

빅 데이터 시대의 정보 거버넌스 전략

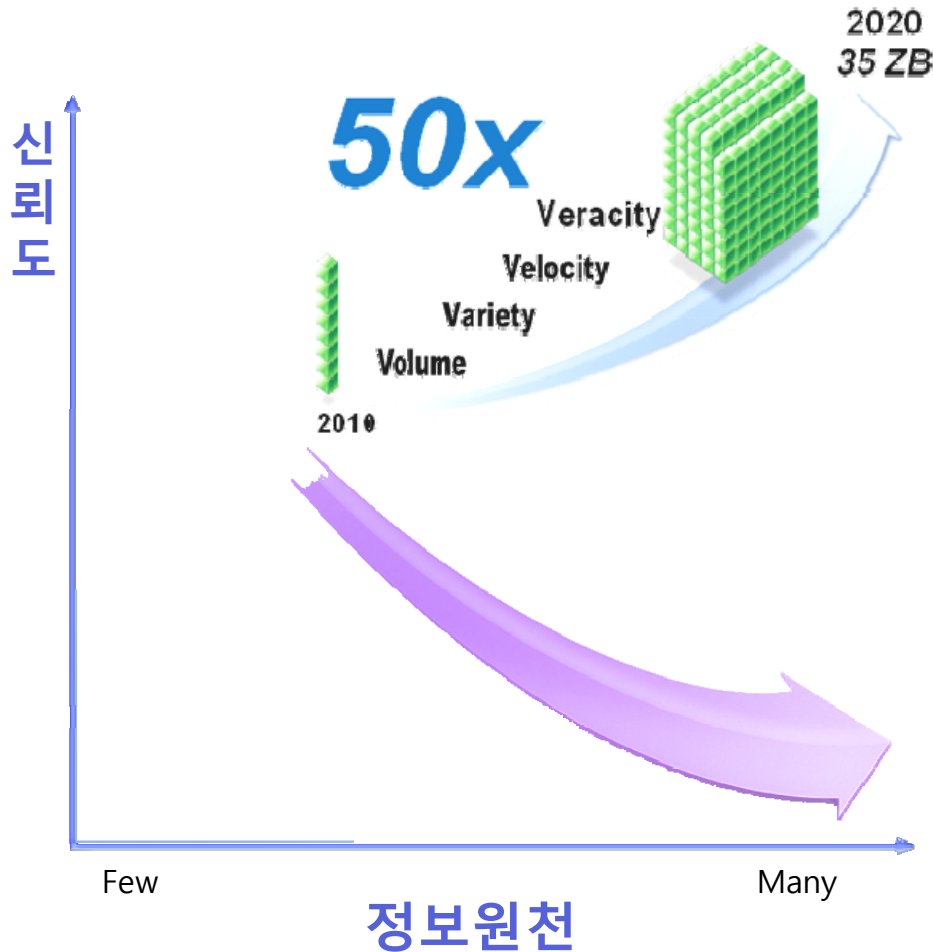
한국IBM 소프트웨어그룹 정보관리사업부
조가원 차장 (kwcho@kr.ibm.com)



빅 데이터 시대의 도래



빅 데이터 시대, 폭증하는 데이터



1 in 3 신뢰할 수 없는 데이터

조직의 리더 중 1/3은 의사 결정에 필요한 정보를 제대로 얻지 못하거나 신뢰할 수 없는 경우가 대부분이라고 말한다.

1 in 2 필요한 정보의 부재

업무 담당자 중 1/2는 업무에 필요한 정보에 대한 권한이 없다고 말한다.

60% 너무 방대한 데이터

CEO 중 60%는 경영 의사 결정을 위해 빠르게 데이터를 캡처/이해하는 향상된 기술이 필요하다.

Source: 2012 IBM GBS CEO Study: http://w3.ibm.com/ibm/resource/gbs_ceostudy2012.html



빅 데이터를 활용한 새로운 통찰력이 필요한 시대

Financial Services

- Fraud detection
- 360° View of the Customer



Utilities

- Weather analysis
- Smart grid management



Transportation

- Logistics optimization
- Traffic congestion



IT

- System Log Analysis
- Cyber security



Health/Life Sciences

- Epidemic early warning
- ICU monitoring



Retail

- 360° View of the Customer
- Real-time promotions



Telecommunications

- Geomapping / marketing
- Network monitoring



Law Enforcement

- Multimodal surveillance
- Cyber security detection

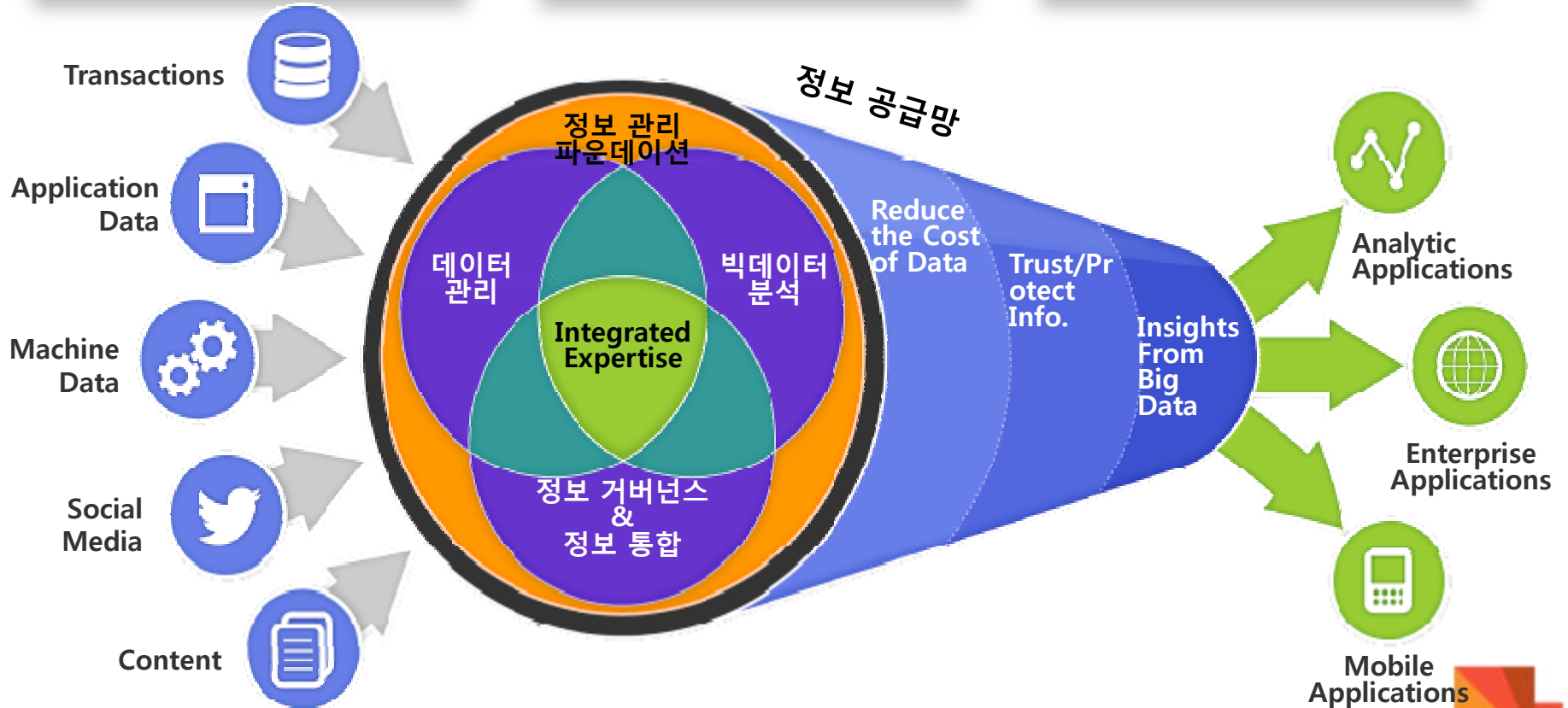


정보 공급망과 빅 데이터 관리 이슈

비용 절감

정보 신뢰도

통찰력



정보 거버넌스



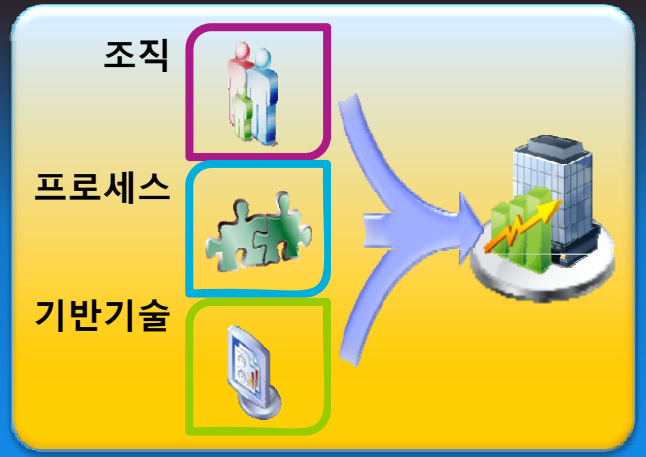
빅 데이터 플랫폼



정보 거버넌스

정보 거버넌스는

정보 품질, 마스터데이터, 보안, 정보수명주기 등 정보를 기업의 비즈니스를 위한 핵심 자산으로 최적화하고, 보호하며, 활용하기 위한 방안



Continued Innovation and market leadership

IBM Defines the IIG Market



2003

정보거버넌스 의회 건립



2005

성숙도 모델 개발



2006

2008

2010

2011

2013



정보 거버넌스 지원을 위한 IBM InfoSphere

포괄적인

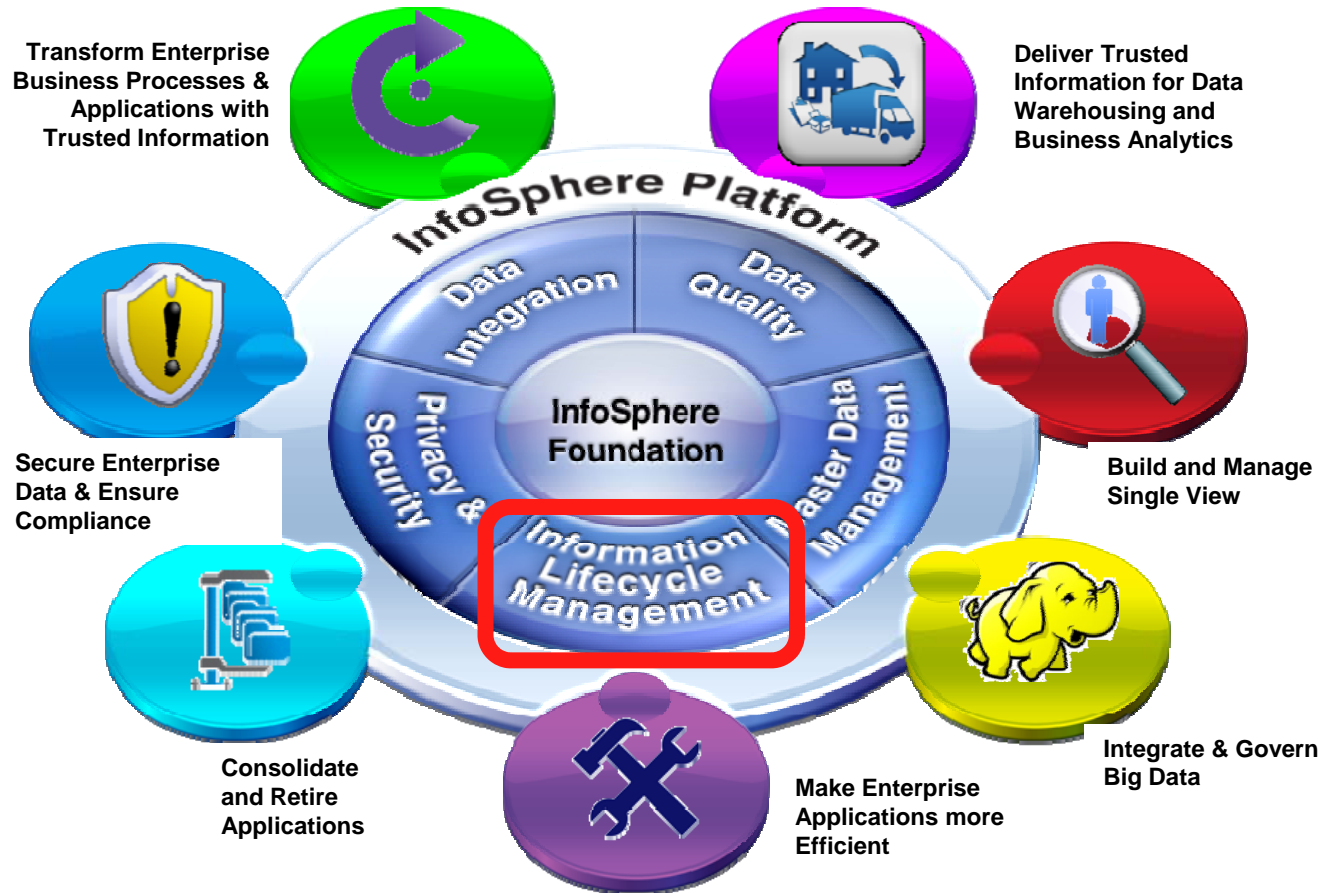
Most mature and comprehensive capabilities across all of IIG

통합된

Integrated capabilities designed to address enterprise use cases

똑똑한

Prebuilt, Automated, Proactive



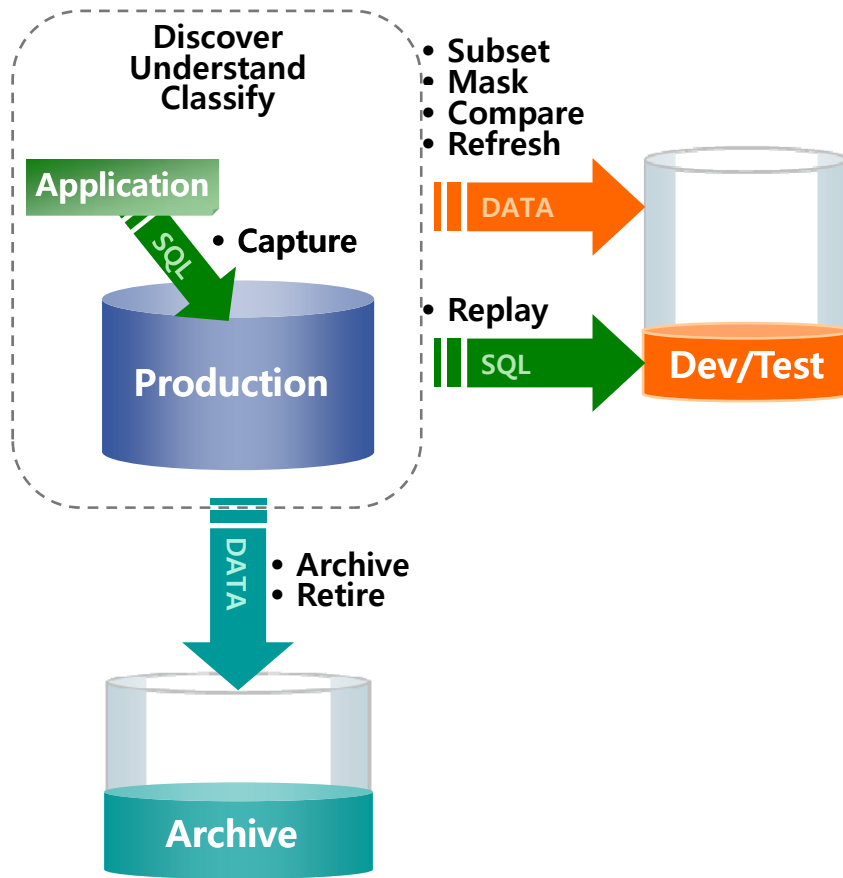
InfoSphere is the market leader in every category of Information Integration and Governance



정보 거버넌스를 통한 수명주기 관리 통찰력의 실행



정보 거버넌스 기반의 정보 수명주기 관리



Discover

- 복잡한 데이터베이스 내 업무 연관도에 의한 데이터 정의, 민감 데이터 식별을 통해 데이터 관리 프로젝트 일정 가속화

Manage Test Data

- 민감 데이터가 정재된 적정 사이즈의 테스트 데이터 관리를 통해 비용 절감, 위험 감소 및 애플리케이션 릴리즈 일정 가속화

Capture and Replay

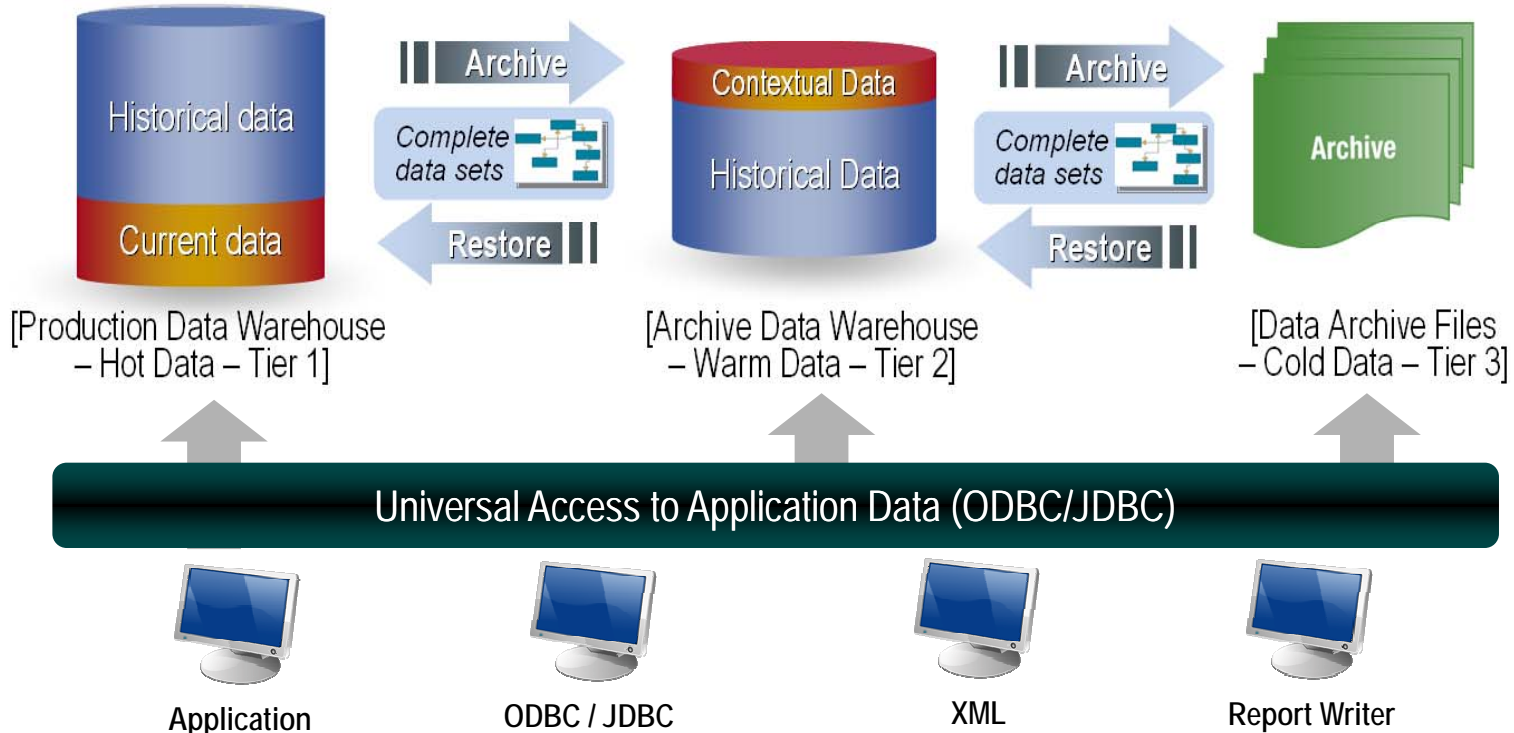
- 실운영 워크로드를 이용한 테스트 시나리오 수행을 통해 테스트 품질 향상: 고객 만족도 제고, 비용 절감 및 SLA 충족

Archive

- S/W, H/W, 스토리지 등 시스템 관리 비용 절감
- 운영 애플리케이션 성능 저하 방지
- 백업/복구 시간 단축
- 업그레이드/마이그레이션 지원
- 데이터 보관 규제 준수 및 시스템 최적화



아카이빙을 통한 정보 수명주기 관리



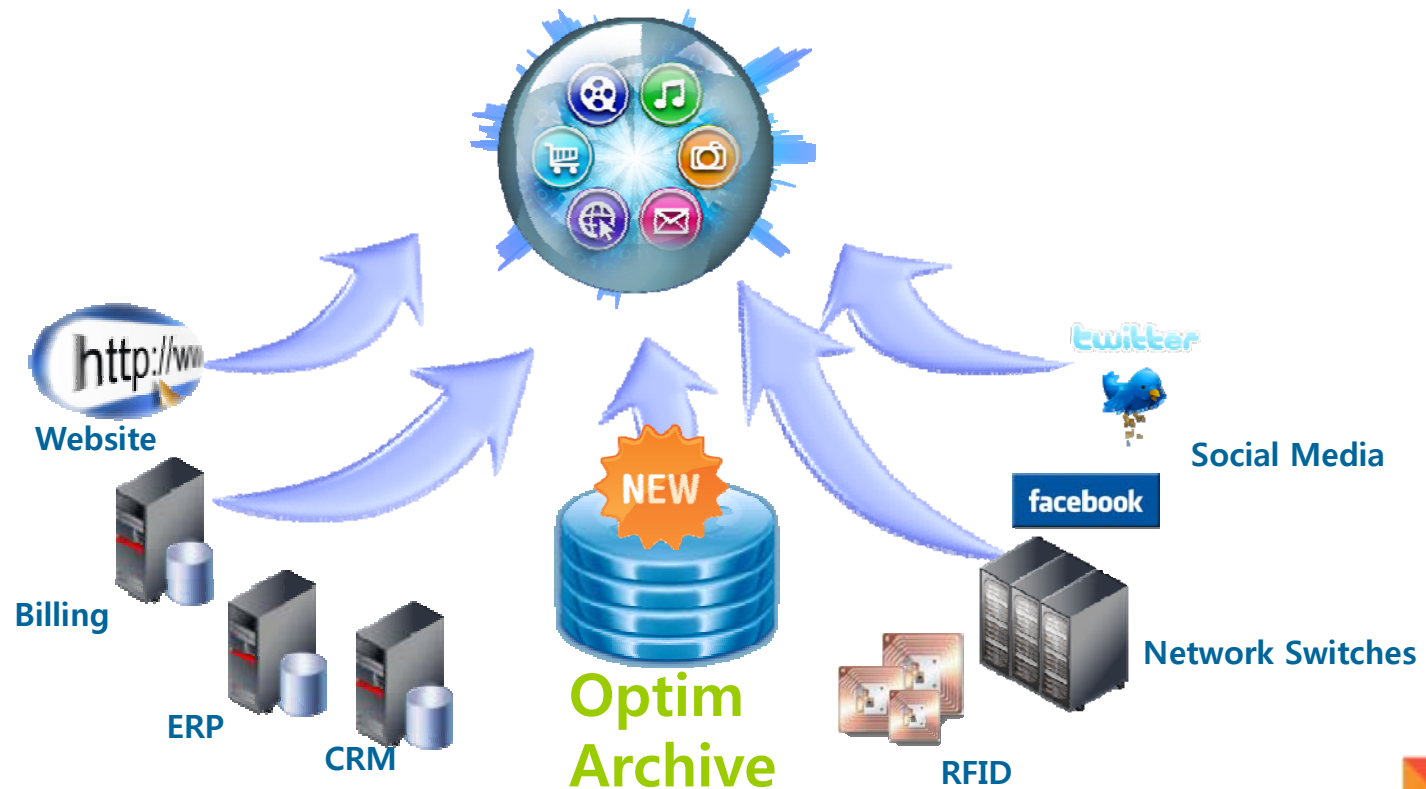
Data Archiving is an intelligent process for **moving** inactive or infrequently accessed data that still has **value**, while providing the ability to **search and retrieve** the data and **monitor access** to archived files



전사 아카이빙 데이터를 활용한 빠른 의사결정

정보 거버넌스 - 정보 수명주기 관리			
탐색/정의	개발/	최적화/ 아카이빙	폐기

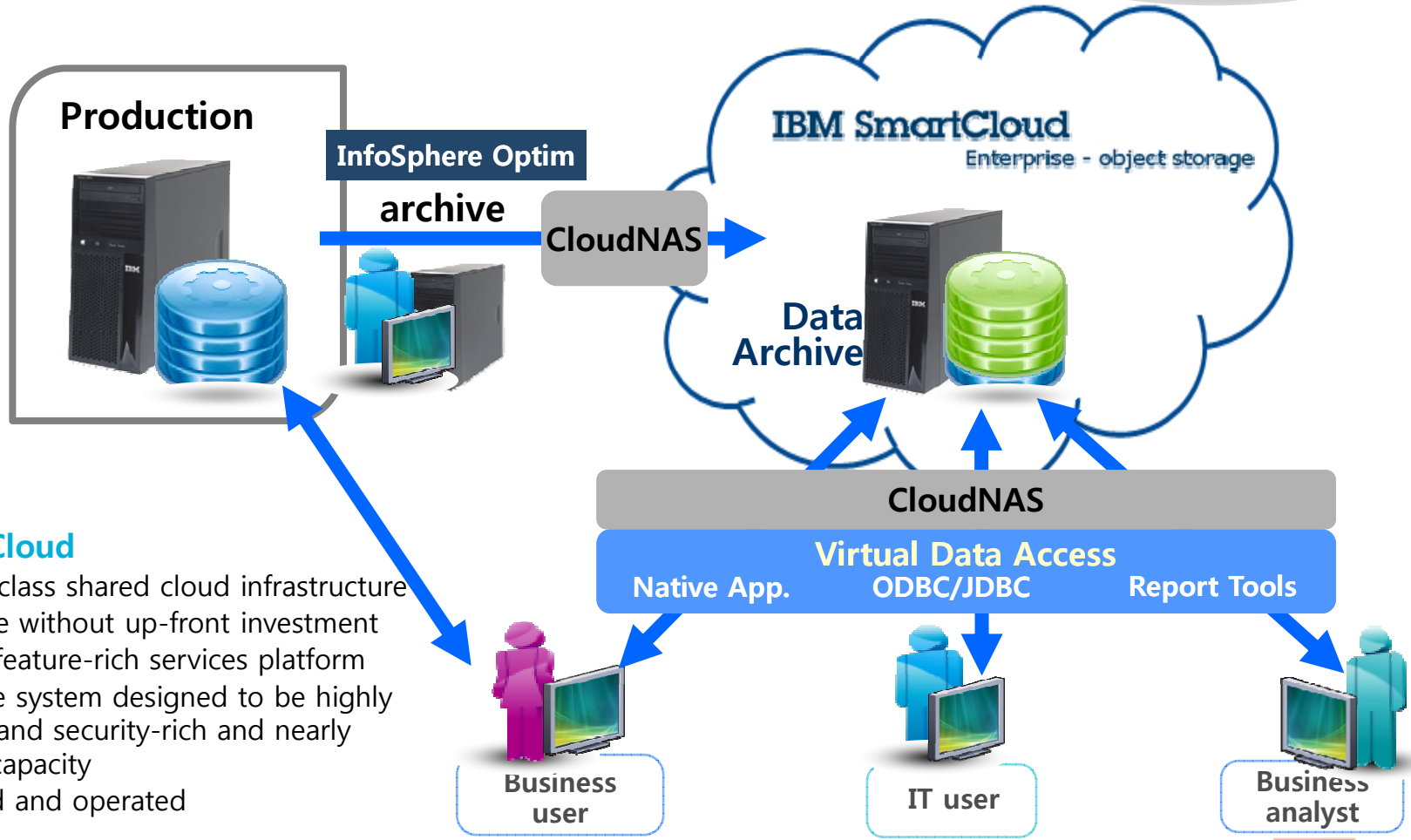
활용 가능한 전사 데이터를 비용 효율적으로 관리하고 분석
unstructured, structured, streaming, archive



IBM Smart Cloud Enterprise를 활용한 아카이빙

정보 거버넌스 - 정보 수명주기 관리

탐색/정의 개발/ **최적화/ 아카이빙** 폐기



IBM Smart Cloud

- Enterprise-class shared cloud infrastructure
- Pay-per-use without up-front investment
- A flexible, feature-rich services platform
- Internet file system designed to be highly accessible and security-rich and nearly unlimited capacity
- IBM owned and operated

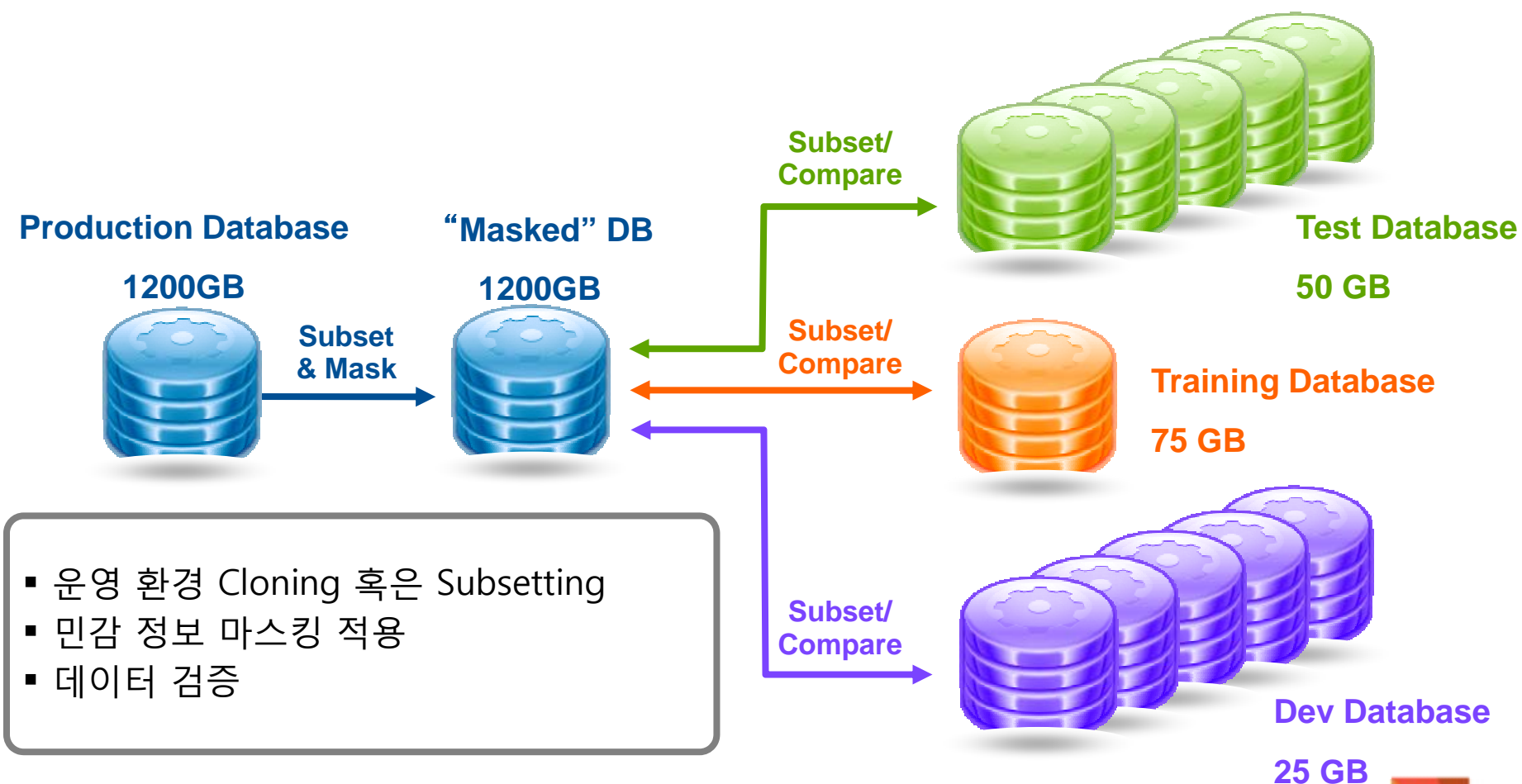
More at: ibm.com/smartcloud/solutions/enterprise



개발/테스트 데이터 관리

정보 거버넌스 - 정보 수명주기 관리

탐색/가	개발/테스트	하/방	통합/폐기
------	---------------	-----	-------



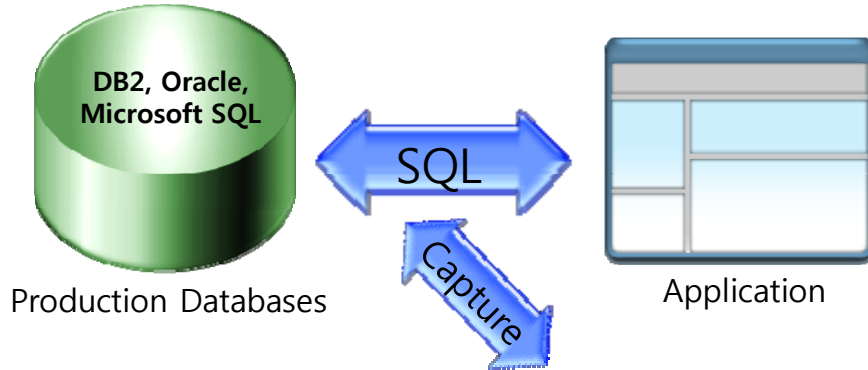
- 운영 환경 Cloning 혹은 Subsetting
- 민감 정보 마스킹 적용
- 데이터 검증



개발/테스트 워크로드 재생

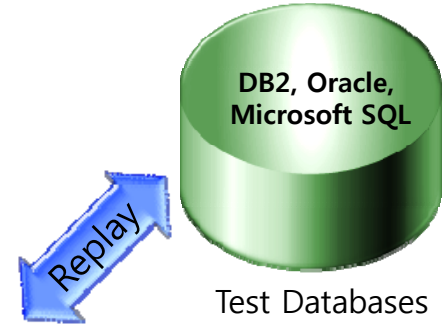
정보 거버넌스 - 정보 수명주기 관리

탐색/	개발/테스트	하/	통합/폐기
-----	--------	----	-------



InfoSphere Optim Query Capture and Replay

운영 워크로드 재생



- 원본 애플리케이션 수행 타이밍
- 실행 순서 정의
- 트랜잭션 그룹 정의
- 독립 레벨 정의
- 성능 분석 지표
- 주요 SQL & 애플리케이션 파악



정보 거버넌스 지원을 위한 IBM InfoSphere

포괄적인

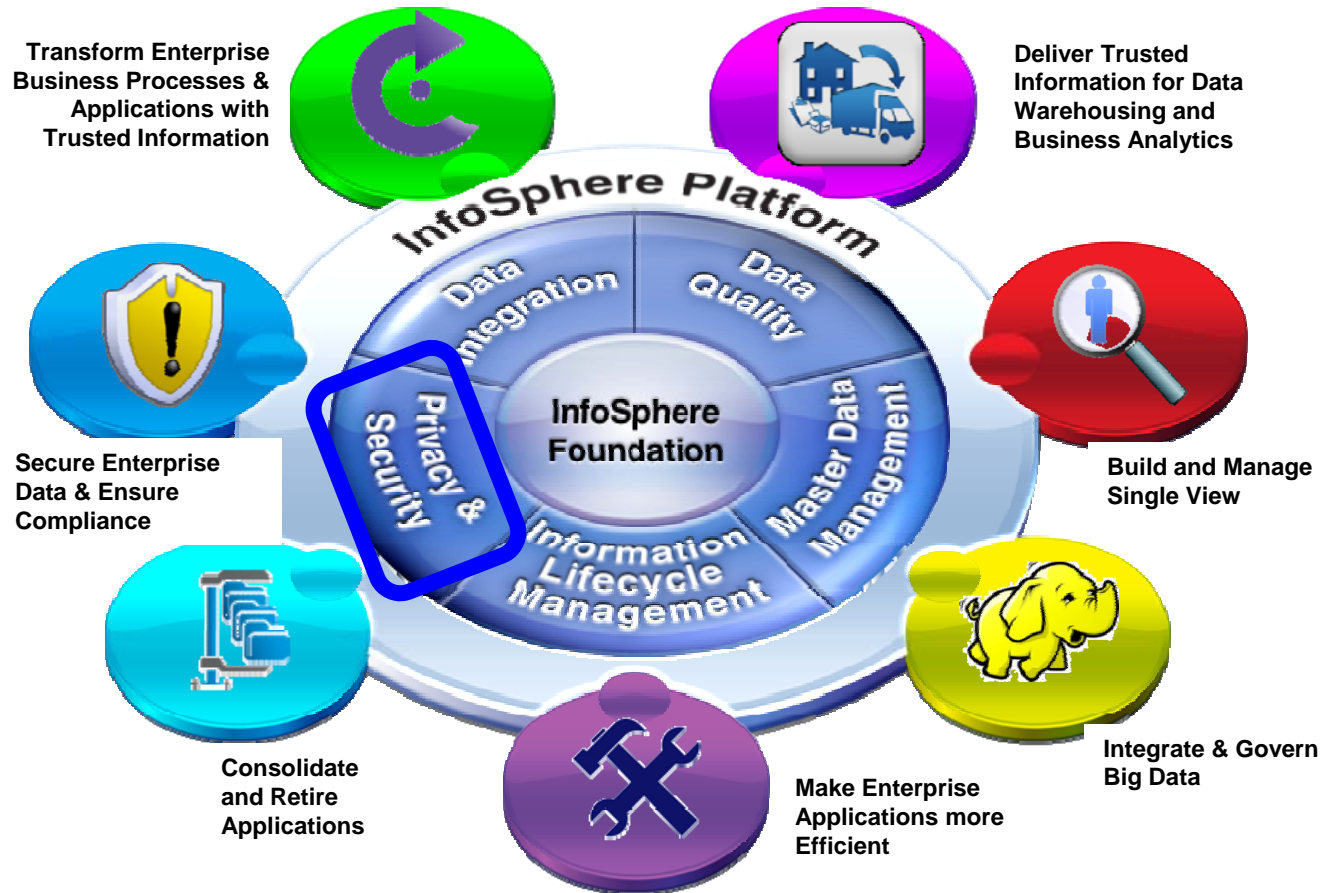
Most mature and comprehensive capabilities across all of IIG

통합된

Integrated capabilities designed to address enterprise use cases

똑똑한

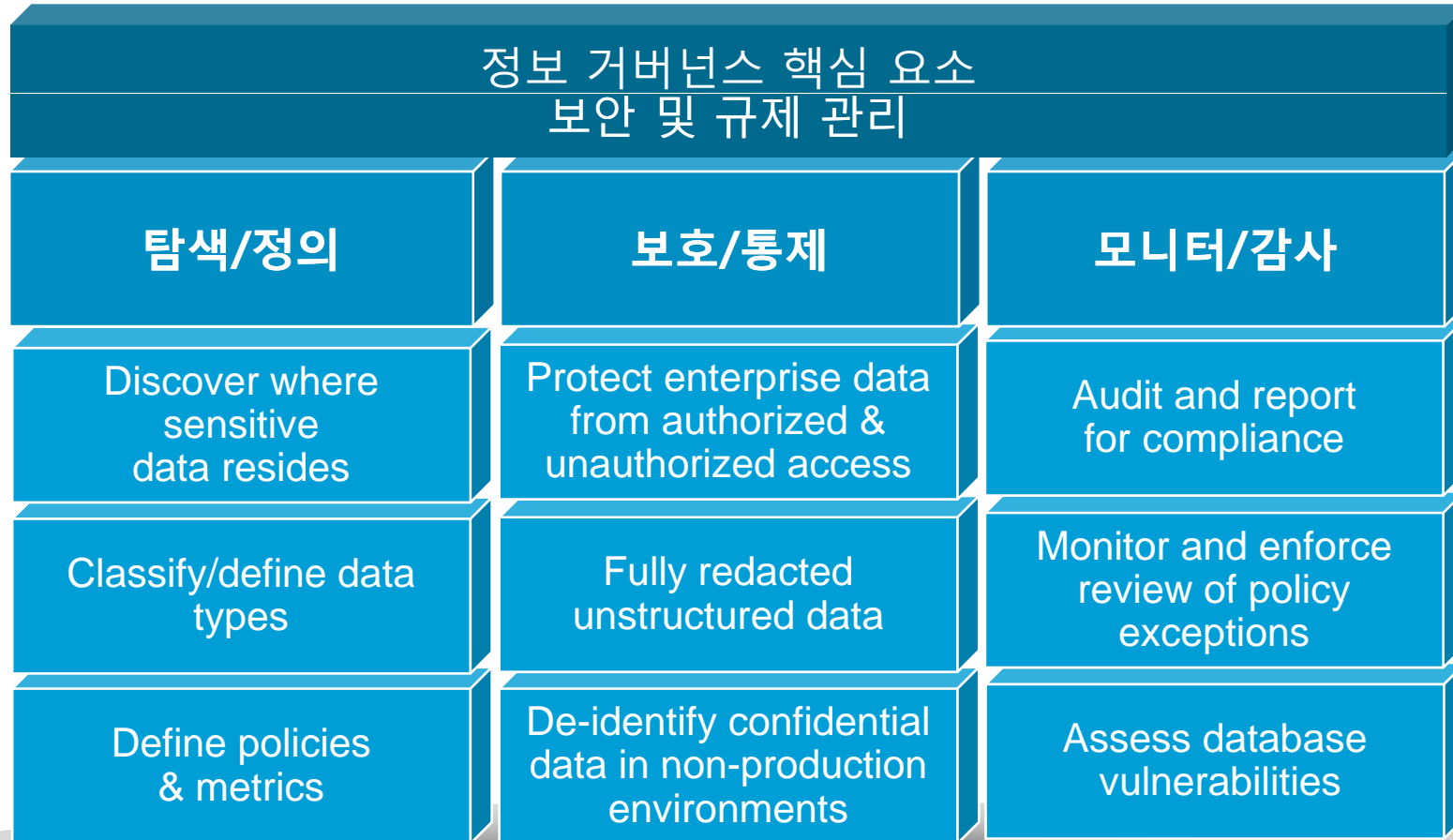
Prebuilt, Automated, Proactive



InfoSphere is the market leader in every category of Information Integration and Governance



정보 거버넌스를 통한 보안/규제 관리 통찰력의 실행



“데이터 보안 전략은 데이터베이스 감사 및 모니터링, 패치 관리, 데이터 마스킹, 접근 제어, 탐색/분류, 그리고 변경 관리를 포함해야 합니다.

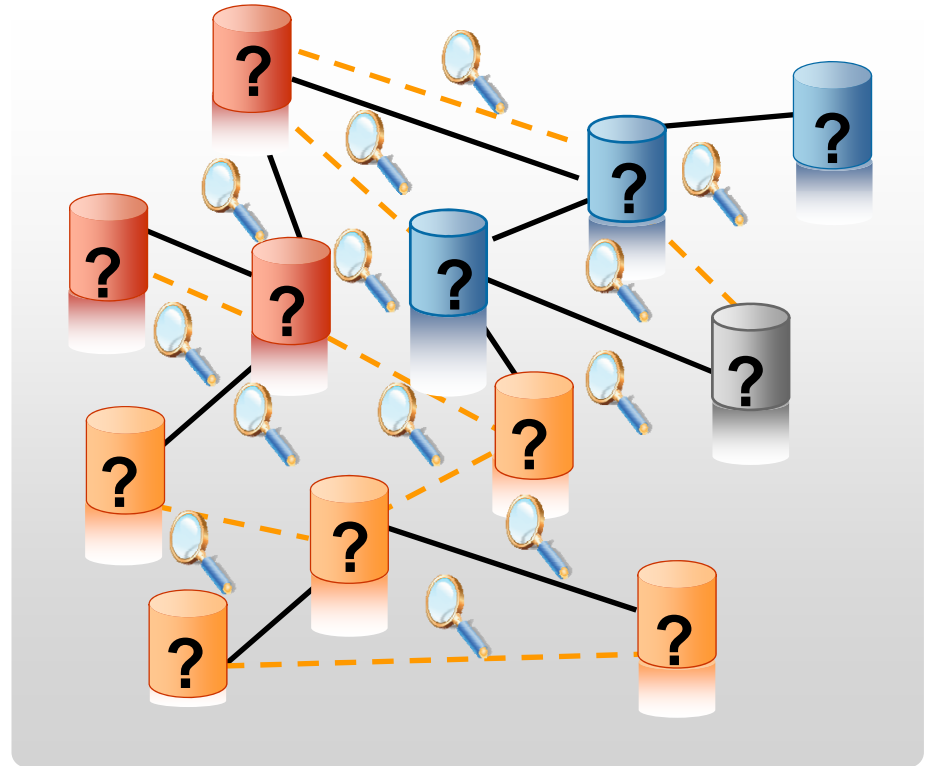
-- Why Enterprise Database Security Strategy Has Become Critical, Forrester Research, Inc, July 13, 2011



정보를 이해하고, 의미를 정의



- 전사내 데이터 소스 및 위치 파악
- 민감 데이터 인식 및 분류
- 연관도 파악
- 정보 수명주기에 걸친 정책 수립



“ 데이터 보안 통제 구현은 데이터의 탐색, 분류, 정책의 수립으로부터 시작된다.
 -- Why Enterprise Database Security Strategy Has Become Critical, Forrester Research, Inc, July 13, 2011



정보의 연관도와 민감 데이터 탐색

정보 거버넌스 - 보안 및 규제 관리

탐색/정의

보호/통제

모니터/감사

민감 데이터 연관도 탐색

System A Table 1		System A Table 15		
Number	Name	Patient	Result	Test
3544600986	Alex Felltham	3802468	N	53
5728150000	Benson Gale	4400745	N	53
3786100000	Bob Smith	5061085	N	53
6783802468	Eileen Ranchman	5567193	N	72
4035567193	Fred Simpson	6123913	Y	47
8037409934	John Smith	6736304	N	34
4306123913	Jamie Slattery	7409934	N	34
9525061085	Jim Johnson	8150928	N	47
4594182715	Martin Aston	8966020	N	34
1288966020				

Patient ID # embedded within another field

System Z Table 25	
Code	Name
53	Streptococcus pyogenes
72	Pregnancy
32	Alzheimer Disease
47	H1N1
24	Dermatamycoses

Compound sensitive data:
Test results could potentially be revealed.

- 연관도와 민감 데이터는 단순 검색으로 확인 불가
 - 필드내에 결합되어 있는 민감 데이터
 - 테이블/시스템 간 연계 관계에 의해서 확인 가능한 민감 데이터
- 테이블 개수/로우 수에 따라 검색 복잡도 증가



비정형 데이터 자동 마스킹

- 데이터 공유가 필요한 기밀 데이터 보호를 위해 비정형 데이터 내 민감 정보 마스킹
- 빠르고, 정확하고, 효율적인 자동화 마스킹 프로세스 지원
- 텍스트, 그래픽 및 양식 기반 문서 형태의 비정형 데이터를 보호
- 데이터 조회에 대한 사용자 역할 정책 제어
- 선택적 수기 리뷰 등 배치 워크플로우 프로세스 자동화

JOHN SMITH
401 MAIN STREET APT 2076
AUSTIN, TX 78745-4548

이름/주소 마스킹

[Redacted Name]
[Redacted Address]
AUSTIN, TX 78745-4548

Patient: Mary Jones
SSN: [SSN]
Phone: [Phone]

...: Associated signs and symptoms include aching joints, redness

...

의사 관점

Patient: Mary Jones
SSN: 123-45-6789
Phone: 786-543-2100

...: Associated signs and symptoms include [Redacted]

[Redacted]

...

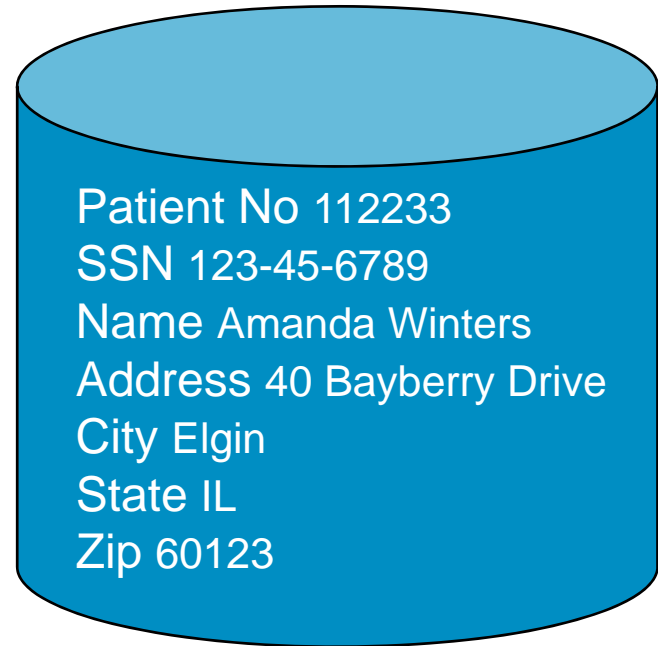
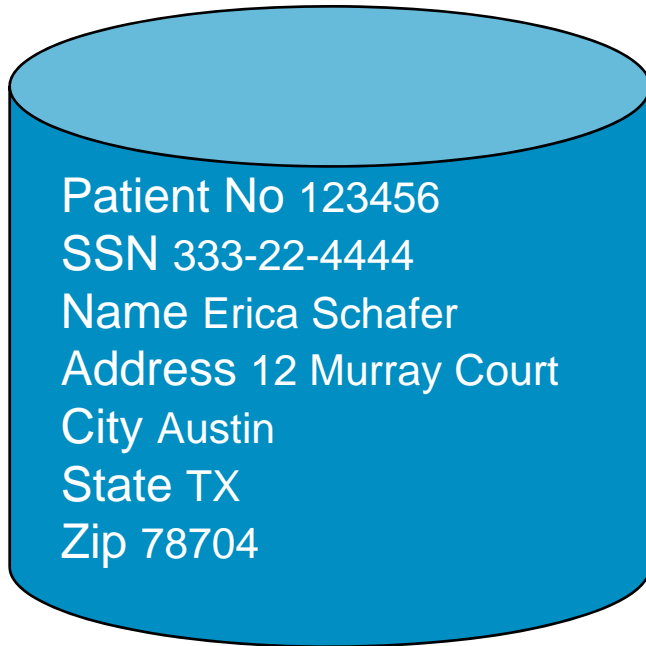
회계 직원 관점



비운영 환경 내 민감데이터 변환

정보 거버넌스 - 보안 및 규제 관리

탐색/정 **보호/통제** 터/감사



- 개발/테스트 환경과 같은 비운영 환경 내 데이터 마스킹
- 비운영 환경 보안 지원
- 컨텍스트 인식 가능한 데이터를 통해 빠른 테스트 프로세스 지원
- 데이터 연관도 인식
- ERP/CRM 애플리케이션 및 개발 애플리케이션 마스킹



비인가자 애플리케이션 조회시 데이터 마스킹

정보 거버넌스 - 보안 및 규제 관리

탐색/정

보호/통제

터/감사

Patient No 123456
 SSN 333-22-4444
 Name Erica Schafer
 Address 12 Murray Court
 City Austin
 State TX
 Zip 78704



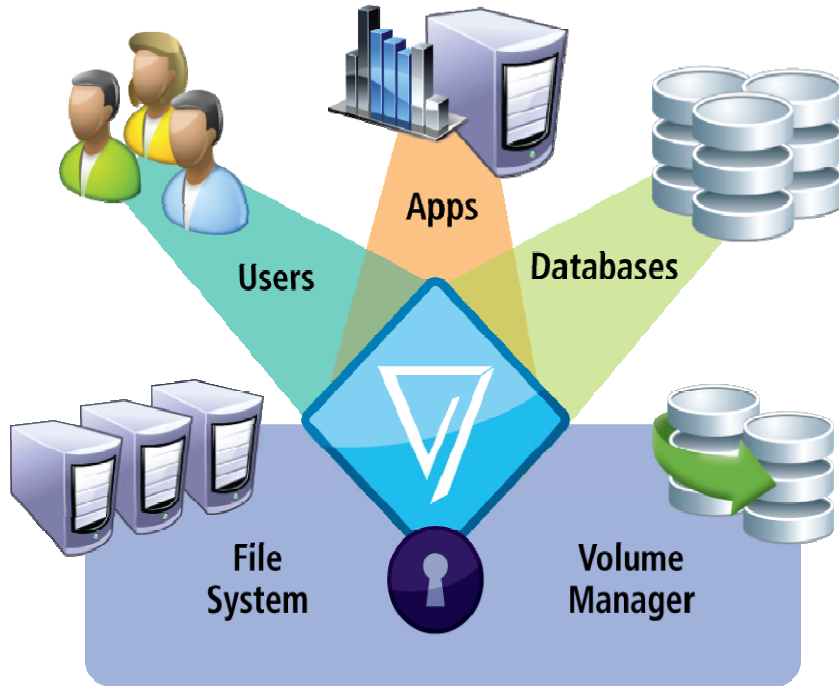
Patient Information			
Patient No.	112233	SSN	123-45-6789
Name	Amanda Winters		
Address	40 Bayberry Drive		
City	Elgin	State	IL
Zip	60123		

- 민감 데이터 활용이 필요한 업무 로직 지원
- 의심 접근 내역에 대한 실시간 데이터 마스킹
- 데이터 접근에 대한 역할 기반 정책 수립 지원



온라인/오프라인 데이터 암호화

정보 거버넌스 - 보안 및 규제 관리
 탐색/정 **보호/통제** 터/감사



- 암호/복호화를 위한 키 관리
- 권한 기반 암호/복호화 지원
- 원본 데이터의 안전한 보관

John Smith
 401 Main Street Apt 2076
 Austin, TX 78745-4548

암호화



*&^\$!@#)(
 ~|” +_)? \$%~::~>>
 %^\$#%&, >< <>?_)^-^%~~

*&^\$!@#)(
 ~|” +_)? \$%~::~>>
 %^\$#%&, >< <>?_)^-^%~~

복호화



John Smith
 401 Main Street Apt 2076
 Austin, TX 78745-4548

Personal identifiable information is encrypted making it meaningless without a proper key.



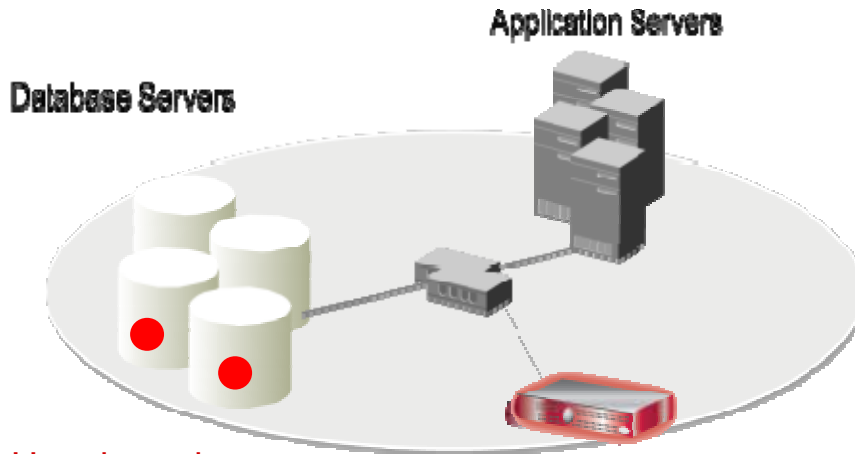
실시간 모니터링/감사 및 DB 보호

정보 거버넌스 - 보안 및 규제 관리

탐색/정의

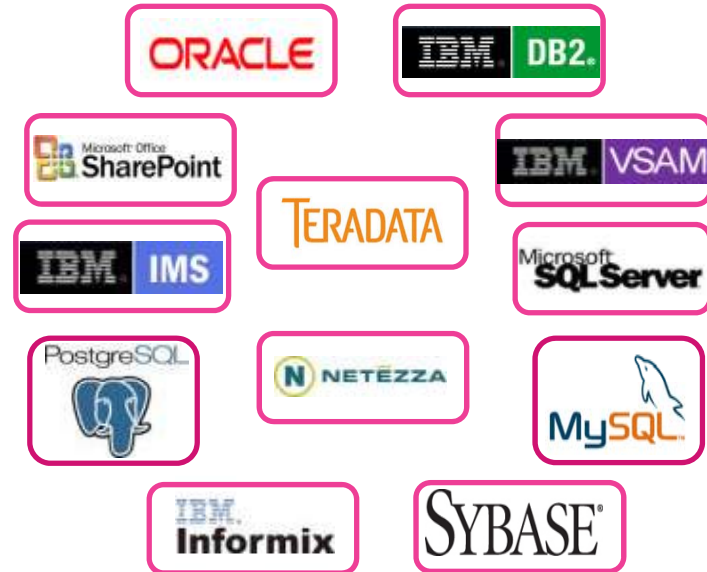
보호/통제

모니터/감사



Host-based Probes (S-TAPs)

Collector



- 지속적으로 모든 데이터 베이스 활동들을 모니터(Super-user의 local 접근포함)
- 이기종 환경 , Cross-DBMS 솔루션지원
- DBMS 나 어플리케이션 변경이 필요 없음
- 최소 성능 영향 (약 5% 이하)

- DBA 나 침입자에 의한 로그 위변조방지
- 자동화된 컴플라이언스 레포팅,결제상신 (SOX, PCI, NIST, etc.)
- 실시간 정책 & 감사 세분화
 - Who, what, when, where, how

DB보안에서 빅 데이터 보안으로

정보 거버넌스 - 보안 및 규제 관리

탐색/정의

보호/통제

모니터/감사

DATA

NETEZZA

IBM DB2

IBM Informix

IBM IMS

IBM VSAM

Big Data Environments

이기종 지원

hadoop

Microsoft Office SharePoint

ORACLE

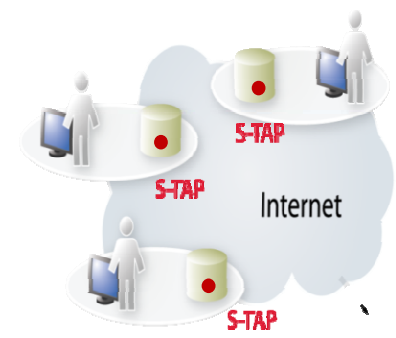
MySQL

Microsoft SQL Server

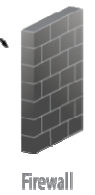
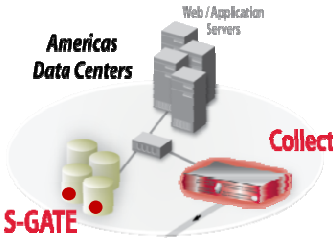
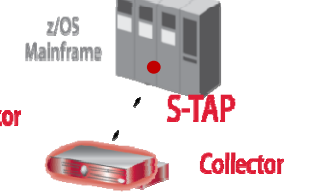
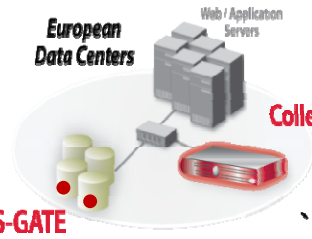
SYBASE

TERADATA

PostgreSQL



Remote Locations & Outsourcers



Integration with LDAP, IAM, SIEM, TSM, Remedy, ...



Hadoop 모니터링

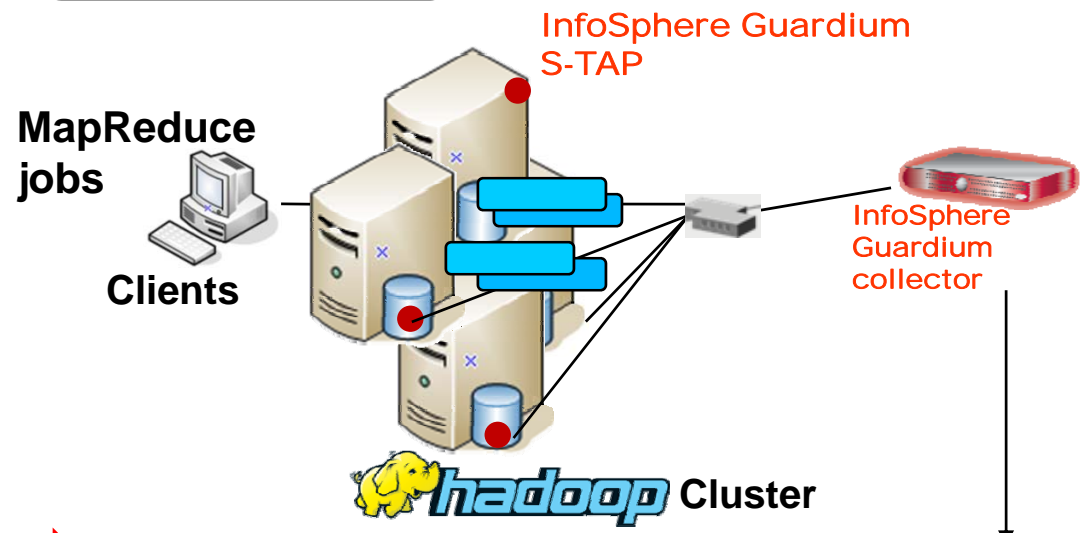
정보 거버넌스 - 보안 및 규제 관리

탐색/정의

보호/통제

모니터/감사

트랜잭션 로그 내역이 Collector로 전송됨



- ### Big Data Security Design Goals
- Minimal impact to the Big Data server resources
 - No Big Data configuration changes
 - Separation of Duties from privilege users
 - Granular audit trail of Big Data activity
 - Real-time alerting
 - Minimal impact to the network
 - 100% Big Data visibility
 - Heterogeneous support



Sensitive data alert!

Hadoop - Unauthorized MapReduce Jobs

Start Date: 2012-09-05 11:06:20 End Date: 2012-09-06 14:06:20
 Alases: OFF Sqt: LIKE %
 Username: LIKE %authusr% name: LIKE %

Timestamp	Server Type	Client IP	Server IP	MapReduce User	MapReduce Name	MapReduce Job	Source Program
2012-09-05 14:01:26.0	HADOOP	9.70.146.2119	70.146.211 (node=/user/data/output_unauth/_logs/history/0.0.0.0_1345143729831_job_201208161502_0018_authusr_hack', client=DFSCClient_-881434930')	authusr	hack	ob_201208161502_0018HADOOP CLIENT PROGRAM	
2012-09-05 14:01:10.0	HADOOP	9.70.146.2119	70.146.211 (path=/user/data/output_unauth/_logs/history/0.0.0.0_1345143729831_job_201208161502_0018_authusr_hack', permissionMask=493,client=DFSCClient_-881434930',overWrite=true,replicationFact=1,numBytes=67108864)	authusr	hack	ob_201208161502_0018HADOOP CLIENT PROGRAM	

InfoSphere Guardium reporting and alerting



정보 거버넌스 기반의 전사 데이터 보안

DB 내 민감한 데이터의 식별

민감한 데이터 및 전사 데이터 연관도 자동 탐색

- 업무 데이터 그룹 정의를 위한 논리적인 데이터 연관도 파악
- 민감한 데이터 자동 탐색
- 데이터 프로토타입 및 로직에 대한 리버스 엔지니어링

InfoSphere Discovery/InfoSphere Guardium

비운영 환경 내 실 데이터 변환

개발, 테스트 DB 내 민감한 데이터 보호

- 테스트 시 민감한 데이터 보호를 위한 베스트 프랙티스
- 실제적인 유사 데이터를 사용하여 일관성 있는 정보 변환
- 비운영 환경 내 내부유출자 및 외부 침입자의 정보 유출 방지

InfoSphere Optim Data Masking

DB 모니터링 /취약성 평가

이기종 환경 내 DB 정보 보호를 위한 필수 안전망 제공

- 지속적인 실시간 데이터베이스 활동 모니터링
- 미승인 접근 및 유해 활동 탐지를 위한 정책 기반 DB 접근제어
- DB 보안 취약성 평가
- DB 변경 내역 감지 및 접근 차단

InfoSphere Guardium DAM/VA Solution

데이터 암호화

고성능 데이터 암호화

- 애플리케이션 변경 없이 테이블 단위 암호화
- DBA와 보안담당자의 직무 분리
- 통합적이고 중앙 집중적인 정책 및 키 관리

InfoSphere Guardium Data Encryption

비정형 데이터 문서 내 주요 정보 변환

문서/양식 내 민감한 데이터에 대한 보호

- 텍스트, 그래픽, 양식 기반 데이터에 대한 정보 변환 지원
- 자동화를 통한 효율성 향상 및 수작업 대비 비용 절감 효과
- 정책 기반 사용자별 데이터 조회 통제

InfoSphere Guardium Data Redaction

컴플라이언스 및 다양한 법적 규제 준수



여러분의 현재는 어디입니까?

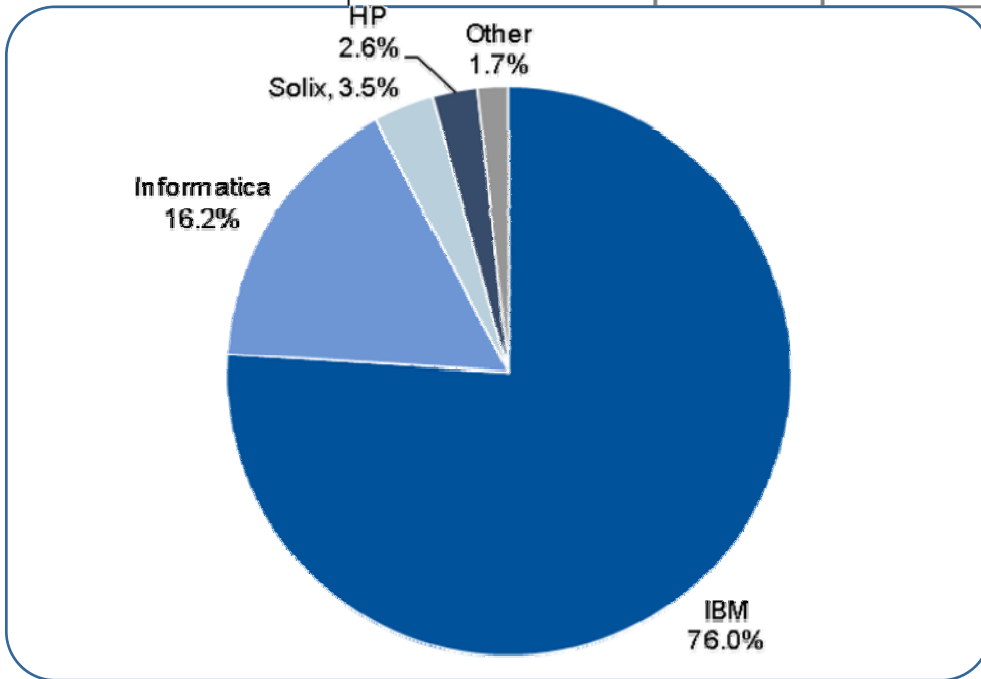


InfoSphere Optim: 데이터 관리 시장을 선도하는 마켓 리더!

- ❖ Optim은 IDC의 '월드와이드 데이터베이스 개발 및 관리 툴 시장' 부문 보고서와 가트너의 '아카이빙 소프트웨어 시장' 부문 보고서에서 각각 56%와 76%에 이르는 시장점유율을 차지한 전세계 1위 솔루션입니다.

Worldwide Database Archiving and Information Life-Cycle Management Revenue by Vendor, 2007-2009 (\$M)

	2007	2008	2009	2008 Share (%)	2009 Share (%)	2008-2009 Growth (%)
IBM	44.3	46.1	50.4	54.0	56.2	9.5
HP	12.0	14.1	13.4	14.1	13.4	0.0
Solix	9.8	11.6	10.9	11.6	10.9	-1.2
Other	3.7	4.1	4.1	4.1	4.1	5.9
Informatica	3.5	4.1	3.9	4.1	3.9	0.2
Total	79.4	87.9	88.5	87.9	88.5	6.0
	10.3	12.1	11.5	12.1	11.5	0.2
Total	89.8	100.0	100.0	100.0	100.0	5.3

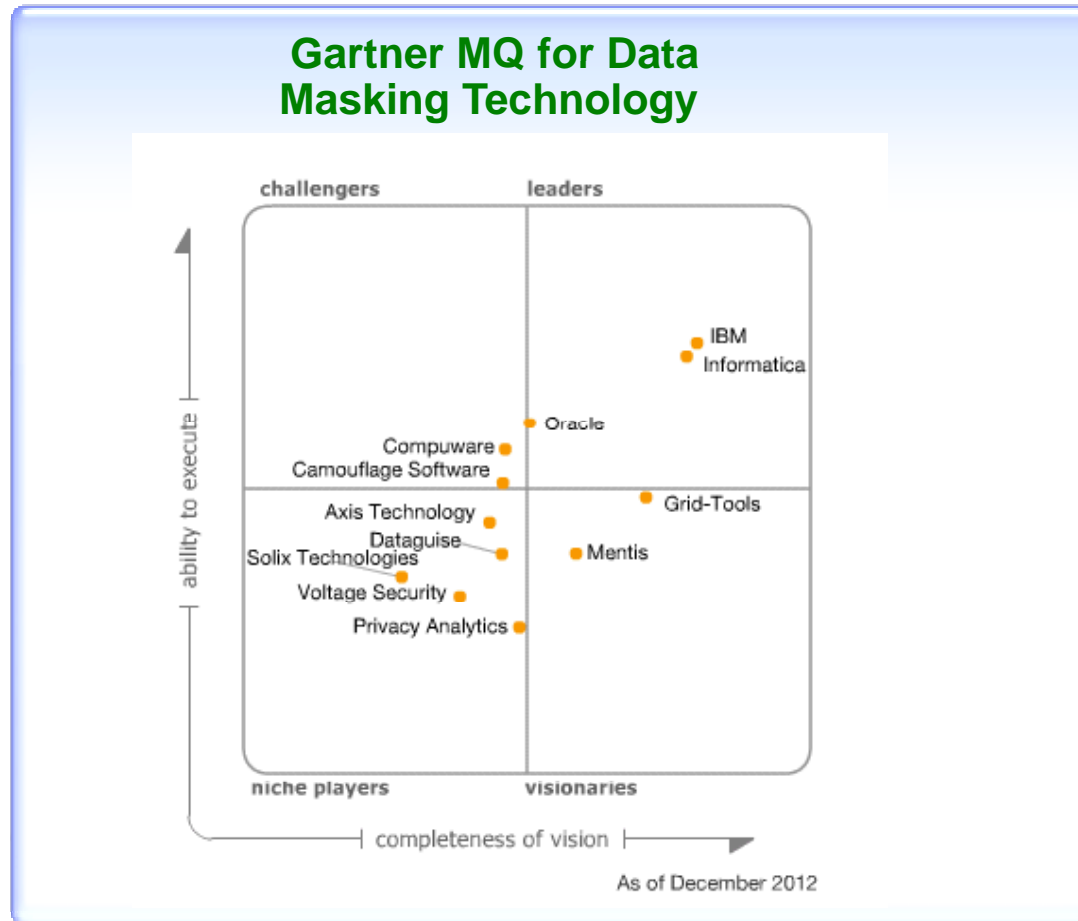


Source: Gartner, Inc., "Market Trends: World, Database Archiving Market Continues Rapid Growth, 2011", S. Childs/A Dayley, September 2011



InfoSphere Optim: 데이터 마스킹 시장을 선도하는 마켓 리더!

❖ Optim은 가트너의 데이터 마스킹 영역의 Magic Quadrant에서 입증된 리더 솔루션입니다.

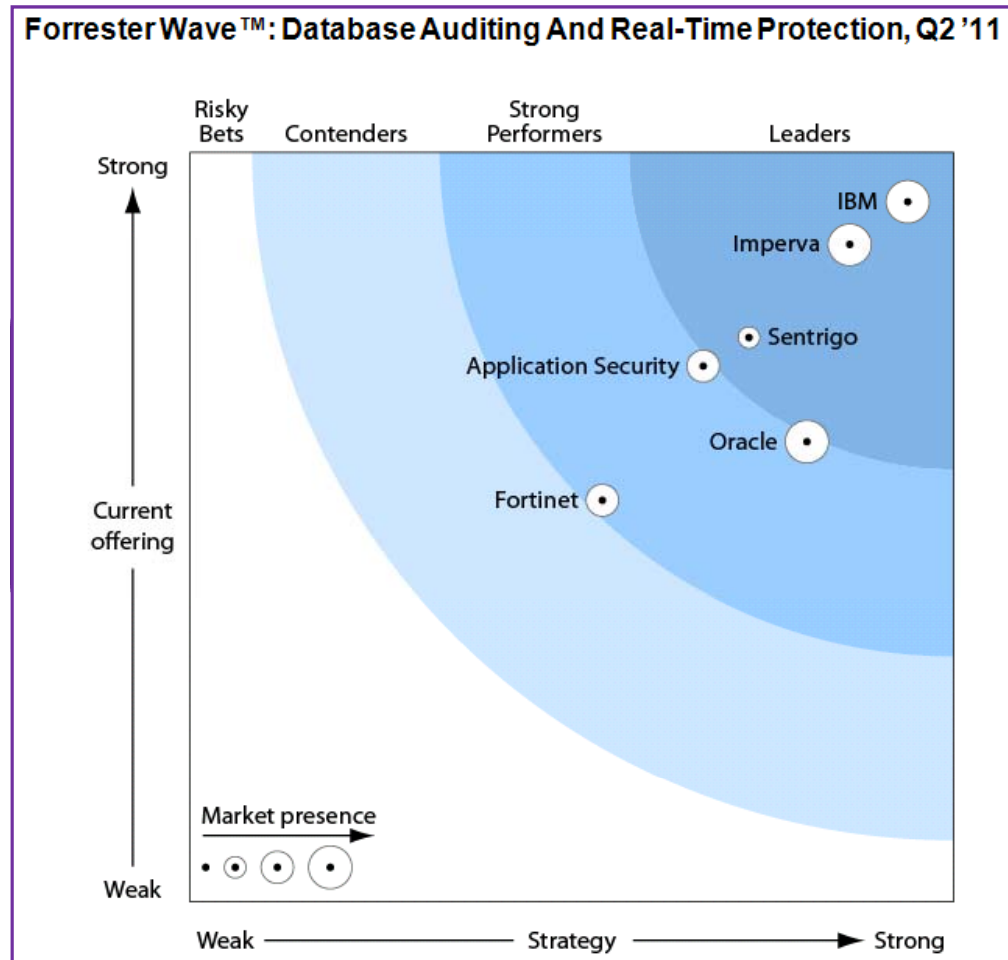


The Magic Quadrant is copyrighted 2012 by Gartner, Inc. and is reused with permission. The Magic Quadrant is a graphical representation of a marketplace at and for a specific time period. It depicts Gartner's analysis of how certain vendors measure against criteria for that marketplace, as defined by Gartner. Gartner does not endorse any vendor, product or service depicted in the Magic Quadrant, and does not advise technology users to select only those vendors placed in the "Leaders" quadrant. The Magic Quadrant is intended solely as a research tool, and is not meant to be a specific guide to action. Gartner disclaims all warranties, express or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.



InfoSphere Guardium: DB 보안 시장을 선도하는 마켓 리더!

- ❖ Guardium은 Forrester Wave에서 입증된 DB 감사 및 실시간 보호 영역의 리더 솔루션입니다.



2007

Source: The Forrester Wave™: Database Auditing And Real-Time Protection, Q2 2011, May 6, 2011. The Forrester Wave is copyrighted by Forrester Research, Inc. Forrester and Forrester Wave are trademarks of Forrester Research, Inc. The Forrester Wave is a graphical representation of Forrester's call on a market and is plotted using a detailed spreadsheet with exposed scores, weightings, and comments. Forrester does not endorse any vendor, product, or service depicted in the Forrester Wave. Information is based on best available resources. Opinions reflect judgment at the time and are subject to change.



데이터 관리 진단 워크샵 Client Value Engagement

CVE 워크샵 개요

- ILM을 비롯한 데이터 관리를 통해 얻어지는 비용 절감과 추가적인 매출 창출에 대한 비즈니스 이윤을 정량화할 수 있도록 분석하는 워크샵
- 전략적인 파트너십 고객을 위해 무료로 제공
- 일정: 평균 2-3주 소요

Tasks	Week 1	Week 2	Week 3
Client Introduction	[Progress Bar]		
Develop CVE Charter		[Progress Bar]	
Conduct CVE Interviews		[Progress Bar]	
Validate CVE Analysis		[Progress Bar]	
Final Presentation			[Progress Bar]

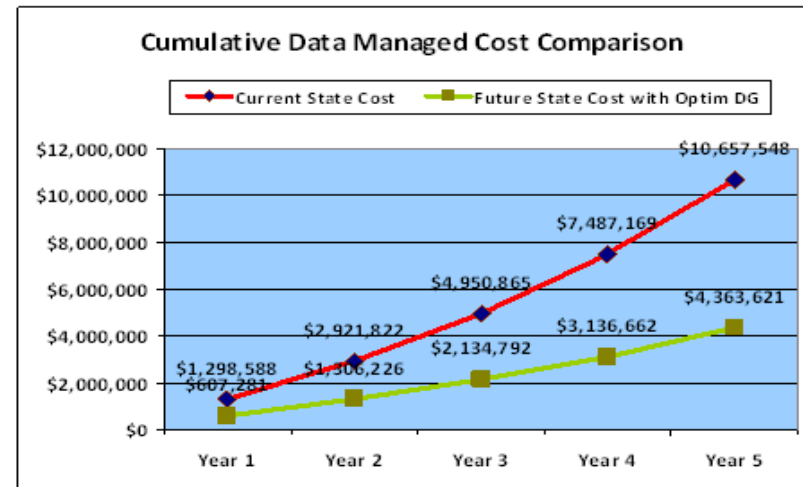
- 진행 프로세스



CVE 분석 보고서 샘플

- 분석 영역
 - ✓ 빅 데이터
 - ✓ 정보 수명주기 관리
 - ✓ DB 보안
 - ✓ 데이터베이스
 - ✓ 정보 통합

3-Year Savings: \$3.060 Million



Thank
YOU

