에서는 연계 및 영향 보기의 압축 보기를 표시하도록 선택할 수 있습니다. 압축 연계 및 영향 보기에는 리소스 수준에서 요약된 연계 및 영향 다이어그램이 표시됩니다.

자세한 내용은 *Informatica 10.4.0 Enterprise Data Catalog 사용자 가이드*에서 **연계 및 영향 보기** 장을 참조하십시오.

- 버전 10.4.0에서는 SAP Business Warehouse, SAP BW/4HANA, IBM InfoSphere DataStage 및 Oracle Data Integrator 소스가 런타임에 액세스할 수 없거나 오프라인인 경우 이러한 소스에서 메타데이터를 추출할 수 있습니다.

자세한 내용은 *Informatica 10.4.0 Catalog 관리자 가이드*를 참조하십시오.

- 버전 10.4.0에서는 SAP Business Warehouse 및 SAP BW/4HANA 데이터 소스에서 메타데이터를 추출할 수 있습니다.

자세한 내용은 *Informatica 10.4.0 Enterprise Data Catalog 스캐너 구성 가이드*를 참조하십시오.

## 데이터 미리보기 및 프로비저닝

버전 10.4.0에서는 카탈로그에서 데이터 검색을 완료한 후 데이터 프로비저닝을 수행할 수 있습니다. 데이터 프로비저닝은 데이터를 대상으로 이동해 추가 분석을 수행하는 데 유용합니다. 관계형 소스에 대한 데이터를 미리 보고 데이터를 대상으로 이동하기 전에 확인할 수 있습니다.

데이터 미리보기 및 프로비저닝에 대한 자세한 내용은 *Informatica 10.4.0 Catalog 관리자 가이드* 및 *Informatica 10.4.0 Enterprise Data Catalog 사용자 가이드*를 참조하십시오.

## 독립 실행형 스캐너 유틸리티의 지원되는 리소스 유형

버전 10.4에서는 런타임이나 오프라인에서 액세스할 수 없는 경우 다음 외부 소스에서 메타데이터를 추출할 수 있습니다.

- Amazon Redshift
- Amazon S3
- Apache Cassandra
- Axon
- Azure Data Lake Store
- Azure Microsoft SQL Data Warehouse
- Azure Microsoft SQL Server
- Business Glossary
- 사용자 지정 연계
- 데이터베이스 스크립트
- Erwin
- Glue
- Google BigQuery
- Google Cloud Storage
- Informatica Cloud Service
- IBM Cognos
- Microsoft Azure Blob Storage
- Microsoft SQL Server 통합 서비스
- PowerBI
- QlikView Business Intelligence
- Salesforce
- SAP HANA
- SAP PowerDesigner
- Snowflake
- Workday

자세한 내용은 *Informatica 10.4 Enterprise Data Catalog 관리자 가이드*의 "오프라인 및 액세스할 수 없는 리소스의 메타데이터 추출" 장을 참조하십시오.



## REST API

버전 10.4에서는 다음과 같은 Informatica Enterprise Data Catalog REST API를 사용할 수 있습니다.

- **Data Provision REST API.** 기존 REST API 기능과 더불어, 사용자에게 데이터 프로비저닝이 제공되는지 여부를 확인하고 데이터 프로비저닝을 지원하는 리소스를 나열할 수 있습니다.
- **연계 필터 REST API.** 연계 필터를 생성, 업데이트, 나열 또는 삭제할 수 있습니다.
- **모델 정보 REST API.** 기존 REST API 기능과 더불어, 미리 정의된 슬라이더 패킷, 슬라이더 패킷 정의 및 연계 필터 정의를 나열할 수 있습니다.
- **모델 수정 REST API.** 기존 REST API 기능과 더불어, 슬라이더 패킷 정의를 생성, 업데이트 및 삭제할 수 있습니다.
- **모니터링 정보 REST API.** 개체 내보내기 유형, 개체 가져오기 유형, 리소스 내보내기 유형 및 검색 내보내기 유형의 작업을 포함하는 작업을 제출하거나 나열할 수 있습니다.
- **개체 하위 개수 REST API.** 개체의 총 하위 자산 수를 나열할 수 있습니다.
- **제품 정보 REST API.** 릴리스 버전, 빌드 버전 및 빌드 날짜를 포함하는 Enterprise Data Catalog에 대한 세부 정보를 나열할 수 있습니다.

REST API에 대한 자세한 내용은 *Informatica 10.4 Enterprise Data Catalog REST API* 참조를 참조하십시오.

## Enterprise Data Preparation

이 섹션에서는 버전 10.4.0의 새로운 Enterprise Data Preparation 기능에 대해 설명합니다.

### Data Lake 액세스 관리

버전 10.4.0에서는 Data Lake의 Hive 스키마 및 HDFS 위치를 나타내는 Enterprise Data Catalog 리소스에 대한 사용 권한을 Enterprise Data Preparation 사용자 및 사용자 그룹에 할당하여 Data Lake에 대한 액세스를 관리할 수 있습니다.

특정 스키마 또는 위치에 대한 사용 권한을 사용자 또는 사용자 그룹에 부여하면 사용자가 가져오기, 게시 또는 업로드 작업을 수행할 경우 응용 프로그램에는 사용자에게 사용 권한이 있는 스키마 및 위치만 표시됩니다.

자세한 내용은 *Enterprise Data Preparation 10.4.0 관리자 가이드*를 참조하십시오.

### 데이터 소스로 Microsoft Azure Data Lake Storage 사용

버전 10.4.0에서는 Microsoft Azure Data Lake Storage Gen2 위치를 데이터 소스로 사용할 수 있습니다.

### Data Lake에 파일 게시

버전 10.4.0에서는 파일 형태로 준비된 데이터를 Data Lake의 HDFS 위치에 게시할 수 있습니다.

데이터를 게시할 때 Data Lake에 데이터를 쓰기 위한 파일 유형을 선택할 수 있습니다. 예를 들어 데이터를 원본으로 구분된 값 파일로 게시하도록 선택하면 데이터가 Data Lake에 .csv 파일로 기록됩니다.

자세한 내용은 *Enterprise Data Preparation 10.4.0 사용자 가이드*를 참조하십시오.

### Data Lake에 파일 업로드

버전 10.4.0에서는 로컬 드라이브에서 Data Lake로 여러 파일 형식을 업로드할 수 있습니다.

다음과 같은 업로드 옵션에서 선택할 수 있습니다.

#### Data Lake에 직접 파일 업로드.

UTF-8 형식의 쉼표로 구분된 파일, Avro 파일, JSON 파일 또는 Parquet 파일은 데이터를 미리 볼 필요 없이 로컬 드라이브에서 Data Lake로 직접 업로드할 수 있습니다. 데이터를 미리 보지 않고 파일을 업로드하려는 경우 이 옵션을 선택할 수 있습니다.

#### CLAIRE가 파일 구조를 결정한 다음 Data Lake에 파일을 업로드하도록 허용.

쉼표로 구분된 파일 또는 Microsoft Excel 스프레드시트의 데이터를 Data Lake에 업로드할 수 있습니다. 이러한 파일을 업로드할 경우 Enterprise Data Preparation에서는 CLAIRE에 포함된 검색 엔진을 사용하여 파일의 구조를 결정하고 데이터의 미리보기를 표시합니다.

이 옵션을 사용하여 Excel 스프레드시트를 업로드하는 경우 CLAIRE 엔진에서 스프레드시트의 시트 및 테이블을 검색합니다. 미리 보려는 시트 및 테이블을 선택할 수 있습니다.

**참고:** Excel 스프레드시트 업로드는 미리보기 기능입니다.

#### 파일 구조를 정의한 다음 Data Lake에 파일 업로드.

쉼표로 구분된 파일의 데이터를 로컬 드라이브에서 Data Lake로 업로드할 수 있습니다. 이러한 파일을 업로드할 때 데이터를 미리 보고, 파일의 구조를 지정하고, 요구 사항에 맞도록 열 특성을 구성할 수 있습니다. 파일을 업로드하기 전에 열 특성을 수정해야 하는 경우 이 옵션을 선택할 수 있습니다.

자세한 내용은 *Enterprise Data Preparation 10.4.0 사용자 가이드*를 참조하십시오.

## Informatica 매핑

이 섹션에서는 버전 10.4.0에 포함된 Informatica 매핑의 새로운 기능에 대해 설명합니다.

### 매핑 매개 변수에 매핑 출력 바인딩

버전 10.4.0에서는 매핑 출력을 유지하고 원시 환경이나 Spark 엔진에서 실행되도록 배포된 매핑에 바인딩할 수 있습니다.

매핑 출력을 생성합니다. 이 출력을 매핑 매개 변수에 바인딩하면 다음번 매핑 실행에서 해당 값을 사용할 수 있습니다. 매핑을 실행할 때 데이터 통합 서비스는 매핑 출력의 값을 매핑 매개 변수에 전달합니다. 매핑 출력을 유지하려면 `infacmd ms runMapping` 명령의 `-RuntimeInstanceName` 옵션을 사용하여 런타임 인스턴스 이름을 지정해야 합니다.

이제 매핑 출력을 매개 변수에 바인딩할 수 있도록 Developer tool의 매핑 속성 보기에 바인딩 열이 포함됩니다.

배포된 매핑의 매핑 출력에 대한 자세한 내용은 *Informatica 10.4.0 Developer 매핑 가이드*에서 "매핑 출력" 장을 참조하십시오.

## infacmd ms 명령

다음 테이블에는 새/업데이트된 **infacmd ms** 명령이 설명되어 있습니다.

명령	설명
deleteMappingPersistedOutputs	배포된 매핑에 대한 모든 지속형 매핑 출력을 삭제하는 새 명령입니다. 삭제할 출력은 응용 프로그램의 이름과 매핑의 런타임 인스턴스 이름을 사용하여 지정합니다. 특정 출력을 삭제하려면 <code>-OutputNamesToDelete</code> 옵션을 사용하십시오.
getMappingStatus	이제 작업 이름을 반환하는 업데이트된 명령입니다. <code>infacmd ms runMapping</code> 에서 런타임 인스턴스 이름을 정의한 경우 작업 이름은 해당 런타임 인스턴스 이름입니다.
listMappingPersistedOutputs	배포된 매핑에 대한 지속형 매핑 출력을 나열하는 새 명령입니다. 출력은 응용 프로그램의 이름과 매핑의 런타임 인스턴스 이름을 기반으로 나열됩니다.

자세한 내용은 *Informatica 10.4.0 명령 참조*에서 "infacmd ms 명령 참조" 장을 참조하십시오.

## CLAIRE 권장 사항 및 인사이트

버전 10.4.0에서는 CLAIRE 인공 지능을 사용하여 개발하는 동안 매핑을 조정하고 보정할 수 있습니다.

권장 사항을 활성화하면 매핑을 개발할 때 매핑에서 CLAIRE가 자동으로 실행되어 매핑을 보정하거나 조정하는데 사용할 수 있는 권장 사항이 표시됩니다.

프로젝트 또는 프로젝트 폴더 내의 매핑에서 CLAIRE 분석을 실행할 수도 있습니다. 매핑 그룹을 분석할 경우 CLAIRE는 매핑 간의 유사성에 대한 인사이트를 표시합니다.

권장 사항 및 인사이트에 대한 자세한 내용은 *Data Engineering Integration User Guide*에서 참조하십시오.

## 매핑 최적화 프로그램 수준 업데이트

버전 10.4.0에서는 `infacmd ms UpdateOptimizationDefaultLevel` 명령을 사용하여 응용 프로그램의 매핑 최적화 프로그램 수준을 기본 수준인 자동으로 설정할 수 있습니다.

이 명령을 실행할 때 응용 프로그램 이름을 지정해야 합니다. `UpdateOptimizationDefaultLevel`은 응용 프로그램의 모든 매핑에 대한 최적화 프로그램 수준을 설정합니다.

자세한 내용은 *Informatica 10.4.0 명령 참조* 및 *Informatica 10.4.0 Developer 매핑 가이드*를 참조하십시오.

# Informatica 변환

이 섹션에서는 버전 10.4.0에 포함된 Informatica 변환의 새로운 기능에 대해 설명합니다.

## 주소 유효성 검사기 변환

이 섹션에서는 새로운 주소 유효성 검사기 변환 기능에 대해 설명합니다.

주소 유효성 검사기 변환에는 다음과 같은 국가에 대해 추가적인 주소 기능이 포함됩니다.

## 미국

버전 10.4에서는 주소 유효성 검사기가 미국 주소에서 MC를 MSC(Mail Stop Code)의 대체 버전으로 인식합니다.

버전 10.4에 포함되는 주소 확인 소프트웨어 엔진의 기능 및 작동에 대한 종합적인 정보는 *Informatica Address Verification 5.15.0 개발자 가이드*를 참조하십시오.

# Informatica 워크플로우

이 섹션에서는 버전 10.4.0의 새로운 Informatica 워크플로우 기능에 대해 설명합니다.

## Amazon EMR 클러스터 생성 태스크 고급 속성

버전 10.4.0에서는 워크플로우에서 Amazon EMR 연결을 사용하여 클러스터 생성 태스크를 구성할 때 새로운 고급 속성을 사용할 수 있습니다.

클러스터 생성 태스크에서 다음과 같은 기능을 구성할 수 있습니다.

- 루트 장치 EBS 볼륨 크기. EBS 루트 장치 볼륨의 크기(GB)입니다.
- 사용자 지정 AMI ID. 사용자 지정 Amazon Linux AMI(Amazon Machine Image)의 ID입니다.
- 보안 구성. 클러스터의 인증 및 암호화를 위한 보안 구성의 이름입니다.

자세한 내용은 *Data Engineering Integration 10.4.0 사용자 가이드* 및 *Informatica® 10.4.0 Developer 워크플로우 가이드*를 참조하십시오.

# Informatica 설치

이 섹션에서는 10.4.0의 새로운 설치 기능에 대해 설명합니다.

## PostgreSQL

버전 10.4.0에서는 도메인 구성 리포지토리, 모델 리포지토리 및 PowerCenter 리포지토리에 PostgreSQL 데이터베이스를 사용할 수 있습니다. Enterprise Data Preparation의 경우 PostgreSQL 데이터베이스를 추가적인 모델 리포지토리 서비스에만 사용할 수 있습니다.

또한 PostgreSQL이 Linux 또는 Windows에서 작동하도록 psql 클라이언트 응용 프로그램 버전 10.6을 설치할 수 있습니다.

PostgreSQL에 대한 자세한 내용은 *Informatica 10.4.0 설치 가이드*를 참조하십시오.

## 자동 모드의 설치 전(i10Pi) 시스템 검사 도구

버전 10.4.0에서는 자동 모드에서 설치 전(i10pi) 시스템 검사 도구를 실행할 수 있습니다. 자동 모드에서 i10pi를 실행하여 사용자 상호 작용 없이 시스템이 설치 요구 사항을 충족하는지 여부를 확인할 수 있습니다.

자동 모드에서 i10Pi를 실행하는 것에 대한 자세한 내용은 Informatica 10.4.0 설치 가이드를 참조하십시오.

## 자동 설치 속성 파일의 암호 암호화

버전 10.4.0에서는 Informatica 설치 프로그램에 자동 모드에서 서비스를 설치하는 데 필요한 옵션을 포함하는 속성 파일에 설정된 암호를 암호화하는 데 사용할 수 있는 유틸리티가 포함되어 있습니다.

설치 프로그램을 자동 모드로 실행하는 경우 설치 프레임워크에서 암호화된 암호를 해독합니다.

자세한 내용은 *Informatica 설치 및 구성 가이드*를 참조하십시오.

## 지능형 구조 모델

이 섹션에서는 버전 10.4.0의 새로운 지능형 구조 모델 기능에 대해 설명합니다.

### 추가 입력 유형

버전 10.4.0에서는 Intelligent Structure Discovery가 ORC, Avro 및 Parquet 입력을 처리할 수 있습니다.

자세한 내용은 *Data Engineering Integration 10.4.0 사용자 가이드*를 참조하십시오.

### 디자인 타임에 샘플에서 모델 생성

버전 10.4.0에서는 복합 데이터 개체를 만들 때 선택한 샘플 파일을 기반으로 Intelligent Structure Discovery에서 모델을 생성할 수 있습니다. 그런 다음 Informatica Intelligent Cloud Services Data Integration에서 모델을 세밀하게 조정할 수 있습니다.

이 기능에서는 XML, JSON, ORC, AVRO 및 Parquet 샘플 파일을 지원합니다.

자세한 내용은 *Data Engineering Integration 10.4.0 사용자 가이드*를 참조하십시오.

### 식별되지 않은 데이터 처리

버전 10.4.0에서는 Intelligent Structure Discovery에서 모델의 식별되지 않은 데이터에 대한 출력을 구조화된 JSON 형식으로 정렬합니다.

자세한 내용은 *Data Engineering Integration 10.4.0 사용자 가이드*를 참조하십시오.

## PowerCenter

이 섹션에서는 버전 10.4.0의 새로운 PowerCenter 기능에 대해 설명합니다.

### HTTP 변환

버전 10.4.0에서는 HTTP 변환에 최종 URL 구성을 위한 메서드(SIMPLE PATCH, SIMPLE PUT 및 SIMPLE DELETE)도 포함됩니다.

부분 업데이트를 수행할 수 있으며 SIMPLE PATCH 메서드에서 입력 데이터가 완전한 본문일 필요가 없습니다. 이 메서드를 사용하여 입력 포트의 데이터를 리소스에 대한 패치로 업데이트할 수 있습니다.

SIMPLE PUT 메서드로 문서의 완전한 대체를 수행할 수 있습니다. 입력 포트에서 HTTP 서버에 대한 단일 데이터 블록으로 데이터를 생성할 수 있습니다. 데이터가 이미 존재하는 경우 입력 포트에서 HTTP 서버에 대한 단일 데이터 블록으로 데이터를 업데이트할 수 있습니다.

SIMPLE DELETE 메서드로 HTTP 서버에서 데이터를 삭제할 수 있습니다.

HTTP 변환을 위한 기본 URL을 매개 변수화할 수도 있습니다.

이전에는 SIMPLE GET 및 SIMPLE POST의 두 메서드에 대해서만 최종 URL 구성을 지정할 수 있었습니다. 또한 HTTP 변환에 대한 최종 URL을 매개 변수화할 수 없었습니다.

자세한 내용은 *PowerCenter 10.4.0 변환 가이드*에서 "HTTP 변환" 장을 참조하십시오.

## 연결

이 섹션에서는 버전 10.4.0에서 새로 추가된 연결 기능에 대해 설명합니다.

### Oracle 연결을 위한 사용자 가장

버전 10.4.0에서는 PowerCenter의 Oracle 연결에서 가장 사용자 이름을 지정하여 Oracle 소스 및 대상에 연결할 수 있습니다. 가장 사용자를 지정하면 사용자를 대신하여 Oracle에 연결하고 읽기 또는 쓰기 작업을 수행할 수 있습니다.

자세한 내용은 *Informatica 10.4.0 PowerCenter 워크플로우 기본 가이드*를 참조하십시오.

### SAP HANA 데이터베이스 모델링 보기 지원

버전 10.4.0에서는 다음과 같은 유형의 SAP HANA 데이터베이스 모델링 보기에서 데이터를 읽을 수 있습니다.

- 분석 보기
- 특성 보기
- 계산된 보기

자세한 내용은 *Informatica 10.4.0 PowerCenter Designer 가이드*를 참조하십시오.

# PowerExchange 어댑터

이 섹션에서는 버전 10.4.0의 새로운 PowerExchange 어댑터 기능에 대해 설명합니다.

## PowerExchange Adapters for Informatica

이 섹션에서는 버전 10.4.0의 새로운 Informatica 어댑터 기능에 대해 설명합니다.

### PowerExchange for Amazon Redshift

버전 10.4.0에서는 PowerExchange for Amazon Redshift에 다음과 같은 기능이 포함됩니다.

- AWS Databricks 환경에서 매핑을 실행할 수 있습니다.
- JDBC URL 연결 속성에서 클러스터 지역 이름을 지정한 경우에도 클러스터 지역 연결 속성에서 클러스터 지역 이름을 선택할 수 있습니다.
- Amazon Redshift에서 데이터를 읽을 때 null 값을 유지할 수 있습니다.
- Amazon Redshift에 데이터를 쓸 때 일괄 처리당 준비 파일의 수를 지정할 수 있습니다.
- CDC 소스에서 Amazon Redshift 대상으로 데이터를 쓸 때 레코드 순서를 유지할 수 있습니다.

자세한 내용은 *Informatica 10.4.0 PowerExchange for Amazon Redshift 사용자 가이드*를 참조하십시오.

## PowerExchange for Amazon S3

버전 10.4.0에서는 PowerExchange for Amazon S3에 다음과 같은 기능이 포함됩니다.

- AWS Databricks 환경에서 매핑을 실행할 수 있습니다.
- AssumeRole을 사용하는 임시 보안 자격 증명을 사용하여 AWS 리소스에 액세스할 수 있습니다.
- 런타임 시 읽기 및 쓰기 작업 속성에서 데이터 형식 유형 및 스키마를 매개 변수화할 수 있습니다.

자세한 내용은 *Informatica 10.4.0 PowerExchange for Amazon S3 사용자 가이드*를 참조하십시오.

## PowerExchange for Google BigQuery

버전 10.4.0에서는 PowerExchange for Google BigQuery에 다음과 같은 기능이 포함됩니다.

- Google Dataproc 클러스터를 사용하여 Spark 엔진에서 매핑을 실행할 수 있습니다.
- 최적화된 Spark 모드에서 매핑을 실행하여 매핑의 성능을 개선할 수 있습니다. 최적화된 Spark 모드를 사용하여 데이터를 읽는 경우 사용할 파티션의 수를 지정할 수 있습니다. 고급 읽기 및 쓰기 작업 속성에서 매핑을 일반 모드에서 실행할지, 아니면 최적화 모드에서 실행할지를 지정할 수 있습니다. 최적화된 Spark 모드는 매핑 성능을 향상시킵니다.
- SQL 재정의의 구성하여 Google BigQuery 소스에서 데이터를 추출하는 데 사용되는 기본 SQL 쿼리를 재정의할 수 있습니다.
- NUMERIC 데이터 유형의 데이터를 Google BigQuery에서 읽거나 쓸 수 있습니다. NUMERIC 데이터 유형은 전체 자릿수가 38자리이고 소수점 이하 자릿수가 9자리인 정확한 숫자 값입니다. NUMERIC 데이터 유형을 읽거나 쓸 때 데이터 통합 서비스에서는 NUMERIC 데이터 유형을 10진수 변환 데이터 유형에 매핑하며 허용되는 전체 자릿수는 최대 38자리이고 소수점 이하 자릿수는 최대 9자리입니다.

자세한 내용은 *Informatica 10.4.0 PowerExchange for Google BigQuery 사용자 가이드*를 참조하십시오.

## PowerExchange for Google Cloud Storage

버전 10.4.0에서는 PowerExchange for Google Cloud Storage에 다음 기능이 포함됩니다.

- Google Dataproc 클러스터를 사용하여 Spark 엔진에서 매핑을 실행할 수 있습니다.
- Google Cloud Storage 소스에서 데이터를 읽을 때 다음과 같은 Google Cloud Storage 데이터 개체 읽기 작업 고급 속성을 구성할 수 있습니다.

### Google Cloud Storage 경로

Google Cloud Storage 데이터 개체에서 선택한 파일에 대한 Google Cloud Storage 경로를 재정의합니다.

다음 형식을 사용합니다.

gs://<버킷 이름> 또는 gs://<버킷 이름>/<폴더 이름>

### 소스 파일 이름

Google Cloud Storage 데이터 개체에서 지정한 Google Cloud Storage 소스 파일 이름을 재정의합니다.

### 디렉터리 여부

Google Cloud Storage 경로 데이터 개체 읽기 작업 고급 속성에서 지정한 폴더에서 사용할 수 있는 모든 파일을 읽습니다.

자세한 내용은 *Informatica 10.4.0 PowerExchange for Google Cloud Storage 사용자 가이드*를 참조하십시오.

## PowerExchange for Microsoft Azure Blob Storage

버전 10.4.0에서는 PowerExchange for Microsoft Azure Blob Storage에 다음 기능이 포함됩니다.

- 런타임 시 읽기 및 쓰기 작업 속성에서 데이터 형식 유형 및 스키마를 매개 변수화할 수 있습니다.
- Microsoft Azure Blob Storage 연결을 생성하는 동안 공유 액세스 서명 인증을 사용할 수 있습니다.

자세한 내용은 *Informatica 10.4.0 PowerExchange for Microsoft Azure Blob Storage 사용자 가이드*를 참조하십시오.

## PowerExchange for Microsoft Azure SQL Data Warehouse

버전 10.4.0에서는 가상 네트워크(VNet)에 상주하는 Microsoft Azure SQL Data Warehouse 끝점에서 데이터를 읽고 쓸 수 있습니다.

자세한 내용은 *Informatica 10.4.0 PowerExchange for Microsoft Azure SQL Data Warehouse 사용자 가이드*를 참조하십시오.

## PowerExchange for Salesforce

버전 10.4.0에서는 PowerExchange for Salesforce에 다음과 같은 기능이 포함됩니다.

- 버전 45.0, 46.0 및 47.0의 Salesforce API를 사용하여 Salesforce 연결을 생성하고 Salesforce 개체에 액세스할 수 있습니다.
- 상위 개체의 공유 항목을 나타내는 공유 개체의 쿼리에 대해 기본 키 청크를 활성화할 수 있습니다. 상위 개체가 지원되는 경우에만 공유 개체에 대한 기본 키 청크가 지원됩니다. 예를 들어 CaseHistory를 쿼리하려는 경우 상위 개체인 Case에 대해 기본 키 청크가 지원되어야 합니다.
- 표준 API를 사용하여 Lead 및 Case 대상 개체에 대한 레코드를 삽입, 업데이트 또는 upsert할 때 레코드의 특성을 재할당하는 할당 규칙을 생성할 수 있습니다.

자세한 내용은 *Informatica 10.4.0 PowerExchange for Salesforce 사용자 가이드*를 참조하십시오.

## PowerExchange for SAP NetWeaver

버전 10.4.0에서는 PowerExchange for SAP NetWeaver에 다음과 같은 기능이 포함됩니다.

- SAP 테이블 판독기 매핑에 대한 HTTPS 스트리밍을 구성할 수 있습니다.
- SAP NetWeaver 시스템 버전이 7.50 이상인 경우 SAP 테이블 판독기를 사용하여 ABAP CDS 보기에서 데이터를 읽을 수 있습니다.
- 다음과 같은 데이터 유형이 있는 필드가 포함된 SAP 테이블에서 데이터를 읽을 수 있습니다.
  - DF16\_DEC
  - DF32\_DEC
  - DF16\_RAW
  - DF34\_RAW
  - INT8
  - RAWSTRING
  - SSTRING
  - STRING

자세한 내용은 *Informatica PowerExchange for SAP NetWeaver 10.4.0 사용자 가이드*를 참조하십시오.



## PowerExchange for Snowflake

버전 10.4.0에서는 PowerExchange for Snowflake에 다음과 같은 기능이 포함됩니다.

- Databricks 환경에서 Snowflake 매핑을 실행할 수 있습니다.
- Snowflake 개체를 매핑의 동적 소스 및 대상으로 사용할 수 있습니다.
- **대상 생성** 옵션을 사용하여 Snowflake 대상을 생성할 수 있습니다.
- 매핑에서 Snowflake 대상에 대한 대상 스키마 전략을 구성할 수 있습니다. 사용 가능한 옵션에서 선택하여 기존 대상 스키마를 유지하거나 스키마가 없는 경우 대상을 생성할 수 있습니다. 또한 대상 스키마 전략 옵션을 매개 변수 값으로 지정할 수 있습니다.
- Snowflake 고급 대상 속성에 데이터 통합 서비스에서 대상에 쓰는 동안 거부된 레코드를 쓰는 데 사용할 거부된 파일 이름 및 경로를 지정할 수 있습니다.
- Snowflake ODBC 연결의 ODBC 공급자 유형이 Snowflake인 경우 푸시다운 최적화를 구성하여 변환 논리를 Snowflake 데이터베이스에 푸시할 수 있습니다.
- 데이터 통합 서비스 속성에서 EnableSDKDecimal38 사용자 지정 플래그를 구성하여 38비트 전체 자릿수 및 소수점 이하 자릿수의 10진수 데이터 유형인 데이터를 읽고 쓸 수 있습니다.

자세한 내용은 *Informatica 10.4.0 PowerExchange for Snowflake 사용자 가이드*를 참조하십시오.

## PowerExchange for HDFS

버전 10.4.0에서는 PowerExchange for HDFS에 다음과 같은 기능이 포함됩니다.

- 런타임 시 읽기 및 쓰기 작업 속성에서 데이터 형식 유형 및 스키마를 매개 변수화할 수 있습니다.
- 읽기 또는 쓰기 작업에 대한 복합 파일 데이터 개체의 스키마를 형식화할 수 있습니다.

자세한 내용은 *Informatica 10.4.0 PowerExchange for HDFS 사용자 가이드*를 참조하십시오.

## PowerExchange Adapters for PowerCenter

이 섹션에서는 버전 10.4.0의 새로운 PowerCenter 어댑터 기능에 대해 설명합니다.

### PowerExchange for Google BigQuery

버전 10.4.0에서는 NUMERIC 데이터 유형의 데이터를 Google BigQuery에서 읽거나 쓸 수 있습니다. NUMERIC 데이터 유형은 전체 자릿수가 38자리이고 소수점 이하 자릿수가 9자리인 정확한 숫자 값입니다. NUMERIC 데이터 유형을 읽거나 쓸 때 PowerCenter 통합 서비스에서는 NUMERIC 데이터 유형을 10진수 변환 데이터 유형에 매핑하며 허용되는 전체 자릿수는 최대 28자리이고 소수점 이하 자릿수는 최대 9자리입니다.

자세한 내용은 *PowerCenter용 Informatica 10.4.0 PowerExchange for Google BigQuery 사용자 가이드*를 참조하십시오.

### PowerExchange for Google Cloud Storage

버전 10.4.0에서는 Google Cloud Storage 소스에서 데이터를 읽을 때 다음과 같은 Google Cloud Storage 데이터 개체 읽기 작업 고급 속성을 구성할 수 있습니다.

#### Google Cloud Storage 경로

Google Cloud Storage 데이터 개체에서 선택한 파일에 대한 Google Cloud Storage 경로를 재정의합니다.

다음 형식을 사용합니다.

gs://<버킷 이름> 또는 gs://<버킷 이름>/<폴더 이름>

## 소스 파일 이름

Google Cloud Storage 데이터 개체에서 지정한 Google Cloud Storage 소스 파일 이름을 재정의합니다.

## 디렉터리 여부

**Google Cloud Storage 경로** 데이터 개체 읽기 작업 고급 속성에서 지정한 폴더에서 사용할 수 있는 모든 파일을 읽습니다.

자세한 내용은 *PowerCenter용 Informatica 10.4.0 PowerExchange for Google Cloud Storage 사용자 가이드*를 참조하십시오.

## PowerExchange for Greenplum

버전 10.4.0에서는 PowerExchange for Greenplum을 사용하여 Greenplum의 데이터를 읽을 수 있습니다. Greenplum 소스가 Greenplum에서 데이터를 추출하는 방법을 결정할 수 있도록 특정 세션 속성을 구성할 수 있습니다.

Greenplum 세션을 실행하여 데이터를 읽을 경우 PowerCenter 통합 서비스에서는 Greenplum의 파일 배포 프로그램인 Greenplum 데이터베이스 병렬 파일 서버(gpfdist)를 호출하여 데이터를 읽습니다.

자세한 내용은 *PowerCenter용 Informatica 10.4.0 PowerExchange for Greenplum 사용자 가이드*를 참조하십시오.

## PowerExchange for JD Edwards EnterpriseOne

버전 10.4.0에서는 JD Edwards EnterpriseOne API 버전 9.2를 사용하여 JD Edwards EnterpriseOne 연결을 생성하고 JD Edwards EnterpriseOne 개체에 액세스할 수 있습니다.

자세한 내용은 *PowerCenter용 Informatica 10.4.0 PowerExchange for JD Edwards EnterpriseOne 사용자 가이드*를 참조하십시오.

## PowerExchange for Kafka

버전 10.4.0에서는 다음과 같은 SSL 속성을 구성하여 Kafka 브로커에 대한 보안 연결을 활성화할 수 있습니다.

- SSL 모드
- SSL TrustStore 파일 경로
- SSL TrustStore 암호
- SSL 키 저장소 파일 경로
- SSL 키 저장소 암호

Kafka 메시징 브로커를 Kafka 브로커 버전 0.10.1.1 이상을 사용하도록 구성할 수 있습니다.

자세한 내용은 *PowerCenter용 Informatica PowerExchange for Kafka 10.4.0 사용자 가이드*를 참조하십시오.

## PowerExchange for Salesforce

버전 10.4.0에서는 버전 46.0 및 47.0의 Salesforce API를 사용하여 Salesforce 연결을 생성하고 Salesforce 개체에 액세스할 수 있습니다.

자세한 내용은 *PowerCenter용 Informatica 10.4.0 PowerExchange for Salesforce 사용자 가이드*를 참조하십시오.

## PowerExchange for SAP NetWeaver

버전 10.4.0에서는 PowerExchange for SAP Dynamic ABAP Table Extractor를 사용하여 HTTP/HTTPS 스트리밍을 통해 SAP 테이블 및 ABAP CDS(Core Data Services) 보기에서 데이터를 읽을 수 있습니다. SAP NetWeaver 시스템 버전이 7.50 이상인 경우 PowerExchange for SAP Dynamic ABAP Table Extractor를 사용하여 ABAP CDS 보기에서 데이터를 읽을 수 있습니다.

자세한 내용은 *Informatica PowerExchange for SAP NetWeaver 10.4.0 사용자 가이드*를 참조하십시오.

## 보안

이 섹션에서는 버전 10.4.0의 새로운 보안 기능에 대해 설명합니다.

### 서로 다른 SAML ID 공급자를 사용하도록 웹 응용 프로그램 구성

버전 10.4.0에서는 한 도메인에서 실행되는 Informatica 웹 응용 프로그램이 서로 다른 SAML(Security Assertion Markup Language) ID 공급자를 사용하도록 구성할 수 있습니다. 예를 들어 Informatica Administrator는 Active Directory Federation Services를 ID 공급자로 사용하도록 구성하고 Informatica Analyst는 PingFederate를 ID 공급자로 사용하도록 구성할 수 있습니다.

자세한 내용은 *Informatica 10.4.0 보안 가이드*를 참조하십시오.

## 제 2 장

# 새로운 기능(10.2.2 HotFix 1)

이 장에 포함된 항목:

- [명령줄 프로그램 , 36](#)
- [Enterprise Data Catalog, 36](#)

## 명령줄 프로그램

이 섹션에서는 버전 10.2.2 HotFix 1의 새로운 명령에 대해 설명합니다.

### infacmd ldm 명령

다음 테이블에는 새로운 infacmd ldm 명령 옵션이 설명되어 있습니다.

새 옵션	설명
-Force -fr	선택 사항입니다. 백업 모드가 오프라인일 때 백업을 강제하려는 경우 사용합니다. 백업을 강제로 수행하고 기존 백업을 덮어씁니다.
-Force -fr	선택 사항입니다. HDFS 및 Apache Zookeeper의 기존 콘텐츠를 지우려는 경우 사용합니다. 백업 데이터를 강제로 복원합니다.

자세한 내용은 *Informatica 10.2.2 HotFix 1 명령 참조*에서 "infacmd ldm 명령 참조" 장을 참조하십시오.

## Enterprise Data Catalog

이 섹션에서는 버전 10.2.2 HotFix 1의 새로운 Enterprise Data Catalog 기능에 대해 설명합니다.

### Azure Data Lake Storage Gen2

버전 10.2.2 HotFix 1에서는 Azure Data Lake Storage Gen2 소스의 메타데이터를 추출할 수 있습니다. Azure Data Lake Storage Gen2 저장소 계정의 파일 및 폴더에 포함된 메타데이터에 액세스하고 Enterprise Data Catalog로 추출할 수 있습니다.

자세한 내용은 *Informatica 10.2.2 HotFix1 Enterprise Data Catalog 스캐너 구성 가이드*에서 "Azure Data Lake 저장소" 장을 참조하십시오.

## 대/소문자 구분 없는 연결

버전 10.2.2 HotFix 1에서는 리소스를 대/소문자를 구분하지 않는 리소스로 구성할 수 있습니다. 이 옵션을 활성화하면 리소스 이름의 대/소문자에 관계없이 소스와 대상 간의 연결된 연결을 볼 수 있습니다.

대/소문자 구분 없는 연결에 대한 자세한 내용은 *Informatica 10.2.2 HotFix 1 Catalog 관리자 가이드*에서 "리소스 관리" 장을 참조하십시오.

## Enterprise Data Catalog Tableau 확장

버전 10.2.2 HotFix 1에서는 Enterprise Data Catalog Tableau 확장이 Tableau 내에서 사용할 수 있는 Tableau 대시보드의 기본 확장입니다.

Enterprise Data Catalog Tableau 확장을 Tableau Desktop, Tableau Server 및 모든 Tableau 지원 웹 브라우저에서 사용할 수 있습니다. Enterprise Data Catalog 응용 프로그램에서 이 확장을 다운로드하고 Tableau의 대시보드에 추가하십시오.

확장에 대한 자세한 내용은 *Informatica 10.2.2 HotFix1 Enterprise Data Catalog Extension for Tableau* 가이드를 참조하십시오.

## 새 리소스

버전 10.2.2 HotFix 1에서는 다음과 같은 새로운 리소스가 Enterprise Data Catalog에 추가됩니다.

- SAP PowerDesigner. SAP PowerDesigner 데이터 소스에서 메타데이터, 관계 및 연계 정보를 추출할 수 있습니다.
- SAP HANA. SAP HANA 데이터베이스에서 개체 및 연계 메타데이터를 추출할 수 있습니다.

자세한 내용은 *Informatica 10.2.2 HotFix1 스캐너 구성 가이드*를 참조하십시오.

## 프로파일링 웨어하우스 스캐너

버전 10.2.2 HotFix 1에서는 하나 이상의 프로파일링 웨어하우스 리소스를 생성할 수 있습니다. 프로파일링 웨어하우스를 연결하고 리소스에서 스키마를 선택할 수 있습니다. 리소스를 실행하면 스캐너가 연결된 프로파일링 웨어하우스에서 선택한 스키마에 대한 프로필 결과를 추출하고 카탈로그로 마이그레이션합니다.

자세한 내용은 *Informatica 10.2.2 HotFix1 Enterprise Data Catalog 스캐너 구성 가이드*에서 "Informatica Platform 스캐너 구성" 장을 참조하십시오.

## REST API

버전 10.2.2 HotFix 1에서는 다음과 같은 Informatica Enterprise Data Catalog REST API를 사용할 수 있습니다.

- Data Provision REST API. 연결 및 리소스를 반환하거나 업데이트하거나 삭제할 수 있습니다.
- Catalog Model REST API. 기존 REST API에 더해 필드 패킷, 쿼리 패킷 및 검색 탭에 액세스하거나 업데이트하거나 삭제할 수 있습니다.
- Object API. 기존 REST API에 더해 카탈로그 검색 및 제안을 나열할 수 있습니다.

REST API에 대한 자세한 내용은 *Informatica 10.2.2 HotFix 1 Enterprise Data Catalog REST API* 참조를 참조하십시오.

## 검색 개선 사항

버전 10.2.2 HotFix 1에서는 다음과 같은 개선 사항에 따라 Enterprise Data Catalog에서 개선된 검색을 수행할 수 있습니다.

### 큰따옴표 검색

큰따옴표("")를 사용한 자산 검색을 수행하여 카탈로그에서 자산 이름의 변형이 아닌 큰따옴표 내의 자산 이름과 정확히 일치하는 자산을 찾을 수 있습니다.

### 검색 연산자

최신 검색 연산자를 사용하여 정확한 검색 결과를 얻을 수 있습니다. 검색 연산자는 AND, OR, NOT, title 및 description입니다.

### 검색 순위

Enterprise Data Catalog는 순위 알고리즘을 사용하여 검색 결과 페이지에 데이터 자산의 순위를 표시합니다. 검색 순위는 특정 검색 결과에 포함된 다른 자산과 비교한 자산의 우선 순위를 나타냅니다.

### 관련 검색

검색 결과 페이지에서 **관련 검색 표시** 옵션을 활성화하여 관련 자산을 표시할 수 있습니다.

검색 개선 사항에 대한 자세한 내용은 *Informatica 10.2.2 HotFix 1 Enterprise Data Catalog 사용자 가이드*에서 "자산 검색" 장을 참조하십시오.

## 검색 탭

버전 10.2.2 HotFix 1에서는 자산 검색을 수행할 때 동일한 검색 조건을 반복적으로 수행하지 않고도 검색 탭을 사용하여 자산을 검색할 수 있습니다. 검색 탭은 Catalog에 미리 정의된 필터입니다.

검색 탭에 대한 자세한 내용은 *Informatica 10.2.2 HotFix 1 Enterprise Data Catalog 사용자 가이드*에서 "검색 사용자 지정" 장을 참조하십시오.

## 독립 실행형 스캐너 유틸리티의 지원되는 리소스 유형

버전 10.2.2 HotFix 1에서는 런타임이나 오프라인에서 액세스할 수 없는 경우 다음 외부 소스에서 메타데이터를 추출할 수 있습니다.

- Apache Atlas
- Cloudera Navigator
- 파일 시스템
- HDFS
- Hive
- Informatica Platform
- MicroStrategy
- OneDrive
- Oracle Business Intelligence
- SharePoint
- Sybase
- Tableau

자세한 내용은 *Informatica 10.2.2 HotFix 1 Enterprise Data Catalog 관리자 가이드*의 "오프라인 및 액세스할 수 없는 리소스의 메타데이터 추출" 장을 참조하십시오.

## 기술 미리 보기

Enterprise Data Catalog 버전 10.2.2 HotFix 1에는 기술 미리 보기로 제공되는 기능이 포함되어 있습니다.

기술 미리 보기 기능은 지원되지만 보장되지 않으며 프로덕션용으로 준비되지 않았습니다. 비프로덕션 환경에서만 사용할 것을 권장합니다. Informatica는 향후 프로덕션용 GA 릴리스에 미리 보기 기능을 포함할 계획이지만 시장 또는 기술적 상황의 변화에 따라 여의치 않을 경우 이러한 기능을 포함하지 않을 수 있습니다. 자세한 내용은 Informatica 글로벌 고객 지원 센터에 문의하십시오.

버전 10.2.2 HotFix 1에서는 다음과 같은 기능을 기술 미리 보기로 사용할 수 있습니다.

- 버전 10.2.2 HotFix 1에서는 Oracle Data Integrator 데이터 소스의 변환 논리를 포함하여 열 수준에서 데이터 연계에 대한 메타데이터를 추출할 수 있습니다.
- 버전 10.2.2 HotFix 1에서는 IBM InfoSphere DataStage 데이터 소스의 변환 논리를 포함하여 열 수준에서 데이터 연계에 대한 메타데이터를 추출할 수 있습니다.
- 버전 10.2.2 HotFix 1에서는 Oracle 및 SQL Server의 저장 프로시저에 대한 데이터 연계를 열 수준에서 추출할 수 있습니다.
- 버전 10.2.2 HotFix 1에서는 카탈로그에서 데이터 검색을 완료한 후 데이터 프로비저닝을 수행할 수 있습니다. 데이터 프로비저닝은 데이터를 대상으로 이동해 추가 분석을 수행하는 데 유용합니다.  
데이터 미리 보기에 대한 자세한 내용은 *Informatica 10.2.2 HotFix 1 Catalog 관리자 가이드* 및 *Informatica 10.2.2 Hotfix 1 Enterprise Data Catalog 사용자 가이드*를 참조하십시오.
- 버전 10.2.2 HotFix 1에서는 데이터를 대상으로 이동하기 전에 데이터를 미리 보고 평가할 수 있습니다. Oracle 및 Microsoft SQL Server 리소스의 테이블 형식 자산에 대한 데이터만 미리 볼 수 있습니다.  
데이터 미리 보기에 대한 자세한 내용은 *Informatica 10.2.2 HotFix 1 Catalog 관리자 가이드* 및 *Informatica 10.2.2 Hotfix 1 Enterprise Data Catalog 사용자 가이드*를 참조하십시오.

## 제 3 장

# 새로운 기능(10.2.2 서비스 팩 1)

이 장에 포함된 항목:

- [Big Data Management, 40](#)
- [Big Data 스트리밍, 41](#)
- [Enterprise Data Catalog, 41](#)
- [Enterprise Data Preparation, 42](#)
- [PowerExchange Adapters for Informatica, 42](#)

## Big Data Management

이 섹션에서는 버전 10.2.2 서비스 팩 1의 새로운 Big Data Management 기능에 대해 설명합니다.

### Sqoop

버전 10.2.2 서비스 팩 1에서는 다음과 같은 새로운 Sqoop 기능을 사용할 수 있습니다.

#### SSL 활성화 Oracle 및 Microsoft SQL Server 데이터베이스에 연결

Oracle 및 Microsoft SQL Server와 같은 보안 관계형 데이터베이스에 연결하도록 Sqoop을 구성할 수 있습니다.

SSL 활성화 Oracle 또는 Microsoft SQL Server 데이터베이스에 연결하려면 JDBC 연결에서 적절한 JDBC 연결 문자열과 connect 인수를 사용하면 됩니다.

자세한 내용은 *Informatica Big Data Management 10.2.2 서비스 팩 1 사용자 가이드*를 참조하십시오.

#### connection-param-file Sqoop 인수 지원

connection-param-file 인수를 정의하여 SSL 활성화 Oracle 데이터베이스에 연결할 수 있습니다. connection-param-file은 Sqoop이 SSL 활성화 데이터베이스에 연결하기 위해 사용해야 하는 추가 JDBC 매개 변수를 지정하는 속성 파일입니다.

이 파일의 콘텐츠는 표준 Java 속성으로 구문 분석되어 사용자가 연결을 생성할 때 드라이버로 전달됩니다.

JDBC 연결에서 **Sqoop** 인수 필드에 connection-param-file 인수를 지정할 수 있습니다.

다음 구문을 사용합니다.

```
--connection-param-file <parameter_file_name>
```

자세한 내용은 *Informatica Big Data Management 10.2.2 서비스 팩 1 사용자 가이드*를 참조하십시오.



# Big Data 스트리밍

이 섹션에서는 버전 10.2.2 서비스 팩 1의 새로운 Big Data Streaming 기능에 대해 설명합니다.

## Amazon S3 대상

버전 10.2.2 서비스 팩 1부터 Amazon S3에 데이터를 쓰는 스트리밍 매핑을 생성할 수 있습니다.

Amazon S3에 데이터를 쓰는 Amazon S3 데이터 개체를 생성하십시오. Amazon S3를 대상으로 사용하는 Amazon S3 연결을 생성할 수 있습니다. Developer tool 또는 infacmd를 사용하여 Amazon S3 연결을 생성하고 관리할 수 있습니다.

자세한 내용은 *Informatica Big Data Streaming 10.2.2 서비스 팩 1 사용자 가이드*를 참조하십시오.

## TIME\_RANGE 함수

버전 10.2.2 서비스 팩 1부터 조인할 스트리밍 이벤트의 시간 범위를 결정하는 TIME\_RANGE 함수를 조이너 변환에 사용할 수 있습니다.

TIME\_RANGE 함수는 스트리밍 매핑의 조이너 변환에만 적용할 수 있습니다.

### 구문

TIME\_RANGE(EventTime1,EventTime2,Format,Interval)

TIME\_RANGE 함수에 대한 자세한 내용은 *Informatica 10.2.2 서비스 팩 1 변환 언어 참조 가이드*를 참조하십시오.

## 여러 HDFS 파일에 데이터 기록 가능

버전 10.2.2 서비스 팩 1부터 복합 파일 데이터 개체를 사용하여 소스 데이터를 여러 HDFS 파일에 쓸 수 있습니다. 데이터 통합 서비스는 복합 파일 데이터 개체의 FileName 헤더 포트 값을 기반으로 런타임 시 대상 위치에 여러 HDFS 파일을 생성합니다.

자세한 내용은 *Informatica Big Data Streaming 10.2.2 서비스 팩 1 사용자 가이드*를 참조하십시오.

# Enterprise Data Catalog

이 섹션에서는 버전 10.2.2 서비스 팩 1의 새로운 Enterprise Data Catalog 기능에 대해 설명합니다.

## 독립 실행형 스캐너 유틸리티의 지원되는 리소스 유형

버전 10.2.2 서비스 팩 1부터 런타임이나 오프라인에서 액세스할 수 없는 경우 다음 외부 소스에서 메타데이터를 추출할 수 있습니다.

- IBM DB2
- IBM DB2 for z/OS
- IBM Netezza
- JDBC
- PowerCenter

- SQL Server Integration Services

자세한 내용은 *Informatica 10.2.2 서비스 팩 1 Enterprise Data Catalog 관리자 가이드*의 "오프라인 및 액세스할 수 없는 리소스의 메타데이터 추출" 장을 참조하십시오.

## Enterprise Data Preparation

이 섹션에서는 버전 10.2.2 서비스 팩 1의 새로운 Enterprise Data Preparation 기능에 대해 설명합니다.

### 모든 유추된 데이터 유형 되돌리기

버전 10.2.2 서비스 팩 1부터 워크시트의 소스 열에 적용된 모든 유추된 유형 및 데이터 도메인을 원래 유형으로 되돌릴 수 있습니다.

자세한 내용은 *Informatica 10.2.2 서비스 팩 1 Enterprise Data Preparation 사용자 가이드*에서 "데이터 준비" 장을 참조하십시오.

### Avro 및 Parquet 파일 준비

버전 10.2.2 서비스 팩 1에서는 데이터 준비의 첫 번째 단계로 프로젝트에 추가하는 Avro 및 Parquet 파일의 계층 데이터를 샘플링할 수 있습니다. Enterprise Data Preparation은 Avro 또는 Parquet 파일 구조를 플랫 구조로 변환한 후 데이터 준비에 사용할 수 있는 워크시트로 데이터를 제공합니다.

자세한 내용은 *Informatica 10.2.2 서비스 팩 1 Enterprise Data Preparation 사용자 가이드*에서 "데이터 준비" 장을 참조하십시오.

## PowerExchange Adapters for Informatica

이 섹션에서는 버전 10.2.2 서비스 팩 1의 새로운 Informatica 어댑터 기능에 대해 설명합니다.

### PowerExchange for Hive

버전 10.2.2 서비스 팩 1부터 Hive 테이블을 가져올 때 **기본 스키마만 표시** 옵션을 사용하여 검색 결과를 지우고 기본 스키마를 사용하는 테이블만 표시하거나 기본 스키마가 아닌 다른 스키마를 사용하는 테이블을 검색할 수 있습니다.

자세한 내용은 *Informatica PowerExchange for Hive 10.2.2 서비스 팩 1 사용자 가이드*를 참조하십시오.

## 제 4 장

# 새로운 기능(10.2.2)

이 장에 포함된 항목:

- [응용 프로그램 서비스, 43](#)
- [Big Data Management, 44](#)
- [Big Data 스트리밍, 48](#)
- [명령줄 프로그램, 50](#)
- [Enterprise Data Catalog, 54](#)
- [Enterprise Data Lake, 57](#)
- [Informatica Developer, 62](#)
- [Informatica 매핑, 62](#)
- [Informatica 변환, 63](#)
- [PowerExchange Adapters for Informatica, 66](#)

## 응용 프로그램 서비스

이 섹션에서는 버전 10.2.2의 새 응용 프로그램 서비스 기능에 대해 설명합니다.

### 대량 수집 서비스

버전 10.2.2에서는 대량 수집 서비스 프로세스와 외부 구성 요소 사이에 보안 통신을 활성화할 수 있습니다. HTTPS 프로토콜에 대한 HTTPS 포트 번호와 SSL 인증서에 대한 키 저장소 파일을 지정할 수 있습니다.

자세한 내용은 *Informatica 10.2.2 응용 프로그램 서비스 가이드*에서 "대량 수집 서비스" 장을 참조하십시오.

### 메타데이터 액세스 서비스

버전 10.2.2에서는 운영 체제 프로필을 사용하도록 메타데이터 액세스 서비스를 구성할 수 있습니다. 운영 체제 프로필을 사용하여 보안을 개선하고 Hadoop 클러스터에서 메타데이터를 가져오고 미리 볼 때 디자인 타임 사용자 환경을 격리할 수 있습니다.

자세한 내용은 *Informatica 10.2.2 보안 가이드*에서 "사용자 및 그룹" 장을 참조하십시오.

## REST 작업 헵 서비스

버전 10.2.2에서는 REST 응용 프로그램에 대한 REST 작업 헵 서비스를 구성할 수 있습니다. REST 작업 헵 서비스는 Informatica 도메인의 REST 시스템 서비스로, REST API를 통해 Informatica 제품 기능을 외부 클라이언트에 표시합니다.

Administrator 도구 또는 infacmd를 통해 REST 작업 헵 서비스를 구성할 수 있습니다. REST 작업 헵 서비스를 사용하여 응용 프로그램에서 배포된 매핑 작업에 대한 매핑 실행 통계를 볼 수 있습니다.

REST 작업 헵 서비스를 사용하여 데이터 통합 서비스 또는 Hadoop 환경에서 실행되는 빅 데이터 매핑에 대한 매핑 실행 통계를 가져올 수 있습니다.

REST API에 대한 자세한 내용은 *Big Data Management 10.2.2 관리자 가이드*를 참조하십시오.

## Big Data Management

이 섹션에서는 버전 10.2.2의 새로운 Big Data Management 기능에 대해 설명합니다.

### Azure Databricks 통합

버전 10.2.2에서는 you Informatica 도메인을 Azure Databricks 환경에 통합할 수 있습니다.

Azure Databricks는 Microsoft Azure 클라우드 서비스에 최적화된 분석 클라우드 플랫폼입니다. 여기에는 오픈 소스 Apache Spark 클러스터 기술과 기능이 통합되어 있습니다.

Informatica 도메인은 Azure VM 또는 온프레미스에 설치할 수 있습니다. 통합 프로세스는 Hadoop 환경과의 통합과 유사합니다. Databricks 환경에서 클러스터 구성을 가져오는 것을 비롯하여, 통합 태스크를 수행합니다. Informatica 도메인은 토큰 인증을 사용하여 Databricks 환경에 액세스합니다. Databricks 토큰 ID는 Databricks 연결에 저장됩니다.

#### 소스 및 대상

Databricks 환경에 포함된 다음과 같은 소스 및 대상에 대해 매핑을 실행할 수 있습니다.

- Microsoft Azure Data Lake Store
- Microsoft Azure Blob Storage
- Microsoft Azure SQL 데이터 웨어하우스
- Microsoft Azure Cosmos DB

#### 변환

다음 변환을 Databricks 매핑에 추가할 수 있습니다.

- 집계
- 식
- 필터
- 조이너
- 조회
- 노멀라이저
- 순위
- 라우터

분류기

합집합

Databricks Spark 엔진은 Hadoop 환경에서 Spark 엔진이 변환을 처리하는 방식과 거의 같은 방식으로 변환을 처리합니다.

## 데이터 유형

다음과 같은 데이터 유형이 지원됩니다.

배열

Bigint

날짜/시간

10진수

배정밀도

정수

맵

구조

텍스트

문자열

## 매핑

매핑을 구성할 때 Databricks 환경에서 매핑의 유효성을 검사하고 매핑을 실행하도록 선택할 수 있습니다. 매핑을 실행하면 데이터 통합 서비스가 Scala 코드를 생성하고 Databricks Spark 엔진으로 코드를 전달합니다.

## 워크플로우

클러스터 워크플로우를 개발하여 Databricks 환경에서 임시 클러스터를 생성할 수 있습니다.

자세한 내용은 다음 가이드를 참조하십시오.

*Big Data Management 10.2.2 통합 가이드*

*Big Data Management 10.2.2 관리자 가이드*

*Big Data Management 10.2.2 사용자 가이드*

## 계층 데이터

이 섹션에서는 버전 10.2.2에 포함된 계층 데이터에 대한 새로운 기능에 대해 설명합니다.

### 동적 복합 포트

버전 10.2.2에서는 Spark 엔진에서 실행되는 동적 매핑에 동적 복합 포트를 추가할 수 있습니다. 동적 복합 포트를 사용하여 복합 파일에서 계층 데이터의 잦은 스키마 변경을 관리합니다.

동적 복합 포트는 런타임 시 스키마 변경 내용을 기반으로 복합 포트의 새로운 요소나 변경된 요소를 수신합니다. 입력 규칙은 동적 복합 포트의 요소를 결정합니다. 입력 규칙을 기반으로 동적 복합 포트는 업스트림 변환으로부터 복합 포트의 요소 하나 이상을 수신합니다. Spark 엔진의 일부 변환에서 동적 배열, 동적 맵 및 동적 구조와 같은 동적 복합 포트를 사용할 수 있습니다.

자세한 내용은 *Informatica Big Data Management 10.2.2 사용자 가이드*에서 "스키마 변경 내용으로 계층 데이터 처리" 장을 참조하십시오.

## 계층 데이터 미리 보기

버전 10.2.2에서는 Developer tool의 매핑 내에서 계층 데이터를 미리 볼 수 있습니다. 계층 데이터를 미리 보면 빅 데이터 매핑을 디자인하고 디버깅하는 데 도움이 됩니다.

다음 계층 유형이 포함된 매핑에서 소스 및 변환을 미리 보기 지점으로 선택할 수 있습니다.

- 배열
- 구조
- 맵

계층 데이터 미리 보기는 기술 미리 보기에서 사용할 수 있습니다. 기술 미리 보기 기능은 지원되지만 보장되지 않으며 프로덕션용으로 준비되지 않았습니다. 이러한 기능은 비프로덕션 환경에서만 사용하는 것이 좋습니다.

자세한 내용은 *Informatica Big Data Management 10.2.2 사용자 가이드*를 참조하십시오.

## 고가용성

이 섹션에서는 버전 10.2.2의 새로운 고가용성 기능에 대해 설명합니다.

### Big Data 작업 복구

버전 10.2.2에서는 데이터 통합 서비스 노드가 예기치 않게 중지된 경우 데이터 통합 서비스가 Spark 엔진에서 실행되도록 구성된 빅 데이터 작업을 복구할 수 있습니다. 작업이 완료되기 전에 데이터 통합 서비스 노드에 장애가 발생하면 데이터 통합 서비스가 작업을 다른 노드로 전송하고 거기서 노드 장애가 발생한 지점부터 작업 태스크의 처리가 다시 시작됩니다.

빅 데이터 매핑을 복구하려면 데이터 통합 서비스 속성에서 빅 데이터 작업 복구를 활성화하고 `infacmd`에서 작업을 실행해야 합니다.

자세한 내용은 *Informatica Big Data Management 10.2.2 관리자 가이드*에서 "데이터 통합 서비스 처리" 장을 참조하십시오.

### 분산 데이터 통합 서비스 대기열

버전 10.2.2에서는 배포된 빅 데이터 작업에 빅 데이터 복구가 활성화된 경우 데이터 통합 서비스가 분산 대기열을 사용하여 작업 정보를 저장합니다. 분산 대기열은 모델 리포지토리에 저장되고, 리소스를 사용할 수 있게 되면 모든 사용 가능한 데이터 통합 서비스가 대기열에서 작업을 실행할 수 있습니다.

자세한 내용은 *Informatica Big Data Management 10.2.2 관리자 가이드*에서 "데이터 통합 서비스 처리" 장을 참조하십시오.

## 지능형 구조 모델

이 섹션에서는 버전 10.2.2의 새로운 지능형 구조 모델 기능에 대해 설명합니다.

### XML 파일의 별칭

버전 10.2.2에서는 Intelligent Structure Discovery가 지능형 구조 모델을 생성하는 데 사용된 XML 파일에 사용되는 것과 동일한 네임스페이스를 다른 별칭을 사용하여 식별하는 XML 파일을 처리할 수 있습니다.

## 데이터 유형

버전 10.2.2와 Informatica Intelligent Cloud Services의 Winter 2019 3월 릴리스부터 복합 파일 관독기가 지능형 구조 모델을 사용할 경우 Intelligent Structure Discovery는 데이터 유형을 출력 데이터 포트에 전달합니다.

예를 들어 Intelligent Structure Discovery가 날짜가 포함된 필드를 감지한 경우 데이터를 출력 데이터 포트에 문자열이 아닌 날짜로 전달합니다.

## 필드 이름

버전 10.2.2와 Informatica Intelligent Cloud Services의 Winter 2019 3월 릴리스부터 지능형 구조 모델에서 가져오는 복합 파일 데이터 개체의 필드 이름이 숫자 및 예약어로 시작될 수 있으며 다음 특수 문자를 포함할 수 있습니다. |. []{}() \*+ - ? . ^ \$ /

필드 이름이 숫자 또는 예약어로 시작되는 경우 Big Data Management 매핑은 필드 이름 시작 부분에 밑줄(\_)을 추가합니다. 예를 들어 지능형 구조 모델의 필드 이름이 OR로 시작되는 경우 매핑에서는 필드를 \_OR로 가져옵니다. 필드 이름이 특수 문자를 포함하는 경우 매핑에서는 특수 문자를 밑줄로 변환합니다.

## 대형 XML 파일 처리

버전 10.2.2에서는 Intelligent Structure Discovery가 XML 파일을 스트리밍하고 청크의 반복 요소에 대한 데이터를 처리할 수 있습니다. 이렇게 하여 대형 XML 파일을 보다 효율적으로 처리합니다.

## 데이터 드리프트

버전 10.2.2와 Informatica Intelligent Cloud Services의 Winter 2019 3월 릴리스부터 Intelligent Structure Discovery의 데이터 드리프트 처리가 향상됩니다.

Intelligent Structure Discovery에서 입력 데이터가 샘플 파일에 포함되지 않은 필드를 포함하는 경우 데이터 드리프트가 발생합니다. 이 경우 Intelligent Structure Discovery는 정의되지 않은 데이터를 무시하는 것이 아니라 대상의 할당되지 않은 데이터 포트에 데이터를 전달합니다.

## 대량 수집

버전 10.2.2에서는 증분 로드를 실행하여 증분 데이터를 수집할 수 있습니다. 증분 로드를 실행하면 Spark 엔진이 타임스탬프 또는 ID 열을 기반으로 증분 데이터를 가져오고 Hive 또는 HDFS 대상에 증분 데이터를 로드합니다. 데이터를 Hive 대상에 수집하는 경우 Spark 엔진은 소스 테이블에 수행된 스키마 변경 내용도 전파할 수 있습니다.

증분 데이터를 수집하는 경우 대량 수집 서비스는 Sqoop의 증분 가져오기 모드를 활용합니다.

자세한 내용은 *Informatica Big Data Management 10.2.2 대량 수집 가이드*를 참조하십시오.

## 모니터링

이 섹션에서는 버전 10.2.2의 Big Data Management 모니터링과 관련된 새로운 기능에 대해 설명합니다.

### Spark 모니터링

버전 10.2.2에서는 Spark 모니터링에 대한 요약 통계 창에서 사전 작업 태스크와 사후 작업 태스크를 모두 볼 수 있습니다.

사전 작업 및 사후 작업 태스크에 대한 자세한 내용은 *Informatica Big Data Management 10.2.2 사용자 가이드*를 참조하십시오.

## 보안

이 섹션에서는 버전 10.2.2의 Big Data Management 보안과 관련된 새로운 기능에 대해 설명합니다.

### 엔터프라이즈 보안 패키지

버전 10.2.2에서는 엔터프라이즈 보안 패키지가 포함된 Azure HDInsight 클러스터가 지원됩니다.

엔터프라이즈 보안 패키지는 인증에 Kerberos를 사용하고 권한 부여에 Apache Ranger를 사용합니다.

엔터프라이즈 보안 패키지에 대한 자세한 내용은 *Informatica Big Data Management 10.2.2 관리자 가이드*를 참조하십시오.

## 대상

이 섹션에서는 버전 10.2.2에 포함된 대상의 새로운 기능에 대해 설명합니다.

### HDFS 플랫 파일 대상

버전 10.2.2에서는 출력 데이터를 HDFS 대상 파일 및 거부 파일에 추가할 수 있습니다. 출력 데이터를 추가하려면 HDFS 대상이 존재하는 경우 데이터를 추가하도록 선택합니다.

추가된 데이터를 포함하는 파일을 쉽게 관리할 수 있도록 데이터 통합 서비스는 대상 파일 및 거부 파일의 이름에 매핑 실행 ID를 추가합니다.

자세한 내용은 *Informatica Big Data Management 10.2.2 사용자 가이드*에서 "대상" 장을 참조하십시오.

## Big Data 스트리밍

이 섹션에서는 버전 10.2.2의 새로운 Big Data Streaming 기능에 대해 설명합니다.

### Azure 이벤트 헵 데이터 개체

버전 10.2.2에서는 다음과 같은 배포에서 소스로 이벤트 헵이 있는 스트리밍 매핑을 배포할 수 있습니다.

- Amazon EMR
- ADLS 저장소를 포함하는 Azure HDInsight
- Cloudera CDH
- Hortonworks HDP

### Amazon Kinesis 연결의 교차 계정 IAM 역할

버전 10.2.2에서는 교차 계정 IAM 역할을 사용하여 Amazon Kinesis 소스를 인증할 수 있습니다.

교차 계정 IAM 역할을 사용하면 한 AWS 계정의 리소스를 다른 AWS 계정의 사용자와 공유할 수 있으며, 각 계정에서 사용자를 생성할 필요가 없습니다.

자세한 내용은 *Informatica Big Data Streaming 10.2.2 사용자 가이드*를 참조하십시오.



## 지능형 구조 모델

버전 10.2.2에서는 Big Data Streaming에서 지능형 구조 모델을 사용할 수 있습니다.

Kafka, Kinesis 또는 Azure 이벤트 헵 데이터 개체에서 지능형 구조 모델을 통합할 수 있습니다. 매핑에 데이터 개체를 추가할 때 모델로 구문 분석할 수 있는 모든 입력 유형을 처리할 수 있습니다.

데이터 개체는 입력을 허용하며 모델을 생성할 때 사용된 파일에 따라 PDF 형태, JSON, Microsoft Excel, Microsoft Word 테이블, CSV, 텍스트 또는 XML 입력 파일을 구문 분석할 수 있습니다.

자세한 내용은 *Informatica Big Data Streaming 10.2.2 사용자 가이드*를 참조하십시오.

## Big Data Streaming 데이터 개체에 대한 헤더 포트

버전 10.2.2에서는 일부 데이터 개체가 이벤트와 연결된 메타데이터를 나타내는 기본 헤더 포트를 포함합니다. 예를 들어 타임스탬프 포트에는 이벤트가 생성된 시간이 포함됩니다. 헤더 포트를 사용하여 데이터를 그룹화하고 처리할 수 있습니다.

헤더 포트에 대한 자세한 내용은 *Informatica Big Data Streaming 10.2.2 사용자 가이드*를 참조하십시오.

## Amazon Kinesis 연결의 AWS 자격 증명 프로필

버전 10.2.2에서는 Amazon Kinesis 연결 시 인증에 기반한 AWS 자격 증명 프로필을 사용할 수 있습니다.

Amazon Kinesis 연결을 생성할 때 AWS 자격 증명 프로필 이름을 입력할 수 있습니다. 런타임 동안 매핑은 AWS 자격 증명 파일에 나열된 프로필 이름을 통해 AWS 자격 증명에 액세스합니다.

자세한 내용은 *Informatica Big Data Streaming 10.2.2 사용자 가이드*를 참조하십시오.

## Spark 구조화된 스트리밍

버전 10.2.2에서는 Big Data Streaming은 Spark 구조화된 스트리밍을 사용하여 스트리밍 데이터를 처리합니다.

Spark 구조화된 스트리밍은 Spark 엔진에 내장된 확장 가능한 내결함성 오픈 소스 스트리밍 처리 엔진입니다. 이 엔진은 스트리밍 이벤트의 지연 수신을 허용하고 소스 타임스탬프를 기반으로 스트리밍 데이터를 처리할 수 있습니다.

Spark 엔진은 스트리밍 매핑을 지속적으로 실행합니다. 이 엔진은 데이터를 읽고, 데이터를 마이크로 배치로 나누고, 마이크로 배치를 처리하고, 결과를 게시한 다음 대상에 씁니다.

자세한 내용은 *Informatica Big Data Streaming 10.2.2 사용자 가이드*를 참조하십시오.

## 창 변환

버전 10.2.2에서는 창 변환을 생성할 때 다음과 같은 기능을 사용할 수 있습니다.

### 워터마크 지연

워터마크 지연은 지연된 이벤트가 데이터 그룹에 누적되는 임계값 시간을 정의합니다.

워터마크 지연은 지연되어 수신된 데이터를 그룹화하여 처리할 수 있는 기간을 지정할 수 있는 임계값입니다. 이벤트 데이터가 임계값 시간 내에 수신된 경우 데이터가 처리되고 데이터가 해당하는 데이터 그룹에 누적됩니다.

## 창 포트

창 포트는 이벤트를 그룹화하는 데 기준으로 사용할 수 있는 타임스탬프 값이 포함된 열을 지정합니다. 누적된 데이터는 타임스탬프 값을 포함합니다. 창 포트 열을 사용하여 지연 수신된 이벤트 시간 데이터를 그룹화합니다.

자세한 내용은 *Informatica Big Data Streaming 10.2.2 사용자 가이드*를 참조하십시오.

# 명령줄 프로그램

이 섹션에서는 버전 10.2.2의 새로운 명령에 대해 설명합니다.

## infacmd dis 명령

이 항목에서는 새로운 infacmd dis 명령 옵션 및 실행 옵션에 대해 설명합니다.

### 서비스 옵션 업데이트

서비스 옵션 업데이트 명령은 데이터 통합 서비스 속성을 업데이트합니다.

다음 테이블에는 새로운 infacmd dis updateServiceOptions 명령 옵션이 설명되어 있습니다.

새 옵션	설명
-RecycleMode rm	선택 사항입니다. 재사용 모드는 서비스를 다시 시작하고 최신 서비스 및 서비스 프로세스 속성을 적용합니다. 중단 또는 완료 선택합니다. <ul style="list-style-type: none"><li>- 완료. 모든 응용 프로그램을 중지하고 각 응용 프로그램의 모든 작업을 취소합니다. 서비스를 비활성화하기 전에 모든 작업이 취소될 때까지 기다리십시오.</li><li>- 중단. 모든 응용 프로그램을 중지하고 모든 작업을 취소한 후 작업을 중단하고 서비스를 비활성화하십시오.</li></ul> 기본값: 완료

다음 테이블에는 새로운 infacmd dis updateServiceOptions 명령 실행 옵션이 설명되어 있습니다.

새 옵션	설명
ExecutionOptions.BigDataJobRecovery	빅 데이터 작업 복구를 활성화합니다. "true" 또는 "false"로 설정합니다. 기본값은 false입니다.

자세한 내용은 *Informatica 10.2.2 명령 참조*에서 "infacmd dis 명령 참조" 장을 참조하십시오.

## infacmd ihs 명령

다음 테이블에는 새로운 infacmd ihs 명령 옵션이 설명되어 있습니다.

새 옵션	설명
-PrimaryNode -nm	선택 사항입니다. 서비스가 실행되는 기본 노드입니다.
-BackupNodes -bn	선택 사항입니다. 기본 노드를 사용할 수 없는 경우 서비스를 실행할 수 있는 노드입니다.고가용성이 있는 경우 백업 노드를 구성할 수 있습니다.

다음 테이블에는 새로운 infacmd ihs 명령이 설명되어 있습니다.

명령	설명
cleanCluster	Informatica 클러스터 서비스를 정리합니다.

자세한 내용은 *Informatica 10.2.2 명령 참조*에서 "infacmd ihs 명령 참조" 장을 참조하십시오.

## infacmd ipc 명령

다음 테이블에는 더 이상 사용되지 않는 infacmd ipc 명령이 설명되어 있습니다.

명령	설명
ExportToPC	모델 리포지토리 또는 내보내기 파일에서 개체를 내보내고 PowerCenter 개체로 변환합니다.

## infacmd ldm 명령

다음 테이블에는 새로운 infacmd ldm 명령 옵션이 설명되어 있습니다.

새 옵션	설명
-PrimaryNode -nm	선택 사항입니다. Enterprise Data Catalog에 대해 고가용성을 구성하려는 경우 기본 노드 이름을 지정합니다.
-BackupNodes -bn	선택 사항입니다. Enterprise Data Catalog에 대해 고가용성을 구성하려는 경우 씬표로 구분된 백업 노드 이름의 목록을 지정합니다.
-isNotifyChangeEmailEnabled -cne	선택 사항입니다. 자산 변경 알림을 활성화하려면 True를 지정합니다. 기본값은 False입니다.

새 옵션	설명
-ExtraJarsPath -ejp	선택 사항입니다. Informatica 도메인을 설치한 시스템의 디렉터리 경로입니다. 이 디렉터리에는 WANDisco Fusion이 있는 기존 클러스터에 Enterprise Data Catalog를 배포하는 데 필요한 JAR 파일이 포함되어야 합니다.
-ExtraJarsPath -ejp	선택 사항입니다. Informatica 도메인을 설치한 시스템의 디렉터리 경로입니다. 이 디렉터리에는 WANDisco Fusion이 있는 기존 클러스터에 Enterprise Data Catalog를 배포하는 데 필요한 JAR 파일이 포함되어야 합니다.

다음 테이블에는 새로운 `infacmd ldm` 명령이 설명되어 있습니다.

명령	설명
<code>collectAppLogs</code>	카탈로그 서비스를 활성화하기 위해 실행되는 YARN 응용 프로그램의 로그 파일을 수집합니다.
<code>publishArchive</code>	오프라인 모드에서 리소스를 생성하고 스캔을 실행합니다.

자세한 내용은 *Informatica 10.2.2 명령 참조*에서 "infacmd ldm 명령 참조" 장을 참조하십시오.

## infacmd mi 명령

다음 테이블에는 `infacmd mi` 명령에 대한 변경 내용이 설명되어 있습니다.

명령	변경 설명
<code>createService</code>	버전 10.2.2에서는 -HttpsPort, -KeystoreFile 및 -KeystorePassword 옵션을 사용하여 대량 수집 서비스 프로세스가 보안 연결을 사용하여 외부 구성 요소와 통신할지 여부를 지정합니다.
<code>extendedRunStats</code>	버전 10.2.2에서는 소스 테이블에 대한 확장된 실행 통계를 보려면 -RunID 옵션을 사용하여 대량 수집 사양의 RunID를 지정하고 -SourceName 옵션을 사용하여 소스 테이블의 이름을 지정해야 합니다. 소스 테이블이 증분 로드를 사용하여 수집된 경우 실행 통계에 증분 키와 시작 값이 표시됩니다. 이전에는 소스 테이블을 수집한 수집 매핑 작업에 대한 JobID를 지정했습니다. 10.2.2로 업그레이드한 경우 <code>infacmd mi extendedRunStats</code> 를 실행하는 모든 스크립트를 새 옵션을 사용하도록 업데이트해야 합니다.
<code>listSpecRuns</code>	버전 10.2.2에서는 이 명령이 Spark 엔진에서 대량 수집 사양을 실행하기 위해 사용하는 로드 유형을 추가적으로 반환합니다.
<code>runSpec</code>	버전 10.2.2에서는 -LoadType 옵션을 사용하여 대량 수집 사양을 실행할 로드 유형을 지정할 수 있습니다. 로드 유형은 전체 로드 또는 증분 로드일 수 있습니다.

자세한 내용은 *Informatica 10.2.2 명령 참조*에서 "infacmd mi 명령 참조" 장을 참조하십시오.

## infacmd ms 명령

다음 테이블에는 새로운 infacmd ms 명령이 설명되어 있습니다.

명령	설명
abortAllJobs	Spark 엔진에서 실행되도록 구성된 모든 배포된 매핑 작업을 중단합니다. 대기열에 있는 작업, 실행 중인 작업 또는 둘 모두를 중단하도록 선택할 수 있습니다.
createConfigurationWithParams	사용자가 명령줄에서 지정한 클러스터 매개 변수를 통해 클러스터 구성을 생성합니다.
listMappingOptions	응용 프로그램의 매핑 옵션을 나열합니다.
purgeDatabaseWorkTables	데이터 통합 서비스에 대한 빅 데이터 복구가 활성화된 경우 대기열에서 모든 작업 정보를 제거합니다.
updateMappingOptions	응용 프로그램의 매핑 옵션을 업데이트합니다.
updateOptimizationLevel	응용 프로그램의 여러 매핑에 대한 최적화 수준을 업데이트합니다.

자세한 내용은 *Informatica 명령 참조*에서 "infacmd ms 명령 참조" 장을 참조하십시오.

## infacmd oie 명령

버전 10.2.2에서는 oie 플러그 인은 더 이상 사용되지 않으며 oie 플러그 인에 대한 지원은 향후 릴리스에서 제거됩니다. infacmd oie 명령은 tools 플러그 인으로 마이그레이션되었습니다. 자세한 내용은 [“infacmd tools 명령” 페이지 53](#) 항목을 참조하십시오.

## infacmd tools 명령

버전 10.2.2에서는 tools infacmd 플러그 인이 oie 플러그 인을 대체합니다. tools 플러그 인은 oie 플러그 인이 수행하는 개체 가져오기 및 내보내기 작업을 수행하며, 이 플러그 인을 사용하여 몇 가지 추가 작업을 수행할 수 있습니다.

다음 테이블에는 tools 플러그 인으로 마이그레이션되는 infacmd oie 명령이 나열되어 있습니다.

이전 명령	현재 명령
infacmd oie deployApplication	infacmd tools deployApplication
infacmd oie exportObjects	infacmd tools exportObjects
infacmd oie exportResources	infacmd tools exportResources
infacmd oie importObjects	infacmd tools importObjects

다음 테이블에는 새로운 **infacmd tools** 명령이 설명되어 있습니다.

명령	설명
patchApplication	.piar 파일을 사용하여 응용 프로그램 패치를 데이터 통합 서비스에 배포합니다.

자세한 내용은 *Informatica 10.2.2 명령 참조*에서 "infacmd tools 명령 참조" 장을 참조하십시오.

## infasetup 명령

다음 테이블에는 변경된 **infasetup** 명령이 설명되어 있습니다.

명령	설명
DefineDomain	10.2.2에서는 -spid 옵션이 DefineDomain 명령에 추가되었습니다.
updateDomainSamlConfig	10.2.2에서는 -spid 옵션이 updateDomainSamlConfig 명령에 추가되었습니다.

자세한 내용은 *Informatica 10.2.2 명령 참조*에서 "infasetup 명령 참조" 장을 참조하십시오.

# Enterprise Data Catalog

이 섹션에서는 버전 10.2.2의 새로운 Enterprise Data Catalog 기능에 대해 설명합니다.

## 열에 비즈니스 제목 자동 할당

버전 10.2.2에서는 Enterprise Data Catalog가 열 자산의 비즈니스 제목으로 사용할 비즈니스 용어를 유추합니다. 신뢰도 점수가 80% 이상인 비즈니스 용어가 열의 비즈니스 제목으로 유추됩니다. 신뢰도 점수는 데이터 도메인의 수락된 비즈니스 용어, 열 이름 유사점 및 열과 비즈니스 용어 사이의 이름 유사점에 기반하는 내부 알고리즘에 의해 계산됩니다.

자세한 내용은 *Informatica 10.2.2 Enterprise Data Catalog 사용자 가이드*에서 "자산 태스크 수행" 장을 참조하십시오.

## 자산에 대한 사용자 공동 작업

버전 10.2.2에서는 다른 Enterprise Data Catalog 사용자와 자산에 대한 공동 작업을 수행할 수 있습니다. 자산에 대한 공동 작업을 통해 다른 사용자와 상호 작용하고, 자산에 대한 통찰력을 공유하고, 자산과 관련된 쿼리를 수행하고, 모든 자산 변경 내용을 팔로우하고, 자산을 인증할 수 있습니다.

다음과 같은 방법으로 자산에 대한 공동 작업을 수행할 수 있습니다.

### 자산 팔로우

자산을 팔로우하여 카탈로그의 자산 변경 내용을 모니터링할 수 있습니다. 자산을 팔로우하면 다른 사용자가 자산에 수행한 변경 내용과 관련된 알림을 받을 수 있으므로 자산을 모니터링하고 필요한 조치를 취할 수 있습니다.

## 자산 평가 및 검토

카탈로그에서 다섯 등급을 기반으로 자산을 평가하고 검토할 수 있습니다. 자산을 평가하고 검토하여 자산의 품질, 적용성, 유용성 및 가용성과 같은 자산의 여러 측면을 기반으로 자산에 대한 피드백을 제공하십시오.

## 자산 쿼리

카탈로그의 자산에 대해 보다 정확하게 이해하려는 경우 자산에 대해 질문할 수 있습니다. 설명적, 탐색적, 예측적 또는 인과론적인 질문을 하십시오.

## 자산 인증

자산을 인증하여 보증할 수 있습니다. 이렇게 하면 다른 사용자가 해당 자산을 인증되지 않은 자산과 달리 신뢰할 수 있는 자산으로 사용할 수 있습니다.

자세한 내용은 *Informatica 10.2.2 Enterprise Data Catalog 가이드*에서 "자산에 대한 사용자 공동 작업" 장을 참조하십시오.

# 설치 프로그램을 사용하여 Enterprise Data Catalog 응용 프로그램 서비스 생성

버전 10.2.2에서는 Enterprise Data Catalog를 설치한 후 설치 프로그램을 사용하여 Enterprise Data Catalog 응용 프로그램 서비스를 생성할 수 있습니다. 응용 프로그램 서비스를 생성하지 않고 Enterprise Data Catalog를 설치한 경우 설치 프로그램을 사용할 수 있습니다.

설치 프로그램을 사용하여 응용 프로그램 서비스를 생성하는 것에 대한 자세한 내용은 *Informatica Enterprise Data Catalog 10.2.2 설치 및 구성 가이드*를 참조하십시오.

# 사용자 지정 메타데이터 유효성 검사 유틸리티

버전 10.2.2에서는 **you can use a** 독립 실행형 Java 기반 유효성 검사 명령줄 유틸리티를 사용하여 카탈로그에 수집하려는 사용자 지정 메타데이터의 구문 및 의미의 유효성을 검사할 수 있습니다. 사용자 지정 메타데이터는 Enterprise Data Catalog가 리소스를 제공하지 않는 사용자 지정 데이터 소스에서 수집하려는 메타데이터를 나타냅니다.

유틸리티 사용에 대한 자세한 내용은 KB 문서 *방법: 카탈로그에 수집하기 전 사용자 지정 메타데이터의 유효성 검사*를 참조하십시오. 유틸리티를 다운로드하는 지침에 대해서는 Informatica 글로벌 고객 지원 센터에 문의하십시오.

# 변경 알림

버전 10.2.2에서는 팔로우하는 자산이 변경되면 Enterprise Data Catalog가 알림을 표시합니다. 알림 유형에는 응용 프로그램 알림, 변경 전자 메일 알림 및 요약 전자 메일 알림이 포함됩니다.

자세한 내용은 *Informatica 10.2.2 Enterprise Data Catalog 가이드*에서 "자산에 대한 사용자 공동 작업" 장을 참조하십시오.

# Business Glossary 할당 보고서

버전 10.2.2에서는 리소스에 대해 제안된 여러 비즈니스 용어를 수락하거나 거부할 수 있습니다.

자세한 내용은 *Informatica 10.2.2 Enterprise Data Catalog 가이드*에서 "자산 태스크 수행" 장을 참조하십시오.

## 운영 체제 프로필

버전 10.2.2에서는 기본 운영 체제 프로필이 없는 경우 운영 체제 프로필을 선택할 수 있습니다. 데이터 통합 서비스는 운영 체제 프로필 사용자 자격 증명을 사용하여 데이터 검색을 수행합니다. 데이터 검색에는 열 프로필 및 데이터 도메인 검색 프로필이 포함됩니다.

Enterprise Data Catalog의 운영 체제 프로필 사용에 대한 자세한 내용은 *Informatica 10.2.2 Catalog 관리자 가이드*에서 "Enterprise Data Catalog 개념" 장을 참조하십시오.

## REST API

버전 10.2.2에서는 다음과 같은 Informatica Enterprise Data Catalog REST API를 사용할 수 있습니다.

- 비즈니스 용어 REST API. 허용되거나, 유추되거나, 거부된 비즈니스 용어를 반환, 업데이트 또는 삭제할 수 있습니다.
- 카탈로그 이벤트 REST API. 사용자 구성, 전자 메일 구성 및 사용자 구독을 액세스, 업데이트 또는 삭제할 수 있습니다.
- 개체 인증 API. 개체에 대한 인증 속성을 나열, 업데이트 및 삭제할 수 있습니다.
- 개체 설명 API. 데이터 개체에 대한 설명, 회신 및 투표를 나열, 생성, 업데이트 및 삭제할 수 있습니다.
- 개체 검토 API. 검토, 검토에 대한 등급 및 투표를 나열, 생성, 업데이트 및 삭제할 수 있습니다.

REST API에 대한 자세한 내용은 *Informatica 10.2.2 Enterprise Data Catalog REST API 참조*를 참조하십시오.

## 소스 메타데이터 및 데이터 프로필 필터

버전 10.2.2에서는 소스 메타데이터 필터 및 데이터 프로필 필터를 사용하여 리소스 실행에서 소스 테이블 및 뷰를 지정할 수 있습니다. 이러한 필터를 사용할 경우 Enterprise Data Catalog는 특정 테이블 및 뷰에서 소스 메타데이터와 프로필 메타데이터를 추출합니다.

소스 메타데이터 및 데이터 프로필 필터에 대한 자세한 내용은 *Informatica 10.2.2 Catalog 관리자 가이드*에서 "리소스 관리" 장을 참조하십시오.

## 스캐너 유틸리티

버전 10.2.2에서는 Informatica가 오프라인 리소스와 액세스할 수 없는 리소스에서 메타데이터를 추출하는 데 사용할 수 있는 독립 실행형 스캐너 유틸리티를 제공합니다. 이 유틸리티에는 연관된 명령과 함께 순서대로 실행해야 하는 스크립트가 들어 있습니다.

독립 실행형 스캐너 유틸리티에 대한 자세한 내용은 *Informatica 10.2.2 Catalog 관리자 가이드*에서 "오프라인 및 액세스할 수 없는 리소스의 메타데이터 추출" 부록을 참조하십시오.

## 리소스 유형

버전 10.2.2에서는 다음과 같은 데이터 소스 유형에 대한 리소스를 생성할 수 있습니다.

### Google BigQuery

Google BigQuery 데이터 소스의 다음과 같은 자산에서 메타데이터, 관계 및 연계 정보를 추출할 수 있습니다.

- 프로젝트
- 데이터 집합
- 테이블



- 보기

Google BigQuery 데이터 소스 구성에 대한 자세한 내용은 *Informatica 10.2.2 Catalog 관리자 가이드*를 참조하십시오.

#### Workday

Workday 데이터 소스의 다음과 같은 자산에서 메타데이터, 관계 및 연계 정보를 추출할 수 있습니다.

- 서비스
- 항목
- 보고서
- 작업
- 데이터 소스
- 속성
- 비즈니스 개체

Workday 데이터 소스 구성에 대한 자세한 내용은 *Informatica 10.2.2 Catalog 관리자 가이드*를 참조하십시오.

## Enterprise Data Lake

이 섹션에서는 버전 10.2.2의 새로운 Enterprise Data Lake 기능에 대해 설명합니다.

### 활성 규칙 적용

버전 10.2.2에서는 프로젝트에서 활성 규칙을 사용할 수 있습니다.

활성 규칙은 **Developer tool**을 사용하여 개발된 맵셋입니다. 활성 규칙을 사용하여 집계 및 데이터 품질 변환과 같은 복합 변환을 워크시트에 일치 및 통합용으로 적용할 수 있습니다.

활성 규칙은 데이터 집합 내의 모든 행을 입력으로 사용합니다. 여러 워크시트를 선택하여 규칙에 대한 입력으로 사용할 수 있습니다. 응용 프로그램은 규칙 출력이 포함된 워크시트를 프로젝트에 추가합니다.

자세한 내용은 *Informatica 10.2.2 Enterprise Data Lake 사용자 가이드*에서 "데이터 준비" 장을 참조하십시오.

### 중복 행 삭제

버전 10.2.2에서는 중복 값을 포함하는 행을 워크시트에서 삭제할 수 있습니다.

자세한 내용은 *Informatica 10.2.2 Enterprise Data Lake 사용자 가이드*에서 "데이터 준비" 장을 참조하십시오.

### 열 데이터 클러스터 및 범주화

버전 10.2.2에서는 유사한 값을 한 열에 모은 다음 Enterprise Data Lake의 권장 사항을 기반으로 값을 범주화할 수 있습니다. 응용 프로그램에서는 음성 알고리즘을 사용하여 유사한 값을 모은 다음 자주 발생하지 않는 값을 자주 발생하는 값으로 대체할 수 있도록 제안합니다.

자세한 내용은 *Informatica 10.2.2 Enterprise Data Lake 사용자 가이드*에서 "데이터 준비" 장을 참조하십시오.

## CLAIRE 기반 권장 사항

버전 10.2.2에서는 데이터를 준비할 때 응용 프로그램이 임베디드 CLAIRE 기계 학습 검색 엔진을 사용하여 권장 사항을 제공합니다.

프로젝트 페이지를 보는 경우 응용 프로그램이 데이터 연계에 기반한 업스트림 데이터 소스와 문서화된 기본-외래 키 관계에서 파생된 대체 및 추가 권장 사항을 표시합니다.

데이터 준비 중에 워크시트에서 열을 선택하면 응용 프로그램이 열 데이터 유형을 기반으로 데이터를 개선할 수 있는 제안을 열 개요 패널에 표시합니다.

두 워크시트에 대한 조인 작업을 수행하면 응용 프로그램이 기본-외래 키 관계를 사용하여 원하는 키 쌍에 대한 낮은 겹침이 발생할 때 호환되지 않는 샘플링을 나타냅니다.

자세한 내용은 *Informatica 10.2.2 Enterprise Data Lake 사용자 가이드*에서 "데이터 준비" 장을 참조하십시오.

## 조건부 집계

10.2.2에서는 AND 및 OR 논리를 사용하여 프로젝트에서 집계 워크시트를 생성할 때 사용하는 IF 계산에 여러 조건을 적용할 수 있습니다.

- 한 조건에 둘 이상의 열을 포함하려면 모든 연산자와 함께 AND를 사용합니다.
- 한 조건의 열 내에 둘 이상의 값을 포함하려면 IS, IS NOT 및 IS BETWEEN 연산자와 함께 OR를 사용합니다.

자세한 내용은 *Informatica 10.2.2 Enterprise Data Lake 사용자 가이드*에서 "데이터 준비" 장을 참조하십시오.

## 데이터 마스킹

버전 10.2.2에서는 Enterprise Data Lake가 데이터 보안 제품인 Informatica Dynamic Data Masking과 통합되어 데이터 자산에서 중요한 데이터의 마스킹을 활성화할 수 있습니다.

Enterprise Data Lake에서 데이터 마스킹을 활성화하려면 Dynamic Data Masking 서버를 구성하여 데이터 레이크의 데이터 자산에 마스킹 규칙을 적용합니다. 또한 Informatica 도메인을 구성하여 Dynamic Data Masking 서버에 연결하도록 Enterprise Data Lake를 활성화합니다.

Dynamic Data Masking은 Enterprise Data Lake에서 데이터 레이크로 전송된 요청을 가로채고 요청된 자산의 열에 마스킹 규칙을 적용합니다. Enterprise Data Lake 사용자가 마스킹된 데이터가 포함된 열을 보거나 관련 작업을 수행하는 경우 실제 데이터는 적용된 마스킹 규칙에 따라 전체 또는 일부가 가려집니다.

자세한 내용은 *Informatica 10.2.2 Enterprise Data Lake 관리자 가이드*에서 "중요한 데이터 마스킹" 장을 참조하십시오.

## 지역화

버전 10.2.2에서는 사용자 인터페이스가 일본어를 지원합니다. 또한 프로젝트 이름 및 설명에서 비라틴어 문자를 사용할 수 있습니다.

## 분할된 소스 및 대상

버전 10.2.2에서는 Enterprise Data Lake가 가져오기, 게시 또는 복사 작업 중에 분할된 소스에서 데이터를 읽을 수 있습니다. 또한 응용 프로그램이 가져오기, 게시, 복사 또는 업로드 작업 중에 데이터 레이크에서 분할된 대상에 데이터를 추가할 수 있습니다.

## 레시피 단계에 설명 추가

버전 10.2.2에서는 레시피 단계에 설명을 추가할 수 있습니다. 설명을 사용하여 공동 작업을 개선하고 감사 요구 사항을 충족하는 세부 정보를 제공할 수 있습니다.

자세한 내용은 *Informatica 10.2.2 Enterprise Data Lake 사용자 가이드*에서 "데이터 준비" 장을 참조하십시오.

## 레시피를 매핑으로 저장

버전 10.2.2에서는 레시피를 게시하고 새 출력 테이블을 생성하는 대신 레시피를 매핑으로 저장할 수 있습니다.

매핑을 Enterprise Data Lake 서비스와 연결된 모델 리포지토리에 저장하거나 매핑을 .xml 파일로 저장할 수 있습니다. 개발자는 Developer tool을 사용하여 매핑을 검토하고 수정한 다음 시스템 리소스 가용성을 기반으로 적절한 시기에 매핑을 실행할 수 있습니다.

자세한 내용은 *Informatica 10.2.2 Enterprise Data Lake 사용자 가이드*에서 "데이터 준비" 장을 참조하십시오.

## Amazon S3, ADLS, WASB, MapR-FS를 데이터 소스로

버전 10.2.2에서는 다음과 같은 데이터 소스에 저장된 파일의 데이터를 준비할 수 있습니다.

- Amazon S3
- MapR-FS
- Microsoft Azure Data Lake Storage
- Windows Azure Storage Blob

준비하려는 데이터를 포함하는 각 데이터 소스에 대해 Enterprise Data Catalog에서 리소스를 생성해야 합니다. 리소스는 외부 데이터 소스 또는 메타데이터 리포지토리를 나타내는 리포지토리 개체입니다. 리소스에 연결된 스캐너가 리소스에서 메타데이터를 추출하고 메타데이터를 Enterprise Data Catalog에 저장합니다.

Enterprise Data Catalog에서 리소스를 생성하는 것에 대한 자세한 내용은 *Informatica 10.2.2 Catalog 관리자 가이드*에서 "리소스 관리" 장을 참조하십시오.

## 통계 함수

버전 10.2.2에서는 데이터를 준비할 때 다음과 같은 통계 함수를 워크시트의 열에 적용할 수 있습니다.

- AVG
- AVGIF
- COUNT
- COUNTIF
- COUNTDISTINCT
- COUNTDISTINCTIF
- MAX
- MAXIF
- MIN
- MINIF
- STDDEV
- STDDEVIF
- SUM

- SUMIF
- VARIANCE
- VARIANCEIF

자세한 내용은 *Informatica 10.2.2 Enterprise Data Lake 사용자 가이드*에서 "데이터 준비" 장을 참조하십시오.

## 날짜 및 시간 함수

버전 10.2.2에서는 데이터를 준비할 때 다음과 같은 날짜 및 시간 함수를 워크시트의 열에 적용할 수 있습니다.

- ADD\_TO\_DATE
- CURRENT\_DATETIME
- DATETIME
- DATE\_DIFF
- DATE\_TO\_UNIXTIME
- EXTRACT\_MONTH\_NAME
- UNIXTIME\_TO\_DATE
- 날짜를 텍스트로 변환
- 텍스트를 날짜로 변환

자세한 내용은 *Informatica 10.2.2 Enterprise Data Lake 사용자 가이드*에서 "데이터 준비" 장을 참조하십시오.

## 수학 함수

버전 10.2.2에서는 데이터를 준비할 때 다음과 같은 수학 함수를 열에 적용할 수 있습니다.

- EXP
- LN
- LOG
- PI
- POWER
- SQRT

자세한 내용은 *Informatica 10.2.2 Enterprise Data Lake 사용자 가이드*에서 "데이터 준비" 장을 참조하십시오.

## 텍스트 함수

버전 10.2.2에서는 데이터를 준비할 때 다음과 같은 텍스트 함수를 열에 적용할 수 있습니다.

- ENDSWITH
- ENDSWITH\_IGNORE\_CASE
- FIND\_IGNORE\_CASE
- FIND\_REGEX
- FIRST\_CHARACTER\_TO\_NUMBER
- NUMBER\_TO\_CHARACTER
- PROPER\_CASE

- REMOVE\_NON\_ALPHANUMERIC\_CHARACTERS
- STARTSWITH
- STARTSWITH\_IGNORE\_CASE
- SUBSTITUTE\_REGEX
- TRIM\_ALL
- 날짜를 텍스트로 변환
- 숫자를 텍스트로 변환
- 텍스트를 날짜로 변환
- 텍스트를 숫자로 변환

자세한 내용은 *Informatica 10.2.2 Enterprise Data Lake 사용자 가이드*에서 "데이터 준비" 장을 참조하십시오.

## 창 함수

버전 10.2.2에서는 창 함수를 사용하여 워크시트 내의 행 그룹에 대한 작업을 수행할 수 있습니다. 함수가 작동하는 행 그룹을 창이라고 하며, 이러한 창은 파티션 키, 정렬 기준 키 및 선택적 오프셋으로 정의합니다. 창 함수는 창 컨텍스트 내에서 모든 입력 행에 대한 반환 값을 계산합니다.

창 함수를 사용하여 다음과 같은 태스크를 수행할 수 있습니다.

- 이전 또는 후속 행에서 데이터를 검색합니다.
- 행 그룹에 기반한 누적 합계 또는 누적 평균을 계산합니다.
- 행 그룹의 각 행에 순차적인 행 번호를 할당합니다.
- 행의 Null 값을 행 그룹 내의 Null이 아닌 이전 값으로 바꿉니다.
- 로그 파일에 기록된 웹 사이트 방문과 같이, 특정 기간을 기반으로 행을 그룹화하는 데 사용할 수 있는 세션 식별자를 생성합니다.

여러 창 함수를 한 워크시트에 적용할 수 있습니다. 예를 들어 한 함수를 적용하여 창 내의 현재 행 다음에 오는 각 행에 대한 값의 합계를 계산한 다음 다른 함수를 적용하여 동일한 값의 평균을 계산할 수 있습니다.

Enterprise Data Lake는 사용자가 워크시트에 적용한 각 함수의 결과가 포함된 열을 추가합니다.

자세한 내용은 *Informatica 10.2.2 Enterprise Data Lake 사용자 가이드*에서 "데이터 준비" 장을 참조하십시오.

## 감사 이벤트 제거

버전 10.2.2에서는 `infacmd edl purgeevents` 명령을 실행하여 감사 기록 데이터베이스에서 사용자 활동 이벤트를 삭제할 수 있습니다. 필요에 따라 이 명령을 실행하여 데이터베이스에서 프로젝트 기록 이벤트를 삭제할 수 있습니다.

## Spark 실행 엔진

버전 10.2.2에서는 Enterprise Data Lake가 자산 계서와 같은 리소스 사용량이 높은 활동과 Python 변환을 사용하는 활성 규칙 맵셋을 실행하기 위해 Spark 엔진을 사용합니다. 리소스 사용량이 높은 활동에 Spark 엔진을 사용하면 성능이 향상되고 Amazon EMR(Elastic MapReduce) 기반의 Enterprise Data Lake 배포에서 자동 확장/축소를 활성화할 수 있습니다.

# Informatica Developer

이 섹션에서는 버전 10.2.2의 새로운 Developer tool 기능에 대해 설명합니다.

## 응용 프로그램

버전 10.2.2에서는 증분 응용 프로그램을 생성할 수 있습니다. 증분 응용 프로그램은 응용 프로그램 개체의 하위 집합을 업데이트하는 응용 프로그램 패치를 배포하여 업데이트할 수 있는 응용 프로그램입니다. 데이터 통합 서비스는 다른 응용 프로그램 개체를 계속 실행하면서 패치에 포함된 개체를 업데이트합니다.

버전 10.2.2로 업그레이드한 경우 기존 응용 프로그램은 "전체 응용 프로그램"으로 레이블이 지정됩니다. 버전 10.2.2에서 계속 전체 응용 프로그램을 생성할 수 있지만 전체 응용 프로그램을 증분 응용 프로그램으로 변환할 수는 없습니다.

자세한 내용은 *Informatica 10.2.2 Developer Tool 가이드*에서 "응용 프로그램 배포" 및 "응용 프로그램 패치 배포" 장을 참조하십시오.

# Informatica 매핑

이 섹션에서는 버전 10.2.2의 새로운 Informatica 매핑 기능에 대해 설명합니다.

## 데이터 유형

버전 10.2.2에서는 Spark 엔진에서 실행되는 일괄 처리 매핑에서 고정밀도 모드를 활성화할 수 있습니다. Spark 엔진은 최대 38자리의 전체 자릿수로 10진수 값을 처리할 수 있습니다.

자세한 내용은 *Informatica Big Data Management 10.2.2 사용자 가이드*를 참조하십시오.

## 매핑 출력

버전 10.2.2에서는 Spark 엔진의 워크플로우에서 매핑 태스크로 실행되는 일괄 매핑에서 매핑 출력을 사용할 수 있습니다. 매핑 출력을 모델 리포지토리에 유지하거나 매핑 출력을 워크플로우 변수에 바인딩할 수 있습니다.

자세한 내용은 *Informatica 10.2.2 Developer 매핑 가이드*의 "매핑 출력" 장 및 *Informatica 10.2.2 Developer 워크플로우 가이드*의 "매핑 태스크" 장을 참조하십시오.

## 매핑 매개 변수

버전 10.2.2에서는 원시 및 비원시 환경에서 실행되는 집계, 식 및 순위 변환에서 식 매개 변수를 포트 식에 할당할 수 있습니다.

자세한 내용은 *Informatica 10.2.2 Developer 매핑 가이드*에서 "매개 변수를 할당할 위치" 및 "동적 매핑" 장을 참조하십시오.

## 최적화 수준

버전 10.2.2에서는 매핑 및 매핑 태스크에 대한 자동 최적화 수준을 구성할 수 있습니다. 자동 최적화 수준을 사용하면 데이터 통합 서비스가 실행 모드 및 매핑 콘텐츠를 기반으로 최적화를 적용합니다.

새 매핑의 최적화 수준 기본값은 자동입니다.

버전 10.2.2로 업그레이드할 경우 매핑에 구성된 최적화 수준은 동일하게 유지됩니다. 업그레이드된 매핑에서 자동 최적화 수준을 사용하려면 최적화 수준을 수동으로 변경해야 합니다.

자세한 내용은 *Informatica 10.2.2 Developer 매핑 가이드*에서 "최적화 수준" 장을 참조하십시오.

## Sqoop

버전 10.2.2에서는 다음과 같은 새로운 Sqoop 기능을 사용할 수 있습니다.

### 증분 데이터 추출 지원

ID 또는 타임스탬프를 기반으로 증분 데이터 추출을 수행하도록 Sqoop 매핑을 구성할 수 있습니다. 증분 데이터 추출을 사용하면 Sqoop에서 마지막 데이터 추출 이후에 변경된 데이터만 추출합니다. 증분 데이터 추출은 매핑 성능을 향상시킵니다.

### Vertica 연결 지원

Vertica 소스에서 데이터를 읽거나 Vertica 대상에 데이터를 쓰도록 Sqoop를 구성할 수 있습니다.

### Sqoop 통과 매핑에 대한 Spark 엔진 최적화

Spark 엔진에서 Sqoop 소스를 포함하는 통과 매핑을 실행하면 다음과 같은 시나리오에서 데이터 통합 서비스가 매핑 성능을 최적화합니다.

- 사용자 지정 DDL 쿼리로 생성된 Hive 대상에 데이터를 씁니다.
- 사용자 지정 DDL 쿼리로 분할되거나 사용자 지정 DDL 쿼리로 분할되고 버킷 구성된 기존 Hive 대상에 데이터를 씁니다.
- 분할되고 버킷 구성된 기존 Hive 대상에 데이터를 씁니다.

### --infaownername 인수 지원

--infaownername 인수를 구성하여 Sqoop이 데이터 개체의 소유자 이름을 적용해야 하는지 여부를 지정할 수 있습니다.

자세한 내용은 *Informatica Big Data Management 10.2.2 사용자 가이드*를 참조하십시오.

## Informatica 변환

이 섹션에서는 버전 10.2.2에 포함된 Informatica 변환의 새로운 기능에 대해 설명합니다.

### 주소 유효성 검사기 변환

이 섹션에서는 새로운 주소 유효성 검사기 변환 기능에 대해 설명합니다.

주소 유효성 검사기 변환에는 다음과 같은 국가에 대해 추가적인 주소 기능이 포함됩니다.

#### 모든 국가

버전 10.2.2에서는 주소 유효성 검사기 변환이 Informatica가 참조 주소 데이터를 제공하는 모든 국가에서 단일 행 주소 유효성 검사를 지원합니다.

이전 버전에서는 26개 국가에 대해 단일 행 주소 유효성 검사가 지원되었습니다.

단일 행 주소를 확인하려면 완전한 주소 포트에 주소를 입력하십시오. 기본 설정 스크립트가 라틴어 또는 서방어 스크립트가 아닌 국가의 주소를 식별하는 경우 변환에서 주소와 함께 기본 설정 스크립트 속성을 사용합니다.

## 오스트레일리아

버전 10.2.2에서는 오스트레일리아 주소에 주소 보강을 추가하도록 주소 유효성 검사기 변환을 구성할 수 있습니다. 이 보강을 사용하여 오스트레일리아 통계국이 주소를 할당한 지리적 구획 및 구역을 검색할 수 있습니다. 이 구획 및 구역에는 인구조사 컬렉션 지구, 메시 블록 및 통계 영역이 포함됩니다.

변환에서는 다음과 같은 포트를 사용하여 데이터를 보강합니다.

- 인구조사 컬렉션 지구 코드 2006
- 좌표가 부여된 국내 주소 파일 식별자
- 규모가 큰 수도 통계 영역 5자리
- 규모가 큰 수도 통계 영역 이름
- 수준 1 통계 영역 11자리
- 수준 1 통계 영역 7자리
- 수준 2 통계 영역 9자리
- 수준 2 통계 영역 5자리
- 수준 2 통계 영역 이름
- 수준 3 통계 영역 5자리
- 수준 3 통계 영역 이름
- 수준 4 통계 영역 3자리
- 수준 4 통계 영역 이름
- 메시 블록 11자리 2011
- 메시 블록 11자리 2016
- 시/도 또는 권역 코드
- 시/도 또는 권역 이름
- 보조 AU 상태

이러한 포트는 AU 보조 포트 그룹에서 찾을 수 있습니다.

## 볼리비아

버전 10.2.2에서는 주소 유효성 검사기 변환의 볼리비아 주소 구문 분석 및 유효성 검사가 개선됩니다. 또한 볼리비아의 참조 데이터가 업데이트됩니다.

볼리비아의 경우 변환에 다음 개선 사항이 포함됩니다.

- 주소 유효성이 거리 수준까지 검사됩니다.
- 주요 도시의 주소에 대한 좌표가 거리 중간 지점 수준까지 표시됩니다.

## 캐나다

캐나다에서는 다음 기능 및 향상된 기능이 도입됩니다.

### 캐나다 주소에서 글로벌 기본 설정 설명자 속성 지원

버전 10.2.2에서는 요소 설명자의 짧은 형식 또는 긴 형식을 반환하도록 주소 유효성 검사기 변환을 구성할 수 있습니다.

변환은 다음 설명자의 짧은 형식 또는 긴 형식을 반환할 수 있습니다.

- 거리 설명자



- 방향 값
- 건물 설명자
- 하위 건물 설명자

설명자에 대한 출력 형식을 지정하려면 변환에서 글로벌 기본 설정 설명자 속성을 구성합니다. 속성은 영어 및 프랑스어 설명자에 적용됩니다. 기본적으로 변환은 참조 데이터가 지정하는 형식으로 설명자를 반환합니다. 속성에서 **PRESERVE INPUT** 옵션을 선택하면 기본 설정 언어 속성이 글로벌 기본 설정 설명자 속성보다 우선합니다.

#### CH 및 CHAMBER를 하위 건물 설명자로 지원

버전 10.2.2에서는 주소 유효성 검사기 변환이 캐나다 주소에서 CH 및 CHAMBER를 하위 건물 설명자로 인식합니다.

#### 콜롬비아

버전 10.2.2에서는 주소 유효성 검사기 변환의 콜롬비아 주소 거리 데이터 처리가 개선됩니다. 또한 콜롬비아의 참조 데이터가 업데이트됩니다.

#### 프랑스

버전 10.2.2에서는 프랑스 주소에 대한 다음 개선 사항이 도입됩니다.

- 프랑스의 보조 참조 데이터가 개선됩니다.
- 주소 유효성 검사기 변환은 프랑스의 IRIS 단위에 주소를 할당하여 정확도를 개선합니다. 변환은 주소의 집 번호를 사용하여 주소가 속하는 IRIS 단위를 확인합니다. 집 번호를 사용하면 주소가 서로 다른 단위의 경계에 근접한 경우 할당 정확도가 개선될 수 있습니다.

#### 이스라엘

버전 10.2.2에는 이스라엘에 대한 다음과 같은 기능 및 향상된 기능이 도입되었습니다.

##### 이스라엘 주소에 대한 다국어 지원

영어 또는 히브리어로 이스라엘 주소를 반환하도록 주소 유효성 검사기 변환을 구성할 수 있습니다.

기본 설정 언어 속성을 사용하여 변환이 반환하는 주소의 기본 설정 언어를 선택합니다.

이스라엘 주소의 기본 언어는 히브리어입니다. 주소 정보를 히브리어로 반환하려면 기본 설정 언어 속성을 **DATABASE** 또는 **ALTERNATIVE\_1**로 설정하십시오. 주소 정보를 영어로 반환하려면 이 속성을 **ENGLISH** 또는 **ALTERNATIVE\_2**로 설정하십시오.

##### 이스라엘 주소에 대한 다중 문자 집합 지원

주소 유효성 검사기 변환은 히브리어 및 라틴어 문자 집합으로 이스라엘 주소를 읽고 쓸 수 있습니다.

기본 설정 스크립트 속성을 사용하여 주소 데이터에 대한 기본 설정 문자 집합을 선택합니다.

이스라엘 주소의 기본 문자 집합은 히브리어입니다. 기본 설정 스크립트 속성을 라틴어 또는 라틴어-1로 설정하면 변환이 히브리어 주소 데이터를 라틴어 문자로 변환합니다.

#### 페루

버전 10.2.2에서는 주소 유효성 검사기 변환이 페루 주소의 유효성을 집 번호 수준까지 검사합니다. 또한 페루의 참조 데이터가 업데이트됩니다.

#### 스웨덴

버전 10.2.2에서는 스웨덴 주소 거리 이름에 대한 주소 유효성 검사기 변환의 유효성 검사가 개선됩니다.

거리 이름에 대한 변환의 유효성 검사는 다음과 같이 개선됩니다.

- 변환은 문자 **G**로 끝나는 거리 이름을 최종 문자 **GATAN**을 포함하는 동일한 이름의 별칭으로 인식할 수 있습니다.
- 변환은 문자 **V**로 끝나는 거리 이름을 최종 문자 **VÄGEN**을 포함하는 동일한 이름의 별칭으로 인식할 수 있습니다.
- 긴 형식 또는 짧은 형식의 설명자가 사용된 경우 주소 유효성 검사기 변환은 잘못된 설명자가 포함된 거리 이름을 인식하고 수정할 수 있습니다.

예를 들어 변환은 다음 주소의 **RUNIUSV** 또는 **RUNIUSVÄGEN**을 **RUNIUSGATAN**으로 수정할 수 있습니다.

**RUNIUSGATAN 7**

**SE-112 55 STOCKHOLM**

## 미국

버전 **10.2 HotFix 2**에서는 하나 이상의 요일에 우편물을 접수하지 않는 미국 주소를 식별하도록 주소 유효성 검사기 변환을 구성할 수 있습니다.

주소를 식별하려면 *배달이 없는 기간(일)* 포트를 사용합니다. 이 포트에는 일요일부터 토요일까지 주의 요일을 나타내는 7자리 문자열이 들어 있습니다. 문자열에서 각 위치는 서로 다른 요일을 나타냅니다.

주소 유효성 검사기 변환은 주소가 특정 요일에 우편물을 접수하지 않는 경우 포트에서 해당하는 위치에 요일의 첫 문자를 반환합니다. 변환은 주의 다른 요일에 해당하는 위치에서 대시 기호를 반환합니다.

예를 들어 배달이 없는 기간(일) 포트에서 **S----FS**의 값은 해당 주소가 일요일, 금요일 및 토요일에 우편물을 접수하지 않음을 나타냅니다.

기본 모델의 미국에만 해당 포트 그룹에서 배달이 없는 기간(일) 포트를 찾으십시오. 배달이 없는 기간(일) 포트에서 데이터를 수신하려면 인증된 모드에서 주소 유효성 검사기 변환을 실행하십시오. 변환은 **USA5C129.MD** 및 **USA5C130.MD** 데이터베이스 파일에서 포트 값을 읽습니다.

버전 **10.2.2**에 포함되는 주소 확인 소프트웨어 엔진의 기능 및 작동에 대한 종합적인 정보는 *Informatica Address Verification 5.14.0 개발자 가이드*를 참조하십시오.

## 업데이트 전략 변환

버전 **10.2.2**에서는 **Spark** 엔진에서 실행되는 매핑의 업데이트 전략 변환을 사용하여 관계형 대상을 업데이트할 수 있습니다.

이전에는 **Spark** 엔진에서 실행되는 매핑의 업데이트 전략 변환을 사용하여 **Hive** 대상만 업데이트할 수 있었습니다.

자세한 내용은 *Developer 변환 가이드*에서 업데이트 전략 변환 장을 참조하십시오.

# PowerExchange Adapters for Informatica

이 섹션에서는 버전 **10.2.2**의 새로운 **Informatica** 어댑터 기능에 대해 설명합니다.

## PowerExchange for Amazon Redshift

버전 10.2.2에서는 PowerExchange for Amazon Redshift에 다음과 같은 기능이 포함됩니다.

- 다음 지역에서 데이터를 읽거나 쓸 수 있습니다.
  - 중국(닝샤)
  - EU(파리)
- Amazon Redshift 개체를 매핑의 동적 소스 및 대상으로 사용할 수 있습니다.
- 인쇄할 수 있거나 인쇄할 수 없는 ASCII 문자의 8진수 값을 DELIMITER 또는 QUOTE로 사용할 수 있습니다.
- 사전 SQL 및 사후 SQL 명령을 입력하여 매핑의 소스 및 대상 개체에 대한 쿼리를 실행할 수 있습니다.
- 매핑의 읽기 데이터 개체에 대한 SQL 쿼리를 정의하여 기본 쿼리를 재정의할 수 있습니다. Amazon Redshift 데이터베이스에서 지원하는 SQL 문을 입력할 수 있습니다.
- 대형 Amazon S3 개체를 여러 부분으로 다운로드할 때 Amazon S3 개체의 최대 크기를 바이트 단위로 지정할 수 있습니다.
- Amazon Redshift 소스에서 데이터를 읽을 때 고유한 값을 읽을 수 있습니다.
- Amazon S3에 개체를 업로드할 때 개체의 최소 크기와 개체를 독립적인 부분의 집합으로 병렬로 업로드하는 스트레드 수를 지정할 수 있습니다.
- 기존 대상 테이블을 유지하거나, 런타임 시 대상 테이블을 바꾸거나, 대상에 테이블이 존재하지 않는 경우 새 대상 테이블을 생성하도록 선택할 수 있습니다.
- 원시 환경에서 Amazon Redshift 대상에 대한 업데이트 전략 변환을 구성할 수 있습니다.
- Amazon Redshift에 데이터를 쓸 때 런타임 동안 Amazon Redshift 대상 테이블 스키마 및 테이블 이름을 재정의할 수 있습니다.
- 연결 유형이 ODBC인 경우 데이터 통합 서비스는 소스 측 및 전체 푸시다운 최적화를 사용하여 변환 논리를 Amazon Redshift 소스 및 대상에 푸시할 수 있습니다.
- Amazon EMR 버전 5.16과 Cloudera CDH 버전 5.15 및 5.16에서 AWS KMS(AWS Key Management Service)과 함께 서버 측 암호화를 사용할 수 있습니다.
- PowerExchange for Amazon Redshift는 AWS SDK for Java 버전 1.11.354를 지원합니다.

자세한 내용은 *Informatica PowerExchange for Amazon Redshift 10.2.2 사용자 가이드*를 참조하십시오.

## PowerExchange for Amazon S3

버전 10.2.2에서는 PowerExchange for Amazon S3에 다음과 같은 기능이 포함됩니다.

- 다음 지역에서 데이터를 읽거나 쓸 수 있습니다.
  - 중국(닝샤)
  - EU(파리)
  - AWS GovCloud(US)
- Amazon S3 개체를 매핑의 동적 소스 및 대상으로 사용할 수 있습니다.
- 원시 환경이나 Spark 엔진에서 매핑을 실행하여 Avro, 플랫, JSON, ORC 또는 Parquet 파일의 데이터를 읽을 때 와일드카드 문자를 사용하여 소스 디렉터리 이름 또는 소스 파일 이름을 지정할 수 있습니다.
- Amazon S3 버킷에 저장된 개체에 태그를 하나 이상 추가하여 개체를 범주화할 수 있습니다. 각 태그에는 키 값 쌍이 포함됩니다. 키 값 쌍을 입력하거나 키 값 쌍이 포함된 절대 파일 경로를 지정할 수 있습니다.
- 최대 임계값 크기를 지정하여 Amazon S3 개체를 여러 부분으로 다운로드할 수 있습니다.

- Amazon S3에 개체를 업로드할 때 개체의 최소 크기와 개체를 독립적인 부분의 집합으로 병렬로 업로드하는 스레드 수를 지정할 수 있습니다.
- 데이터 개체 읽기 또는 쓰기 작업을 생성할 때 끝점 이름 및 파일의 소스 경로를 포함하는 파일 이름 포트에 있는 데이터를 읽을 수 있습니다.
- 데이터 개체 읽기 또는 쓰기 작업을 생성할 때 포트 탭에서 직접 새 열을 추가하거나 열을 수정할 수 있습니다.
- Avro, JSON, ORC 또는 Parquet 파일을 읽거나 쓰는 매핑 생성 시 포트 탭에서 소스 변환, 대상 변환 또는 다른 모든 변환의 열을 복사하여 데이터 개체 읽기 또는 쓰기 작업에 직접 열을 붙여 넣을 수 있습니다.
- 다른 Amazon S3 파일 형식에 대한 열 예측 속성을 구성한 후에도 열 예측 속성의 스키마 필드에서 열 메타데이터 손실 없이 Amazon S3 파일 형식을 업데이트할 수 있습니다.
- Amazon EMR 버전 5.16과 Cloudera CDH 버전 5.15 및 5.16에서 AWS KMS(AWS Key Management Service)과 함께 서버 측 암호화를 사용할 수 있습니다.
- PowerExchange for Amazon S3는 AWS SDK for Java 버전 1.11.354를 지원합니다.

자세한 내용은 *Informatica PowerExchange for Amazon S3 10.2.2 사용자 가이드*를 참조하십시오.

## PowerExchange for Google BigQuery

버전 10.2.2에서는 **대상 생성** 마우스 오른쪽 단추 클릭 옵션을 사용하여 Google BigQuery 대상을 생성할 수 있습니다.

자세한 내용은 *Informatica PowerExchange for Google BigQuery 10.2.2 사용자 가이드*를 참조하십시오.

## PowerExchange for HBase

버전 10.2.2에서는 PowerExchange for HBase에 다음과 같은 새로운 기능이 포함됩니다.

- HBase 데이터 개체를 생성할 때 운영 체제 프로필을 선택하여 보안을 개선하고 Hadoop 클러스터에서 메타데이터를 가져오고 미리 볼 때 디자인 타임 사용자 환경을 격리할 수 있습니다.  
**참고:** 메타데이터 액세스 서비스가 운영 체제 프로필을 사용하도록 구성된 경우 운영 체제 프로필을 선택할 수 있습니다. 메타데이터 액세스 서비스는 사용자에게 할당된 기본 운영 체제 프로필을 사용하여 메타데이터를 가져옵니다. 사용 가능한 운영 체제 프로필 목록에서 운영 체제 프로필을 변경할 수 있습니다.
- HBase 개체를 매핑의 동적 소스 및 대상으로 사용할 수 있습니다.
- Spark 엔진에서 매핑을 실행하여 HBase 리소스의 데이터를 조회할 수 있습니다.

자세한 내용은 *Informatica PowerExchange for HBase 10.2.2 사용자 가이드*를 참조하십시오.

## PowerExchange for HDFS

버전 10.2.2에서는 PowerExchange for HDFS에 다음과 같은 새로운 기능이 포함됩니다.

- 복합 파일 데이터 개체를 생성할 때 운영 체제 프로필을 선택하여 보안을 개선하고 Hadoop 클러스터에서 메타데이터를 가져오고 미리 볼 때 디자인 타임 사용자 환경을 격리할 수 있습니다.  
**참고:** 메타데이터 액세스 서비스가 운영 체제 프로필을 사용하도록 구성된 경우 운영 체제 프로필을 선택할 수 있습니다. 메타데이터 액세스 서비스는 사용자에게 할당된 기본 운영 체제 프로필을 사용하여 메타데이터를 가져옵니다. 사용 가능한 운영 체제 프로필 목록에서 운영 체제 프로필을 변경할 수 있습니다.
- 원시 환경이나 Spark 엔진에서 매핑을 실행하여 복합 파일 데이터 개체의 데이터를 읽을 때 와일드카드 문자를 사용하여 소스 디렉터리 이름 또는 소스 파일 이름을 지정할 수 있습니다.  
다음 와일드카드 문자를 사용할 수 있습니다.

### ? (물음표)

물음표 문자(?)를 사용하여 임의 문자의 1회 발생을 나타낼 수 있습니다.

### \*(별표)

별표 문자(\*)를 사용하여 임의 문자의 0회 또는 1회 이상의 발생을 나타낼 수 있습니다.

- 복합 파일 개체를 매핑의 동적 소스 및 대상으로 사용할 수 있습니다.
- 복합 파일 개체를 사용하여 복합 파일 시스템에서 데이터를 읽고 복합 파일 시스템에 데이터를 쓸 수 있습니다.
- 원시 환경 또는 **Spark** 엔진에서 매핑을 실행하여 복합 파일 데이터 개체에 데이터를 쓸 때 대상 데이터를 덮어쓸 수 있습니다. 데이터 통합 서비스는 새 데이터를 쓰기 전에 대상 데이터를 삭제합니다.
- 데이터 개체 읽기 또는 쓰기 작업을 생성할 때 끝점 이름 및 파일의 소스 경로를 포함하는 파일 이름 포트에 있는 데이터를 읽을 수 있습니다.
- 이제 데이터 개체 읽기 또는 쓰기 작업을 생성한 후 즉시 데이터 개체 작업을 볼 수 있습니다.
- 데이터 개체 읽기 또는 쓰기 작업을 생성할 때 새 열을 추가하거나 열을 수정할 수 있습니다.
- **Avro**, **JSON**, **ORC** 또는 **Parquet** 파일의 읽기 또는 쓰기 시 직접 소스 변환, 대상 변환 또는 다른 모든 변환의 열을 복사하여 데이터 개체 읽기 또는 쓰기 작업에 열을 붙여 넣을 수 있습니다.

자세한 내용은 *Informatica PowerExchange for HDFS 10.2.2 사용자 가이드*를 참조하십시오.

## PowerExchange for Hive

버전 10.2.2에서는 **PowerExchange for Hive**에 다음과 같은 새로운 기능이 포함됩니다.

- Hive 대상에 대해 다음과 같은 대상 스키마 전략 옵션을 구성할 수 있습니다.
  - **RETAIN** - 기존 대상 스키마 유지
  - **CREATE** - 런타임 시 테이블 생성 또는 바꾸기
  - **APPLYNEWCOLUMNS** - 테이블 변경 및 새 열만 적용
  - **APPLYNEWSHEMA** - 테이블 변경 및 새 스키마 적용
  - **FAIL** - 대상 스키마가 다른 경우 매핑 실패
  - 매개 변수 할당
- 데이터를 로드하기 전에 내부 또는 외부 분할된 **Hive** 대상을 잘라낼 수 있습니다. 이 옵션은 **Hadoop** 환경에서 매핑을 실행할 때 적용할 수 있습니다.
- 원시 모드에서 **Hive**에 대한 읽기 또는 쓰기 변환을 생성하여 **Hive** 소스에서 데이터를 읽거나 **Hive** 대상에 데이터를 쓸 수 있습니다.
- **Hive** 대상에 데이터를 쓸 때 **Hive** 연결에서 다음과 같은 속성을 구성할 수 있습니다.
  - **HDFS**의 **Hive** 준비 디렉터리. **Hive** 준비 테이블에 대한 **HDFS** 디렉터리를 나타냅니다. 이 옵션은 원시 환경의 **Hive** 대상에 데이터를 쓰는 경우 적용되고 필요합니다.
  - **Hive** 준비 데이터베이스 이름. **Hive** 준비 테이블에 대한 네임스페이스를 나타냅니다. 이 옵션은 원시 환경에서 **Hive** 대상에 데이터를 쓰는 매핑을 실행하는 경우 적용됩니다. **Blaze** 또는 **Spark** 엔진에서 매핑을 실행하는 경우 **Hive** 연결에 **Hive** 준비 데이터베이스 이름을 구성하지 않아도 됩니다. 데이터 통합 서비스는 **Hadoop** 연결에 구성된 값을 사용합니다.

자세한 내용은 *Informatica PowerExchange for Hive 10.2.2 사용자 가이드*를 참조하십시오.

## PowerExchange for MapR-DB

버전 10.2.2에서는 MapR-DB에 대한 HBase 데이터 개체를 생성할 때 운영 체제 프로필을 선택하여 보안을 개선하고 Hadoop 클러스터에서 메타데이터를 가져오고 미리 볼 때 디자인 타임 사용자 환경을 격리할 수 있습니다.

**참고:** 메타데이터 액세스 서비스가 운영 체제 프로필을 사용하도록 구성된 경우 운영 체제 프로필을 선택할 수 있습니다. 메타데이터 액세스 서비스는 사용자에게 할당된 기본 운영 체제 프로필을 사용하여 메타데이터를 가져옵니다. 사용 가능한 운영 체제 프로필 목록에서 운영 체제 프로필을 변경할 수 있습니다.

자세한 내용은 *Informatica PowerExchange for MapR-DB 10.2.2 사용자 가이드*를 참조하십시오.

## PowerExchange for Microsoft Azure Blob Storage

버전 10.2.2에서는 PowerExchange for Microsoft Azure Blob Storage에 다음 기능이 포함됩니다.

- Azure Databricks 환경에서 매핑을 실행할 수 있습니다.
- US Government Microsoft Azure 끝점을 구성할 수 있습니다.
- Microsoft Azure Blob Storage에서 데이터 읽거나 데이터를 쓸 때 다음과 같은 형식의 데이터를 압축할 수 있습니다.
  - 없음
  - Deflate
  - Gzip
  - Bzip2
  - Lzo
  - Snappy
- Microsoft Azure Blob Storage 개체를 매핑에서 동적 소스 및 대상으로 사용할 수 있습니다.
- 원시 환경에서 런타임에 데이터 통합 서비스가 데이터를 읽는 파일의 이름을 읽을 수 있습니다.
- 고급 소스 및 대상 속성의 **Blob 컨테이너 재정의**에서 상대 경로를 구성할 수 있습니다.

자세한 내용은 *Informatica PowerExchange for Microsoft Azure Blob Storage 10.2.2 사용자 가이드*를 참조하십시오.

## PowerExchange for Microsoft Azure Cosmos DB SQL API

버전 10.2.2에서는 PowerExchange for Microsoft Azure Cosmos DB SQL API에 다음 기능이 포함됩니다.

- Azure Databricks 환경에서 매핑을 실행할 수 있습니다. PowerExchange for Microsoft Azure Cosmos DB SQL API에 대한 Databricks 지원은 기술 미리 보기에서 사용할 수 있습니다. 기술 미리 보기 기능은 지원되지만 보장되지 않으며 프로덕션용으로 준비되지 않았습니다. 이러한 기능은 비프로덕션 환경에서만 사용하는 것이 좋습니다.

자세한 내용은 *Informatica PowerExchange for Microsoft Azure Cosmos DB SQL API 10.2.2 사용자 가이드*를 참조하십시오.

## PowerExchange for Microsoft Azure Data Lake Store

버전 10.2.2에서는 PowerExchange for Microsoft Azure Data Lake Store에 다음 기능이 포함됩니다.

- Azure Databricks 환경에서 매핑을 실행할 수 있습니다.

- Hadoop 환경에서 실행되는 매핑에서 배열, 구조 및 맵과 같은 복합 데이터 유형을 사용할 수 있습니다. 복합 데이터 유형을 사용하면 해당하는 엔진이 Avro, JSON 및 Parquet 복합 파일의 계층 데이터를 직접 읽고 처리하고 씁니다. 지능형 구조 소스의 경우 읽기 작업만 구성할 수 있습니다.
- 매핑을 생성하여 원시 환경의 기본 데이터 유형만 포함하는 Avro 및 Parquet 파일을 읽고 쓸 수 있습니다.
- 디렉터리를 매핑의 소스로 선택하여 해당 디렉터리에서 여러 파일을 읽을 수 있습니다.
- Microsoft Azure Data Lake Store 개체를 매핑에서 동적 소스 및 대상으로 사용할 수 있습니다.
- 대상 생성 옵션을 사용하여 Microsoft Azure Data Lake Store 대상을 생성할 수 있습니다.

자세한 내용은 *Informatica PowerExchange for Microsoft Azure Data Lake Store 10.2.2 사용자 가이드*를 참조하십시오.

## PowerExchange for Microsoft Azure SQL Data Warehouse

버전 10.2.2에서는 PowerExchange for Microsoft Azure SQL Data Warehouse에 다음 기능이 포함됩니다.

- Azure Databricks 환경에서 매핑을 실행할 수 있습니다.
- 원시 환경 및 Spark 엔진에서 실행되는 매핑에서 US Government Microsoft Azure 끝점을 구성할 수 있습니다.
- Microsoft Azure Blob Storage 컨테이너에서 오류 파일을 생성할 수 있습니다. 오류 파일에는 거부된 행과 거부된 행에 대한 원인이 포함됩니다.
- 원시 환경의 고급 대상 속성에서 일괄 처리 크기를 정의할 수 있습니다.
- 변환 논리를 소스 데이터베이스 및 대상 데이터베이스에 푸시하는 전체 푸시다운 최적화를 구성할 수 있습니다. 데이터베이스 리소스를 사용하여 태스크 성능을 개선하려면 푸시다운 최적화를 사용하십시오.
- Microsoft Azure SQL Data Warehouse 개체를 매핑에서 동적 소스 및 대상으로 사용할 수 있습니다.

PowerExchange for Microsoft Azure SQL Data Warehouse에 대한 전체 푸시다운 최적화 및 동적 매핑 기능은 기술 미리 보기에서 사용할 수 있습니다. 기술 미리 보기 기능은 지원되지만 보장되지 않으며 프로덕션용으로 준비되지 않았습니다. 이러한 기능은 비프로덕션 환경에서만 사용하는 것이 좋습니다.

자세한 내용은 *Informatica PowerExchange for Microsoft Azure SQL Data Warehouse 10.2.2 사용자 가이드*를 참조하십시오.

## PowerExchange for Salesforce

버전 10.2.2에서는 PowerExchange for Salesforce에 다음과 같은 새로운 기능이 포함됩니다.

- 버전 43.0 및 44.0의 Salesforce API를 사용하여 Salesforce 연결을 생성하고 Salesforce 개체에 액세스할 수 있습니다.
- Salesforce 연결에 대한 OAuth를 구성할 수 있습니다.
- 소스 데이터 개체 작업에 대한 원시 식 필터를 구성할 수 있습니다.
- Salesforce 데이터 개체에 대한 다음과 같은 읽기 작업 속성을 매개 변수화할 수 있습니다.
  - SOQL 필터 조건
  - CDC 시작 타임스탬프
  - CDC 끝 타임스탬프
  - PK 청킹 크기
  - PK 청킹 시작 행 ID



Salesforce 데이터 개체에 대한 다음과 같은 쓰기 작업 속성을 매개 변수화할 수 있습니다.

- BULK 성공 및 오류 파일에 대한 접두사 설정
- SFDC 성공 파일 디렉터리
- BULK 오류 파일의 위치 설정

자세한 내용은 *Informatica PowerExchange for Salesforce 10.2.2 사용자 가이드*를 참조하십시오.

## PowerExchange for Snowflake

버전 10.2.2에서는 PowerExchange for Snowflake에 다음과 같은 새로운 기능이 포함됩니다.

- Snowflake 연결의 JDBC URL 매개 변수에 인증 세부 정보를 지정하여 Okta SSO 인증을 구성할 수 있습니다.
- SQL 재정의의 구성하여 Snowflake 소스에서 데이터를 추출하는 데 사용되는 기본 SQL 쿼리를 재정의할 수 있습니다. Snowflake 데이터 개체 읽기 작업 속성에서 SQL 재정의의 지정하십시오.
- Snowflake 테이블에 쓰기 전에 파일을 압축하여 쓰기 성능을 최적화하도록 선택할 수 있습니다. 이는 고급 속성에서 수행하면 됩니다. Snowflake 데이터 개체 쓰기 작업 고급 속성의 추가 쓰기 런타임 매개 변수 필드에서 압축 매개 변수를 켜짐 또는 꺼짐으로 지정할 수 있습니다.
- 데이터 통합 서비스는 Snowflake Spark 커넥터 API를 사용하여 Spark 엔진에서 Snowflake 매핑을 실행합니다.
- Azure 또는 Amazon의 준비 데이터에 대해 활성화된 Snowflake에서 데이터를 읽거나 쓸 수 있습니다.

자세한 내용은 *Informatica PowerExchange for Snowflake 10.2.2 사용자 가이드*를 참조하십시오.

## PowerExchange for Teradata Parallel Transporter API

버전 10.2.2에서는 PowerExchange for Teradata Parallel Transporter API에 다음과 같은 고급 대상 속성의 기능이 포함됩니다.

- 대상에 데이터를 로드하는 동안 Teradata 데이터베이스에서 지원되지 않는 Teradata 유니코드 문자 대신 사용할 교체 문자를 지정할 수 있습니다.
- Teradata 대상에 데이터를 로드하는 동안 지원되지 않는 문자 대신에 사용되는 문자를 지정한 경우 대상 Teradata 데이터베이스에 버전 8.x ~ 13.x 또는 14.x 이상을 지정할 수 있습니다. 교체 문자 특성과 함께 이 특성을 사용합니다. Teradata 대상에 데이터를 로드하는 동안 교체 문자를 지정하지 않은 경우 데이터 통합 서비스는 이 특성을 무시합니다.
- Teradata에 데이터를 쓸 때 런타임 동안 Teradata 대상 테이블 스키마 및 테이블 이름을 재정의할 수 있습니다.

자세한 내용은 *Informatica PowerExchange for Teradata Parallel Transporter API 10.2.2 사용자 가이드*를 참조하십시오.