AUSTRIAN INSTITUTE OF TECHNOLOGY

Database Preservation Evaluation Report SIARD vs. CHRONOS

Preserving complex structures as databases through a record centric approach?

Preserving Transactional Data Briefing Day
Joint Workshop of the
Digital Preservation Coalition and
UK Data Service
17th of March 2016
WCIT, London

Andrew Lindley
Research Engineer
Digital Safety & Security Department
Digital Insight Lab (DIL)
andrew.lindley@ait.ac.at
@lchwardort

AIT Austrian Institute of Technology

- Largest Austrian non-university research facility approx. 1200 employees and total operating income of 118,8 Mio. Euro

- AIT provides **research and technological development** to realize basic innovations for the **next generation of infrastructure related technologies** in the fields of health & environment, energy, mobility and safety & security. These technological research areas are supplemented by the competence in foresight & policy development.

- As a national and international network node at the **interface of science and industry** AIT **enables innovation** through its scientific-technological expertise, market experience, tight customer relationships and high quality research infrastructure.

Our research group the "**Digital Insight Lab**" is based in the department of Digital Safety and Security (DSS) and we provide Services in the research field of **Data Science** within two specific Business Cases

- "Data Engineering and Analytics" (machine learning, prediction and recommendation)
- "Archiving and Preservation"

European Research Projects

- SCAPE: Scalable Preservation http://scape-project.eu/
- E-ARK: European Archival Records and Knowledge Preservation http://www.eark-project.com/

What‘s the motivation

Aren’t databases great products anyway to keep data?

- Database Dump != Database Archive
- **Solutions for database archiving are not part of a standard relational database systems**. According to Forrester only 15% percent of business data are actively required to serve a company’s day-to-day requirements while the vast amount of data could already be moved into an archived state.

- *[..]Terabyte-size transactional databases are harder to manage, increase costs for hardware capacity and database licenses, and drive up requirements for database administrators (DBAs). Yet **85% of production data is inactive**, so information and knowledge management professionals should devise a database archiving strategy that moves inactive data to lower-cost storage and servers, thus improving the manageability, performance, and security of critical production applications[..]*

- Typical data life-cycle can be categorized in
    - Active State (data is generated, modified as part of the production system)
    - Archiving State (data no longer altered but required for fulfilling business processes)
    - Long-Term Archiving (only selected parts of a dataset are kept for retention)

What‘s the motivation

Structured data archiving technologies help IT leaders **retire legacy applications**, **reduce capital and operating expenses**, and meet governance and **compliance requirements**. Gartner evaluate vendors offering products and services that provide archiving for databases and data from enterprise applications.

Structured data archiving is the ability to index, migrate and protect application data in secondary databases or flat files typically located on lower-cost storage for policy-based retention. It makes data available in context and protects it in the event of litigation or an audit.

Structured data archiving addresses:
Storage optimization — It can reduce the volume of data in production and maintain seamless data access. The benefits of using this technology include reduced capital and operating expenditures, improved information governance, improved recoverability, lower risk of regulatory compliance violations, and access to secondary data for reporting and analysis.
Governance — The technology preserves data for compliance when retiring applications. Structured data is often transactional and related to financial accounts or back-office functions (for example, HR, patient enrollment in healthcare and other use cases that might be regulated) that require information governance, control and security, along with the ability to respond to related events such as audits, litigation and investigation. These and other requirements, such as maintaining information context, can prevent organizations from moving data to lower-cost tiers of storage, or adopting other do-it-yourself approaches.
Cost optimization — Structured data archiving and application retirement can result in significant ROI. Structured data in legacy systems, ERP and databases accumulates over years — and, in some cases, over decades — driving up operational and capital expenses.
Data scalability — The technology can manage large volumes of nontraditional data resulting from newer applications that can generate billions of small objects. Scalability to petabytes of capacity is required in these cases.
The desire to leverage archives as a secondary data store for big data analytics is driving the growth of the structured data archiving market. Newer market participants are offering alternate ways for managing archived data that involve virtual copies of databases, extreme compression and native SQL access.

[Source Gartner]

Legal Requirements for Electronic Records Retention in Western Europe

- National Legislative Regulation on data retention
- Taxation, Trade, Liability, Social Insurance, Employment, Data protection law…
- EU regulations and council directives
- Industrial Sector Regulations
- Retention, Accountability, Proof of Evidence: 1-40+ years

www.arma.org/go/prod/v4980 (395 $)

Market Overview

Based on Gartner's estimates, the size of the **structured data archiving and application retirement** market is **$263 million** — with additional related product **revenue at $27 million** — and growing at a compound **annual growth rate of 10%.** The use of this technology has long been viewed as a cost avoidance measure to contain operational and capital expenditures related to data growth, and as a measure to improve factors such as application performance. The market is changing and expanding due to growth in data, application retirement, information governance and big data analysis opportunities.

- **Application Retirement as a Leading Use Case for Structured Data Archiving**
- **The Trend Toward Big Data Analytics and Petabyte-Scale Archives**
- **Growing Importance of Information Governance in Structured Data**

SIARD - CHRONOS

SIARD

- Software Independent Archiving of Relational Databases
- Owned by Swiss National Archives (BAR)
- Both open format to express db archives + software product
- Available under closed source license.
- Quite popular in the GLAM world

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

CHRONOS

- Commercial product owned by CSP
- Emerged through a joint research cooperation at University of Landshut
- CSP is an SME approx 50 employees
- Customers: *Deutsche Telekom, Lufthansa Technik Logistik, E.ON, ING DiBa,, BMW, Audi, etc*

Other Solutions you might are familiar with:

- Princeton Softech -> IBM Optim Solutions, RODADB, HPAIO, Informatica, etc.

How to Evaluate?

Gartner: Magic Quadrant for Structured Data Archiving and Application Retirement

Evaluation Criteria Definitions

Ability to Execute

- Product/Service
- Overall Viability
- Sales Execution/Pricing
- Market Responsiveness/Record
- Marketing Execution
- Customer Experience
- Operations

Completeness of Vision

- Market Understanding
- Marketing Strategy
- Sales Strategy
- Offering (Product) Strategy
- Business Model
- Vertical/Industry Strategy
- Geographic Strategy

Source Gartner, June 2015

https://www.gartner.com/doc/reprints?id=1-2HYPOQ8&ct=150616

Experiment Setup and Scope

Based on AIT Case Study for the Austrian Ministry of Defense

- Report: 3 major sections
  1. Generic Evaluation of Tools and Features
  2. ISO 25010:2011 driven evaluation of software quality aspects based on TR9126 ‚**quality in use' metrics** in areas of efficiency, productivity, security, satisfaction in a specific context and staging environment
  3. Interpretation of research results based on customer requirements

- Testdata extremely important, 3 types
  - Administrative, scientific and document management databases

- Setup: virtualized Windows 7 standard Hardware Desktop, Tools and all dependencies locally, Oracle 11gR2 DB and TPC-C ‚Entry-Order' records enriched with BLOB and CLOB data

- **No aim of providing benchmark information** but rather accompanying documentation of technical features and USPs – no entitlement of functional completeness.

Evaluation Criteria

- Supported Preservation Scenarios
- Exported Elements of an Archived Database
- Pre- and Postprocessing via Database Scripts and Markertables
- Data Retention and Data Controls
- Support of UDTs and Oracle Specifics
- Rights, Roles and User Management
- Archive Data Access and Performance
- Syntactic and Semantic Data Changes
- Existing APIs and Interfaces
- Scalability and Limitations
- Risk Behavior and Dependencies
- Referential Dependencies
- Standard and Compliance
- Data Exchange Formats
- Structure, Setup and Size of the physical Archive
- Specification of Information Lost
- Installation and Delivered Components

Findings

Full Report: http://t.co/yZFvj4xxQv

Database Preservation Evaluation Report - SIARD vs. CHRONOS

Preserving complex structures as databases through a record centric approach?

Andrew Lindley
AIT - Austrian Institute of Technology GmbH
Donau-City-Strasse 1, 1220 Vienna, Austria
andrew.lindley@ait.ac.at

ABSTRACT

Preserving information systems is one of the greatest challenges in digital preservation. In this paper we outline the existing strengths and shortcomings of a record-centric driven preservation approach for relational databases by lining up a state-of-the art industry database archiving tool CHRONOS[1] against SIARD[2] one of the most popular products in the GLAM (galleries libraries archives museums) world. A functional comparison of both software products in the use cases of database retirement, continuous and partial archiving as well as application retirement is presented. The work focuses on a technical evaluation of the software products - organizational and process aspects of digital preservation are out of scope. We explain why preserving complex structures as databases through a record centric approach does not only depend on the amount of information captured in the preservation package and present a brief overview on available functional aspects in CHRONOS that help to address the challenges of application decommissioning. The paper at hand presents the results of a case study which was undertaken 2012 at AIT - Austrian Institute of Technology GmbH.

Categories and Subject Descriptors

H.2.m [Database Management]: Database Applica-

Keywords

Digital Preservation, Database Archiving, Case Study, Technical Evaluation, Decommissioning, Application Retirement

1. INTRODUCTION

Sustained information to our scientific and cultural heritage world is stored digitally. The term digital preservation (DP) summarizes methods and techniques to secure long-term access to digital information. Every information management system, data warehouse, or even the simplest online webstore is backed by a database system. For the last decades relational databases have been the dominant technology in this area mainly due to broad vendor adoption and acceptance of the SQL standard for the relational model. ACID (Atomicity, Consistency, Isolation, Durability) provides principals governing how changes are applied to a database. In the decade of big data some of these principles are loosened with respect to high data volumes and high traffic throughput and niche products as NoSQL databases, key value and tripple stores found their place.[1]

Within the last ten years the digital preservation community was able to achieve a solid understanding of issues and provided solutions and guidance in the domain of document

Evaluation Results

SIARD CHRONOS

1. Supported Preservation Scenarios

Three classification scenarios

1. **Database Retirement:**
   - DB independent transformation, understandability of the physical archive, SQL data access, etc
2. **Continious and Partial Archiving:**
   - Inkl. schema schanges over time, data retention, etc.
3. **Application retirement:**
   - Support for recreating business objects, reporting, data access roles and programmatic access, middleware, etc.)

CHRONOS

- Database Retirement + Continious and Partial Archiving + Application Retirement

SIARD

- Database Retirement

Evaluation Results

SIARD CHRONOS

2. Exported Elements of an Archived Database

An **RDBMS is a complex product** which consists out of Tables, Views, Materialized Views, Indices, Packages, Triggers, Stored Procedures, Functions, Sequences, Scheduler, Check Constraings and Triggers, Queues, Database Links, user Management Access Privileges, and Roles – just to mention the most important constructs

- Which db elements are extracted into an archive, which ones are missed?
- Which elements remain functional after re-importing
- Which ones are preserved but solely serve the purpose of documentation?

Evaluation Results

SIARD CHRONOS

2. Exported Elements of an Archived Database

CHRONOS

- Main Focus: exporting Primary Data and Datatypes
- Supported:
  - Tables, Views, Indices, Packages, Procedures, Functions, Triggers, Sequences, Materialized Views, Scheduler and Check Constraints are supported elements when transferring data into a database archive
- Unsupported:
  - DB Links,
  - Jobs (depricated)
  - user management and definition of roles (not accessible)
    - Most often not depicted at DB level anyway (+LDAP integration)
- When Re-Imported: Triggers, Procedures, Views remain unsupported elements when re-importing data (potential to cause serious damage)

Evaluation Results

SIARD CHRONOS

2. Exported Elements of an Archived Database

SIARD

- Focuses on preserving primary data
- Exclusively supports archiving of core SQL:1999 elements
- Procedures and Functions are minimally supported and documented in a SIARD archive
- Unsupported SQL:1999 elements:
  - Triggers (only useful in live DBs where they can occur)
  - Check Constraints (supported by SIARD, but mostly not accessible)
- Not Defined in SQL:99:
  - Materialized Views (temporary tables)
  - Packages
  - Indices (not defined as database elements but only as performance enhancers)
- User and Roles are archived
- User Defined Datatypes (UDTs) not archived -> not available when SIARD dev started
- Backward Compatibility a major requirement!
- ➔ when restoring a db-archive into RDBMS only tabular content is restored.

Evaluation Results

SIARD CHRONOS

3. Pre- and Postprocessing via DB Scripts and Markertables

In the process of creating an archival package, especially in the scenario of partial and ongoing archiving, it might be necessary to execute pre- and post processing steps on the database as for example preparation or cleanup tasks. Supporting a smooth and integrated continuous archival workflow might require logging some kind of state or placing process markers within a production system. To which degree do the tools offer support for **interacting with a production environment** as executing pre- or post processing scripts **or documenting archival state within the database itself**?

CHRONOS

- Interaction via: Shell Commands, DB Scripts, Marker Tables (granularity table level)

SIARD

- By design **never writes to a database** (read only permissions sufficient)
- Workaround of running scripts via batch file via sqlplus for static pre- and post processing possible

Evaluation Results

	SIARD	CHRONOS
4. Data Retention and Data Controls	■	

Due to legal regulations for example on handling of personal or sensitive data it might be required **to keep and/or delete records after a given period of time** from an archive. Other forms of data retention concern the **periodical refreshment of expiration dates**. The following questions are taken into account:

- Do the systems easily allow to classify and separate archival data from master data items (customer records, etc.).
- Which mechanisms are in place to handle data retention and deletion controls and at which degree of granularity.
- Is it for example possible to connect to external storage systems that ship with built in mechanisms for data retention?
- Which security mechanisms for supervising deletion control mechanisms are in place?

SIARD CHRONOS

Evaluation Results

4. Data Retention and Data Controls

CHRONOS

- ships with modules for **creating archival data retention policies** and fully applies to the requirements of implementing **legal hold** within a repository.
- Policies enforced on exported data (across different storage mechanisms)
- Allows maintaining distributed archival packages
- Adapters for dedicated storage facilities as EMC^2 but also standard file systems
- Deletion of data: two step approval process

SIARD

- Never deletes information
- Data integrity guaranteed by SIARD Suite
- Everything beyond: up to the archivist

SIARD CHRONOS

Evaluation Results

5. Support of UDTs and Oracle Specifics

Clarifies the degree of support for custom Oracle database features such as user-defined datatypes (UDTs) or Oracle specific extensions as PL/SQL, Oracle Spatial and custom built applications with Oracle Forms.

CHRONOS

- Archive Oracle user-defined datatypes in a preliminary form, further support announced by CSP.
- UDTs seen problematically given their inconsistency and incompatibility across different version of Oracle DBs. JDBC driver support required. No cross vendor mapping possible.
- Chronos falsely reactivates disabled check of foreign key constraints when re-importing
- Makes use of native dialects for data export (when querying the db systems)
- PL/SQL (Procedural Language SQL) not supported for querying or archival purpose
- Act as Middleware (APIs) for Oracle Forms applications.

SIARD

- Plans to enhance SIARD format to accommodate UDTs in a SQL99 standardized way
- Oracle specifics: one exception supported: table/column comments -> metadata

Evaluation Results

SIARD CHRONOS

6. Rights, Roles and User Management

Access controls and user management is a core component of a running database environment. This section focuses on the capabilities of the tested database archiving products to offer rights, roles and user management functionality on top of the extracted database archive.

CHRONOS

- Mature user, rights and access management layer out of the box. Tightly integrated throughout all delivered CHRONOS components, highly customizable
- Granularity: protect sensitive data in the archive on db column level
- LDAP integration possible

SIARD

- No user, rights, permission management or custom application views
- Makes use of underlying RDBMS user management component
- Visibility and rights of archiving user determines scope of harvested data -> full export

SIARD CHRONOS

Evaluation Results

7. Archive Data Access and Performance

CHRONOS

- Possibility to execute SQL statements directly on top of archived data on file system
  - With performance measures comparable to standard database systems
  - Hybrid approach**: Custom SQL92 interpreter**, **global search index**, local Btree index on column level as well as **H2/hsqldb in-memory db** systems for SQL JOIN operations
  - Various tweaks for performance fine-tuning (e.g. package split size, etc.)
- Creates database export in a vendor independent, generic, human understandable format (just data and corresponding schematic structure) + added value on top of physical archive which is crucial for management and use of such data.
- Chronos consists of 6 independent server modules (indexing, search, archiving, etc.) a middle ware layer and a different very sophisticated GUI applications on top.
- Support for revisions of data and search operations within revisions, even if database schema has been modified in the meantime

SIARD CHRONOS

Evaluation Results

7. Archive Data Access and Performance

SIARD

- **Performance: of command line tools is solid for large datasets!**
- SiardEdit: Graphical user interface application for exploring archive files. Allows to display, browse, sort primary data and to add or chance archival metadata (but not primary data) -> not suitable for complex research within large archives.

Evaluation Results

SIARD CHRONOS

8. Syntactic and Semantic Data Changes

In the case of **continuous archiving partial datasets remain within the production environment**. Therefore a common scenario which needs to be dealt within is the reaction **to syntactic and semantic changes over time**. Which form of support or traceability do the systems provide for this kind of temporal changes?

CHRONOS:

- Structural changes in the schema as adding additional columns, are automatically detected by CHRONOS
- Data is exported into a separate revision. User is given tools to administrate complex changes.
- Support to automatically transform deposited data via customizable operations for an entire revision (written in Java, full richness of JDK data manipulation!)
- **The actual physical archive however stays authentic, consistent and untouched -> changes are only reflected in CHRONOS middleware**

SIARD:

Cannot be evaluated agains this use case as it exclusively offers support for database retirment and the tools are not designed to cope with semantic or syntactic changes of the underlying data.

Evaluation Results

SIARD CHRONOS

9. Existing APIs and Interfaces

The scenarios archiving, data access and search were evaluated with respect to available programming interfaces.

CHRONOS:

- All server modules offer programmatic access via JDBC, Java RMI and web-services and allow deep system interaction.
- A CHRONOS JDBC class 4 driver provide unified access. Data manipulation via JDBC is not possible.
- Programmatic support of entire process has been tested:
  - setting up, running a db export, re-importing a CHRONOS archive in a db system
- Out of the box support for variety of external facilites as job schedulers, storage solutions, etc.

SIARD:

- Scripting solutions via SiardFromDb and SiardToDb possible -> external configuration file for settings possible e.g. for automation via cron job scheduling.

Evaluation Results

SIARD CHRONOS

11. Risk Behavior and Dependencies

What is the degree of underlying dependencies for a given database archive in subject to system dependencies, vendor / tool locking, or similar objectives?

→Both tools follow the approach of clearly separating the composition and description of the data structure from the actual primary data - this is also reflected on file system level.

CHRONOS:

- Implicit data export format (*)
- Describes Structure in XML + provides a fully interpretable XSD schema file
- Content is store in delimiter file. Binary Files are linked with pointers and can be held externally. In theory all information to properly read and interpret data and therefore manually revive it in case of a vendor crash is available without direct dependencies. **In practice: Non trivial task!**
- Java + JVM, XML and Zip32 Deflate (public domain e.g. used in PNG and OOXML)

- Additional system configuration, documentation regarding technical approval processes, underlying user/role/rights menegement is not part of archival package

Evaluation Results

11. Risk Behavior and Dependencies

SIARD:

- Explicit data export format/specification, based on **SQL99 representation**

Please note:

(*) While SIARD provides an explicit format specification, Chronos requires a mapping between internal form of data representation and the corresponding database configuration and mapping. Chronos explicitly documents the supported datatypes for every vendor and database version BUT treats the cross-vendor and inter-version representation as industrial secret.

Evaluation Results

SIARD CHRONOS

12. Referential Dependencies

In many cases the database does not contain full referential integrity as this is often depicted by external documentation or reflected within a different software layer. In some use cases it many be required to export a given dataset including all referential dependencies?

CHRONOS:

- Allows to automatically deted referential dependencies for master tables
- Has tools to decide how to deal with cyclic references and to which depth
- External dependencies can be remodeled manually within the tool.

SIARD:

- Ability to archive an entire database
  - Refers to a collection of objects that a db user has read access to
  - -> Creation of a dedicated export user with specific access rights
- All foreign keys are resolved

SIARD CHRONOS

Evaluation Results

13. Standard and Compliance

Currently there is no standard in the field of long-term archiving for databases. The SIARD format has become a widely accepted format for the exchange of relational database content within GLAMs.

- Both products are no fully compliant OAIS repositories, but rather create Archival Information Packages (AIPs)
- SQL subset for standard Archiving of ISO-9075-SQL, similar to PDF/A?
  - Entry, intermediate and full level conformance
  - SQL standard more than 2000 pages, far of being fully self-contained

For discussion and ideas on this, please refer to the paper.

CSP CHRONOS – Compliance statement for ISO 14721:2003 – by iKeep
http://ikeep.com/images/CHRONOS-OAIS-Compliance.pdf

Evaluation Results

SIARD CHRONOS

14. Structure, Setup and Size of the physical Archive

The Transaction Processing Performance council database (TPC-C) dump was used to get measures and comparison on the physical size of an exported database archive. Not taken into account in this comparison are parameters which are built up within a database environment that are not easily uniquely assignable. The size of the original source of a database is not a defined value i.e. there is no measurement on the size of an Oracle schema or database index in bytes?

- ***While a SIARD archive required +338% on disc space compared to the database dump a CHRONOS archive is able to decrease the required space by -41%.***

CHRONOS:

- In average we measured a 40-60% reduction of file size compared to database dump
- MD5 checksums applied out of the box, tends to blow up small records
- Archival Split size of 20MB showed the best performance regarding searchability

SIARD:

- Uses a zip container but does not apply any compression algorithm. By applying a post-compression (deflate 32K word size, standard compression) size of an archive can be brought down to +30%

Evaluation Results

SIARD CHRONOS

15. Specification of Information Lost

Which **audit trail capabilities** does the system offer for logging and tracking modifications over time. Is there a way of specifying the **amount of information lost** when exporting data into a long-term archive? One example on a measure which could be applied is the Oracle SQL Minus operation after re-importing a database archive to determine the correct structure and item count against the original data.

- Both SIARD and CHRONOS are idempotent in terms of upload-download-upload delivers the same data types and values.
- **No statement to declare what information is actually lost**
  - during export (UDTs, disabled foreign key constraints, etc.)
  - During cross db version re-import or mapping from native type to SQL99
- Provide logging to track down system behavior
- **No audit trails** or similar tracking features at this level.
- → More in the realm of the enclosing archival system, but could expect from Chronos.

The blurring of the boundary

DB Preservation with Record Centric Approach?

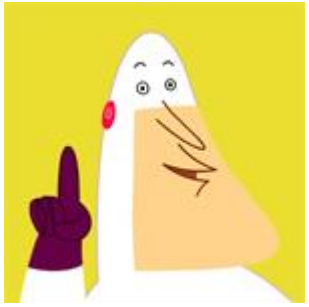
- Brown, A., Lappin, J.: Ecm talk 17: Practical digital preservation (2013) http://traffic.libsyn.com/ecmtalk/ECM_Talk_017.mp3

*ECM podcast on practical digital preservation Adrian Brown, director of the Parliamentary Archives in London mentions the **'blurring of the boundary**' between **digital objects** and the **applications that they are held in** as key challenge the institutions are confronted with. Digital preservation initiatives and projects made great progress in tackling the problem of how to preserve the file formats and the objects themselves but now faces the more complex problem of **how to preserve the information that an application has about the objects it holds**? How to enable digital objects to move from one application to another without losing that information?*

- Preservation is a selection process. Taking (well documented) decisions.
- You need to take the decision: what are the significant apsects your object must keep, independent if it's reflected within the format or the tool itself.
- -> Chronos is a perfect example, exporting ‚core data' in an open archivable way, remodeles data representation; offers clever database like performance on top, providing interfaces to access, query data, to integrate in a fully automated workflow and rebuilds essential infrastructure (as permission management) on top.

Summary & Conclusion 1/3

- Archiving databases either means preserving information or preserving functionality or both
- **Both tested tools proved stable, technically mature in creating a db-archive in a vendor independent long-term preservation format for a rich number of database systems**
- Both tools showed solid performance
- Differences in the number of supported db vendors, SQL elements and internal data representation.



➔No clear recommendation which product to adopt, as supported scope and use cases both tools are able to deliver are highly diverse!

Summary & Conclusion 2/3

- SIARD: designed as reference implementation for SIARD format, **exclusively offers support for database retirement.**

- CHRONOS: commercial product, designed for scalability and industrial needs, provides rich set of tools and end-user applications that allow **both to export a physical archive and to operate on top**. It provides all required bits to handle the required complexity for ongoing/continious and partial archiving
  + SQL92 queries on archived data
  + Database like performance (through indices and in memory dbs)
  + Revisions; syntactic and semantic schema modifications
  + Resolving Cyclic Dependencies
  + Full blown access control and data retention layer
  + Snapshot data export and compliance

Summary & Conclusion 3/3

CHRONOS

- Limited Support of complex objects (as Oracle UDTs)
- Lacking support of audit trails for classification and documentation of information lost.
- No full archival system in terms of OAIS
- Cross vendor-mapping ‚secret‘

~Application Retirement via Archive Explorer: remodelling of business objects, application logic and report functionality

+ Middleware layer to retarget legacy applications
+ Rich set of programmatic interfaces allow to integrate with most of the system‘s functionality as well as to grant access to data via standard mechanisms as JDBC

AUSTRIAN INSTITUTE OF TECHNOLOGY

Questions?

Preserving Transactional Data Briefing Day
Joint Workshop of the
Digital Preservation Coalition and
UK Data Service
17th of March 2016
WCIT, London

Andrew Lindley
Research Engineer
Digital Safety & Security Department
Digital Insight Lab (DIL)
andrew.lindley@ait.ac.at
@lchwardort