Introduction into Database Forensics – Issues to be solved

Antje Raab-Düsterhöft
University of Applied Science Wismar, Technology, Business and Design
Dep. of Computer Science

antje.duesterhoeft@hs-wismar.de
https://it-forensik.fiw.hs-wismar.de/

Database Forensics...

- subfield of digital forensics
- deals with forensic examination of databases

Databases....

- structured collections of data
- allows efficient storage, retrieval and manipulation of information
- controlled by database management systems (DBMS)

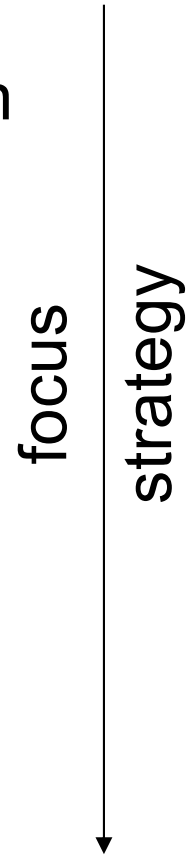
Why Databases are critical assets?

- Databases contains important information
- Database servers now store a greater volume of sensitive data than ever before
- Database security breaches are everywhere and also used in the context of various attacks
  - Use of credit cards
  - Use of medical personal data
  - Use of industrial data
  - …
- Insider attacks: persons can manipulate data for their own purposes

→ Databases are “interesting sources”

The Road to Database Forensics...

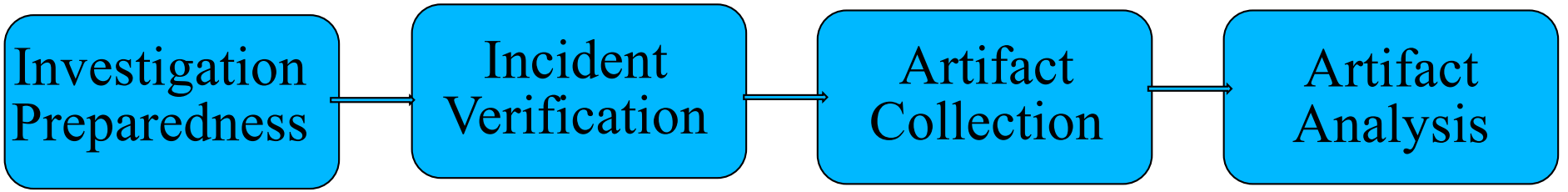
1. Goals
2. Preparation of database forensic investigation
   - Scenarios
   - Kinds of DBMS & Database models
3. Incident verification & Artefact collection
4. Artifact analysis
   - Overview of database artifacts
   - Physical and logical database artifacts
   - Tools
5. Conclusion – Issues to be solved



1. Goals

- Prove or disprove an occurance of a data security breach
- Determine the scope of a database intrution
- Retrace user operations (SQL operations)
- Identifiy data pre- and post-transactions
- Recover previously manipulated or deleted data
- ...

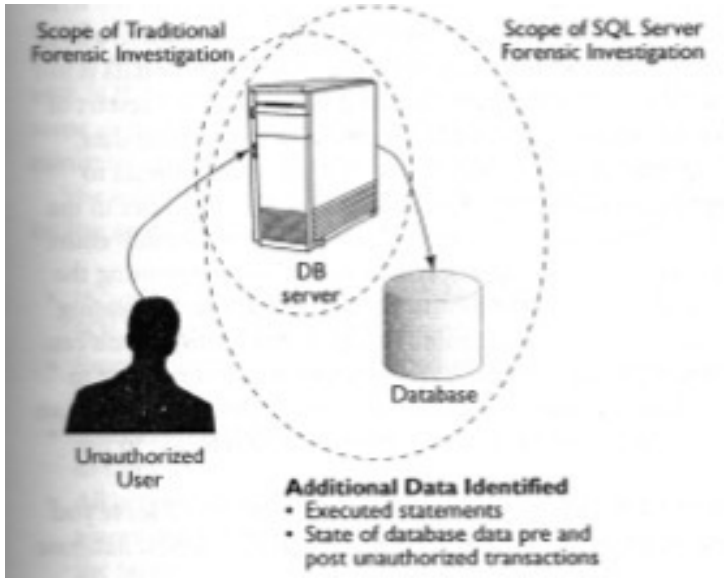
Kevvie Fowler: SQLServer Forensic Analysis, p.49



Kevvie Fowler: SQLServer Forensic Analysis, p.60

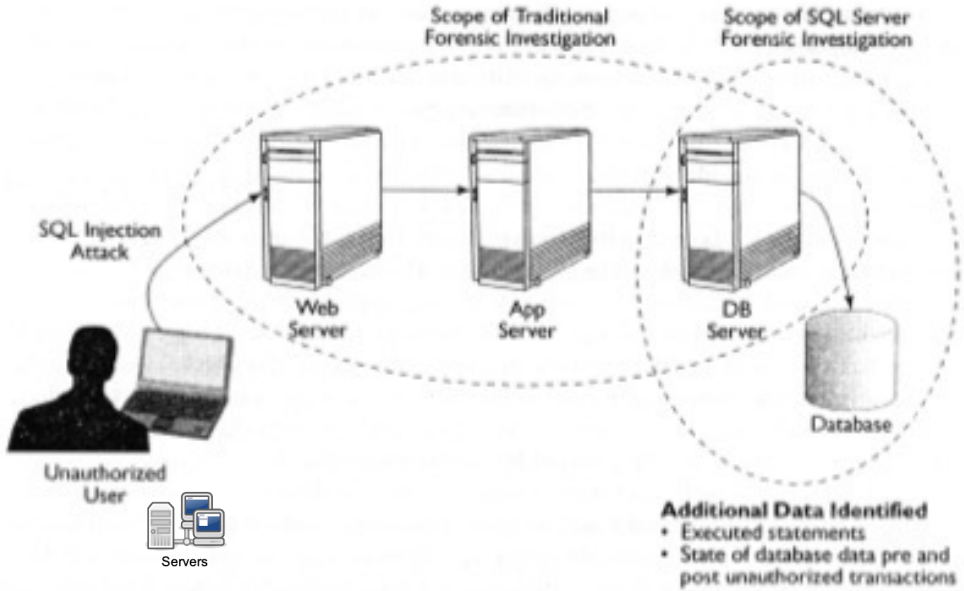
2. Preparation:

Scenario 1

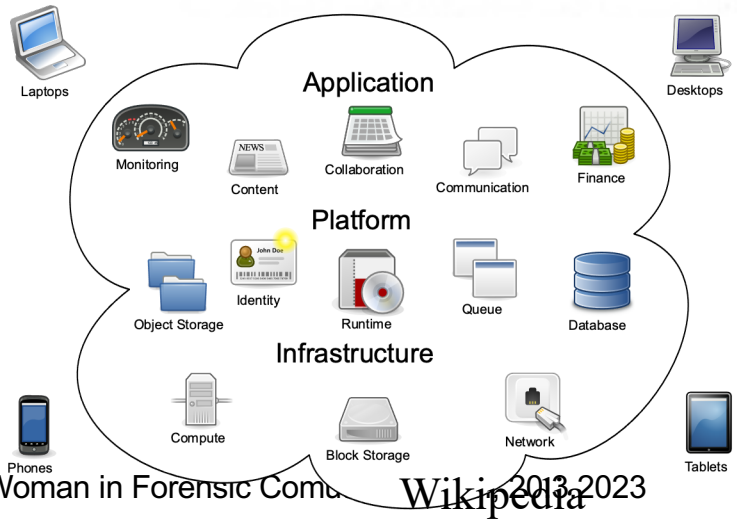


Kevvie Fowler: SQLServer Forensic Analysis, p.53

Scenario 2

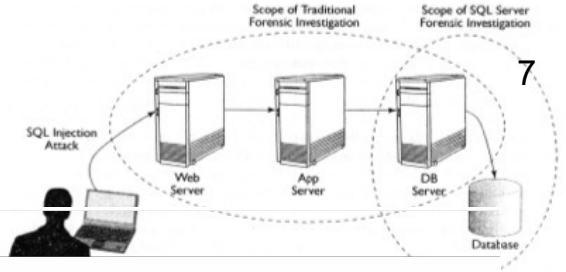


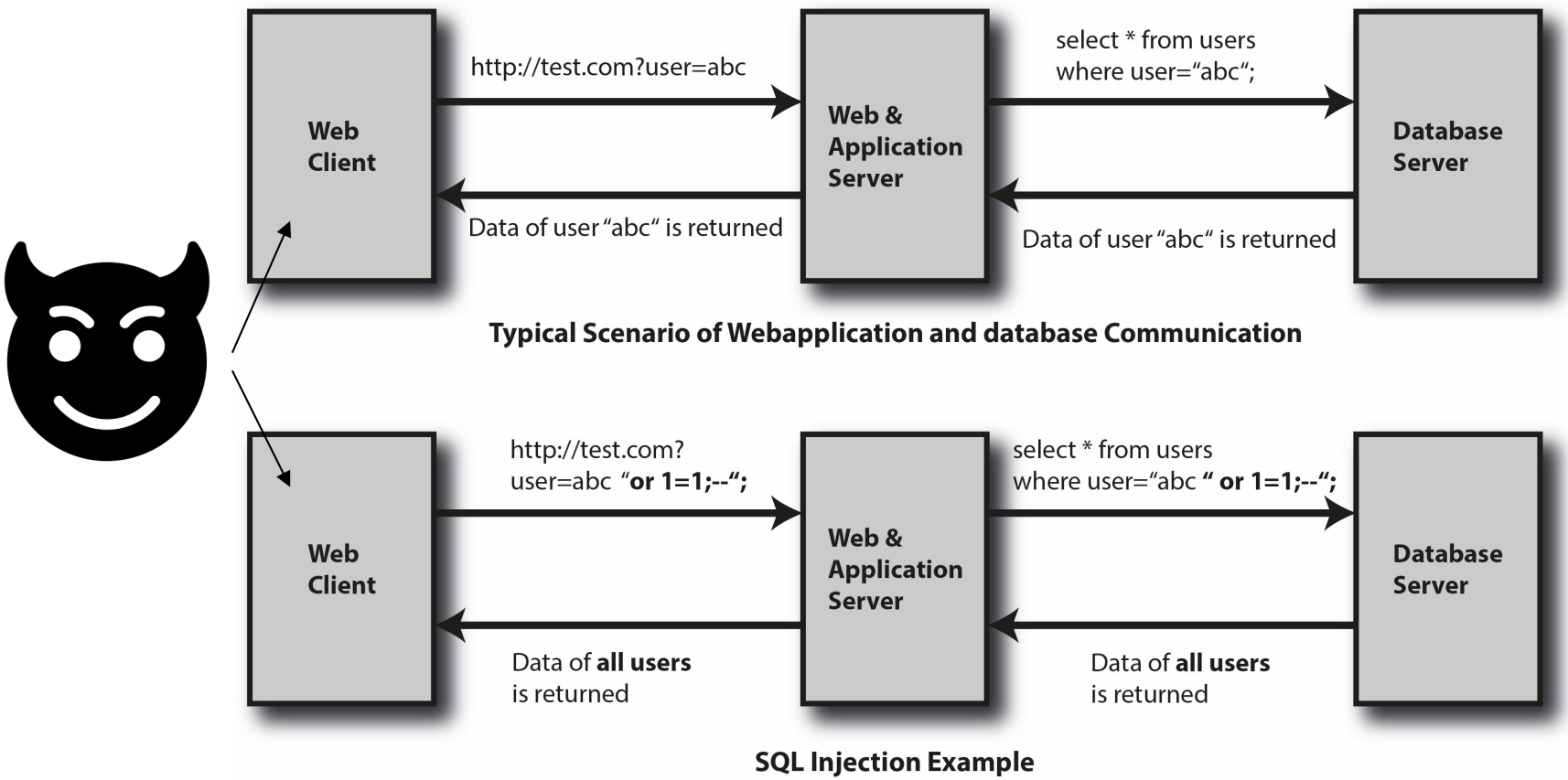
Scenario 3 Cloud system (new)



Antje Raab-Düsterhöft, DFRWS-EU, "Woman in Forensic Comu... Wikipedia 2013 2023

Scenario 2: Example SQL Injection

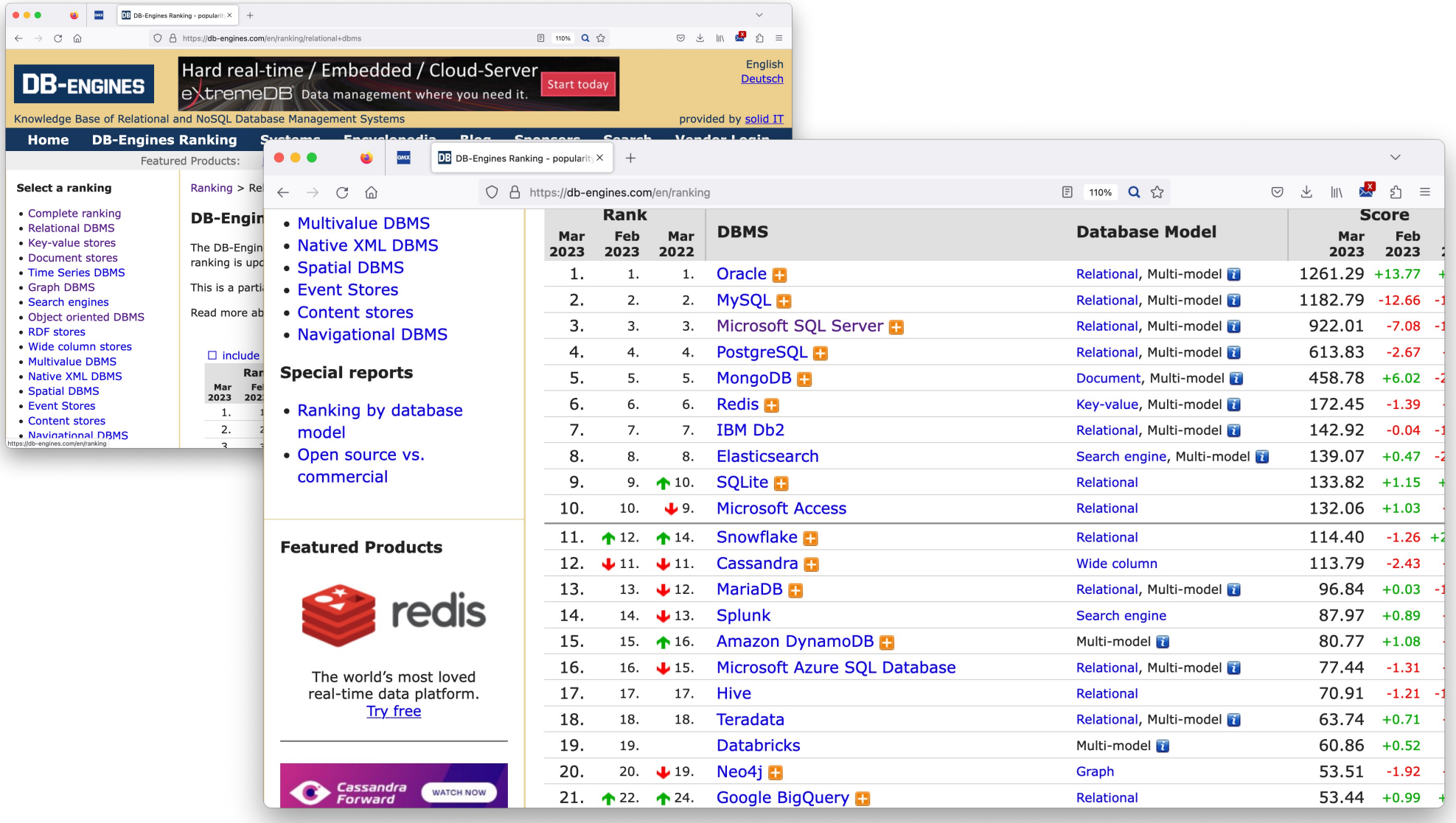




- Application Logs

- Webserver-Log
- Error Log
- Page-File
- Memory
- ...

2. Preparation: Database Management Systems (RDMS)



https://www.db-engines.com (March 2023)

2. Preparation: Types of database models

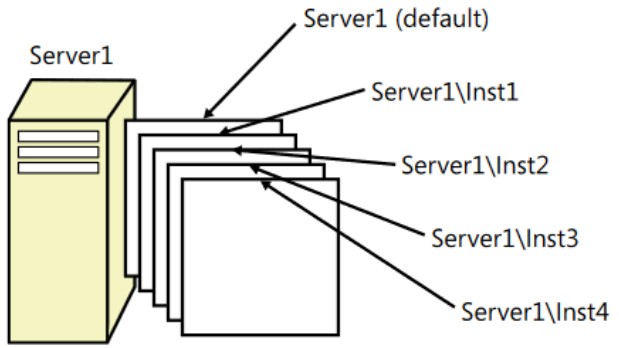
	Relational	NoSQL	Others (deductive, object oriented, ...)
Data model	Structured (table)	Free/ differnet	Free/ different
Audit/ Log Mechanismen	many (some standardized, some differnent)	some/ few	?
Storage	storage management, Some differences	file-based	?
Standard	SQL	no/ own	?

focus

2. Preparation: Strategy

1. Defining or identifiying the scope/ infrastructure
2. Identifiying DBMS, version, configuration, instances, ...
3. Classification of the database model

?



Kevvie Fowler: SQLServer Forensic Analysis, Appendix B, p. 439ff

2. Preparation: Strategy

4. Configure a forensic workstation for the database system

- Use forensic software
- Use/ create/ integrate forensic scripts
  - for live aquisistion/ automated live forensic
  - for memory aquisistion (volatile data)
  - (first) artifact collection

2. Preparation: Database Artifacts

Database artifacts refer to the objects that are created in a database system as part of structure and functionality.

Physical artifacts:

- the actual files that are created on the storage medium when a database is created
- data files, log files, backup files, system information, …

Logical artifacts:

- the database objects that are created and managed within the database
- Scheme, tables, views, indexes, stored procedures, triggers, constraints, …

2. Preparation: Forensic database scripts

- for live aquisistion/ automated live forensics
- for memory aquisistion (volatile data)
  - Caches (data, plan, ...)
  - Databases
  - SQL statements
  - Connections
  - Transaction log
  - Users
  - ...
  - DBObjects
  - Configurations
  - Jobs
  - TimeConfigs
  - Triggers
  - Sessions
  - ...

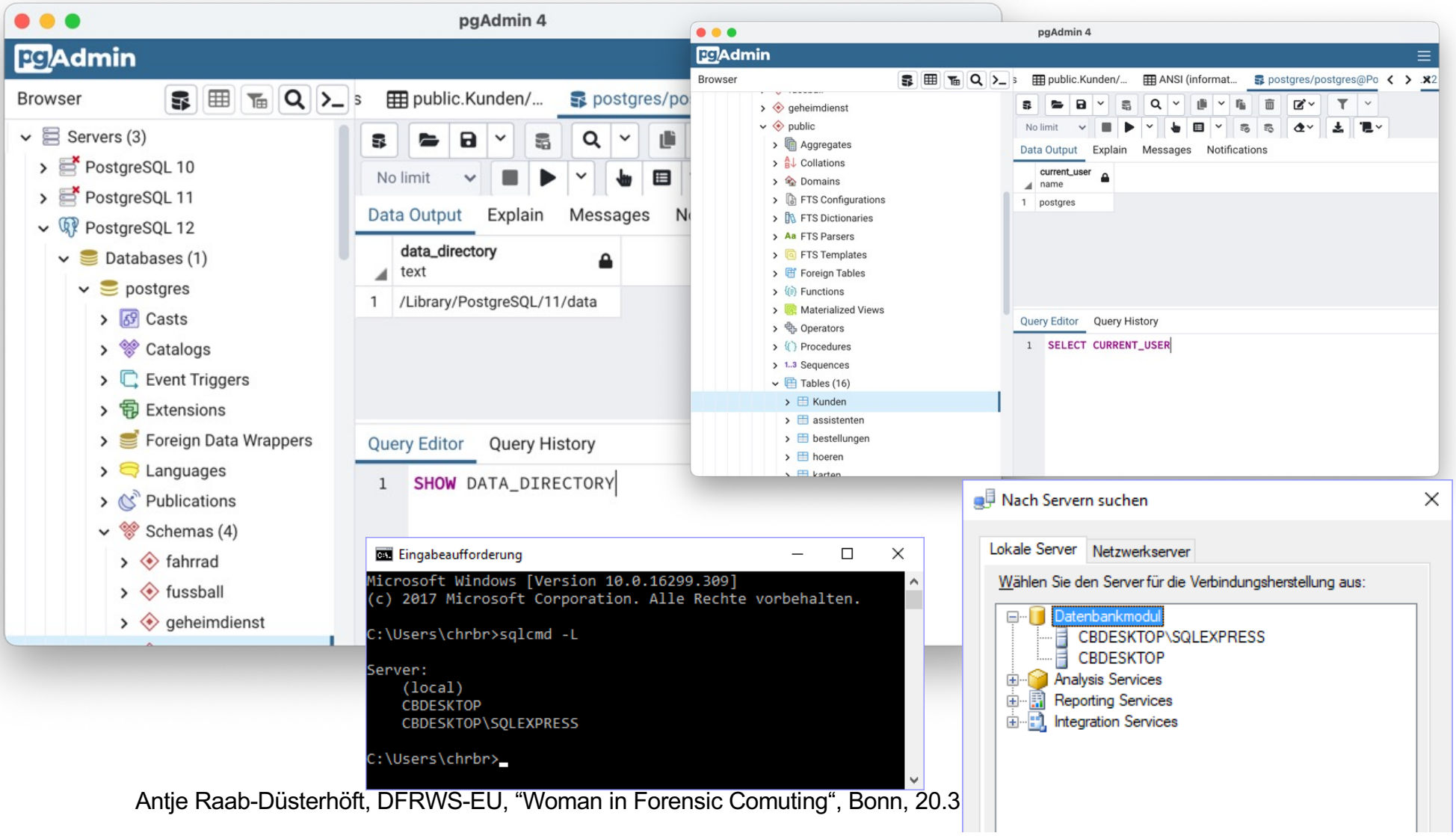
e.g. SQL Server: SELECT * FROM sys.configurations

SQL Server: SELECT * FROM **sys.dm_exec_cached_plans cross apply sys.dm_exec_sql_text(plan_handle)**

Postgres: SELECT current_user

Oracle: SELECT owner,table_name,column_name,data_type
FROM dba_tab_columns
WHERE data_type='DATE' or data_type like 'TIMESTAMP%';

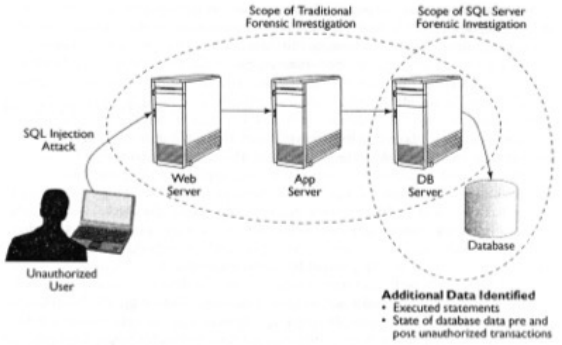
2. Preparation: Forensic database scripts



3. Incident Verification + Artifact Collection

Identifiying first signs of breaches

- Server problems, server traffic
- Network traffic
- An (unauthorized) modification of data
- Compare checksums of database source code
- Confirm the breach/ incident
- Made a decision about shut-down/ seperate the database server



First collection of forensic data (forensically usable)

- Collecting volatile and non-volatile artitacts
- Bit-to-bit image of the database data, query files, query output,..
- Collecting data from the infrastructure

3. Incident Verification + Artifact Collection

General rules for collecting database artifacts:

1. Do not use system libraries, they might be compromised
2. Do not execute DDL/DML statements to avoid changes to data
3. Do not create SQL Server logins, database users or objects
4. Restrict the use of temporary tables and variables in order not to overwrite existing information in the SQL Server buffer
5. Do not modify existing corrections or database objects
6. Do not execute untested statements, errors will be logged and may change the event log

3. Artifact Collection: Starting point

Results of collecting database artifacts

Decision of analysing methods:

1. Actual database
2. Images, that can be used for starting the DBMS with databases
3. Parts of the databases (data), that can be integrated into a virtual environment with the according DBMS

Further data from the short artifact analysis

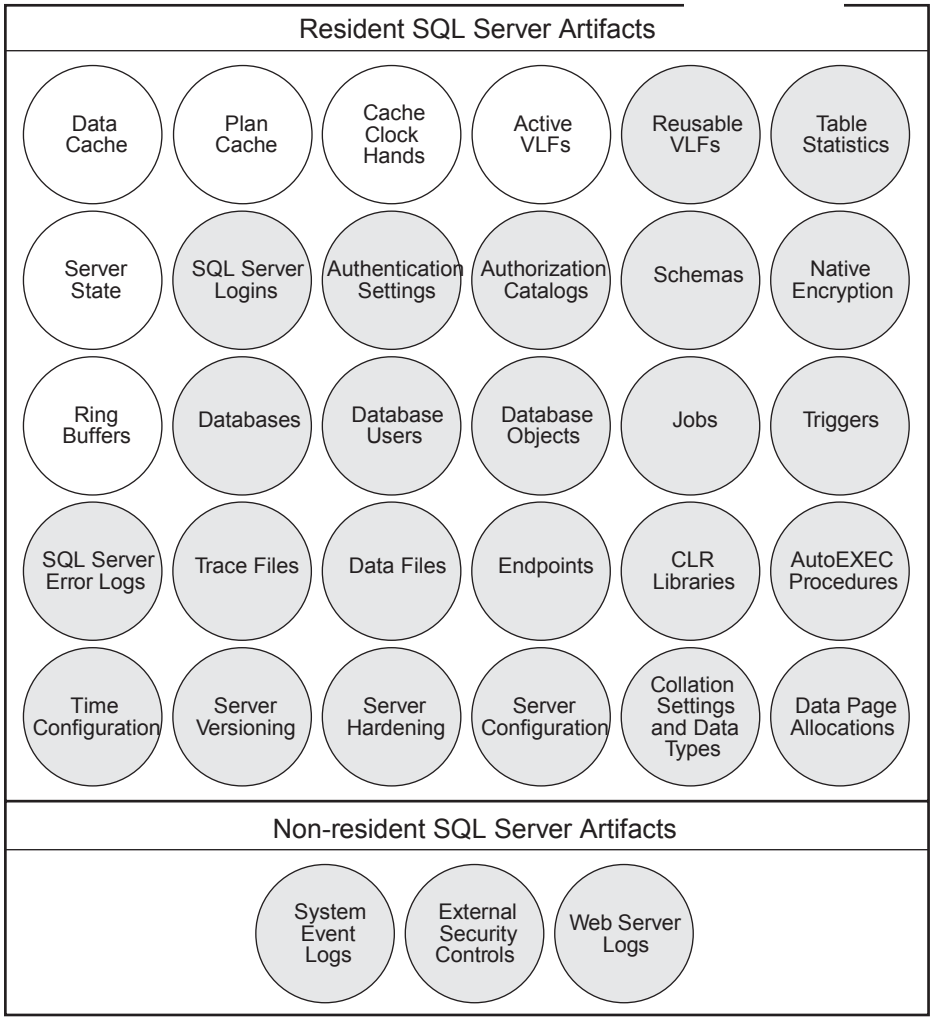
- Incident SQL queries, that were found
- Data, that was queried or modified
- Application Logs/ Web Server Log with IP Adresses (connections to the servers)

3. Artifact Collection: Overview

lost upon MSSQLServer service shutdown — flüchtig

persistent artifacts-those that remain after an MSSQLServer service restart — nicht flüchtig

- DBMS specific
- Artifacts are
  - Logical
  - Physical
  - Volatile
  - Non-volatil
  - Resident
  - Non-resident



Quelle: Fowler, Kevvie, SQL Server Forensic Analysis, 2009, S.65.

4. Artifact Analysis: Logical Artifacts

Starting point: e.g. Incident SQL queries in Web Server Log

First: Get an overview of the logical artifacts

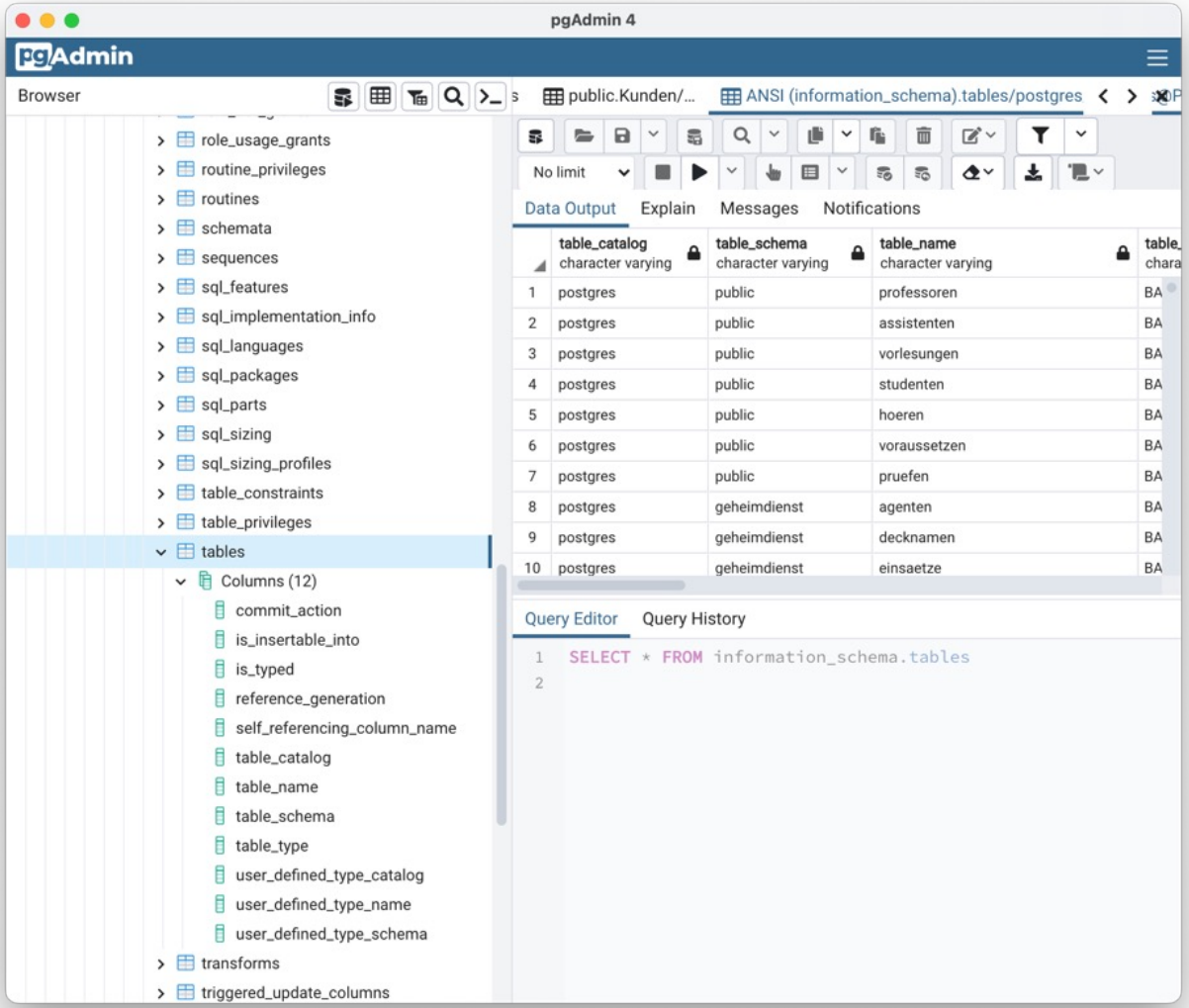
- Which database ist involved?
- Which data are stored in this database?
- Which users were active?
- Which rights does a special user have?
- Which queries were executed?
- ...

→ SQL-Queries for analysis

4. Artifact Analysis: Database structure

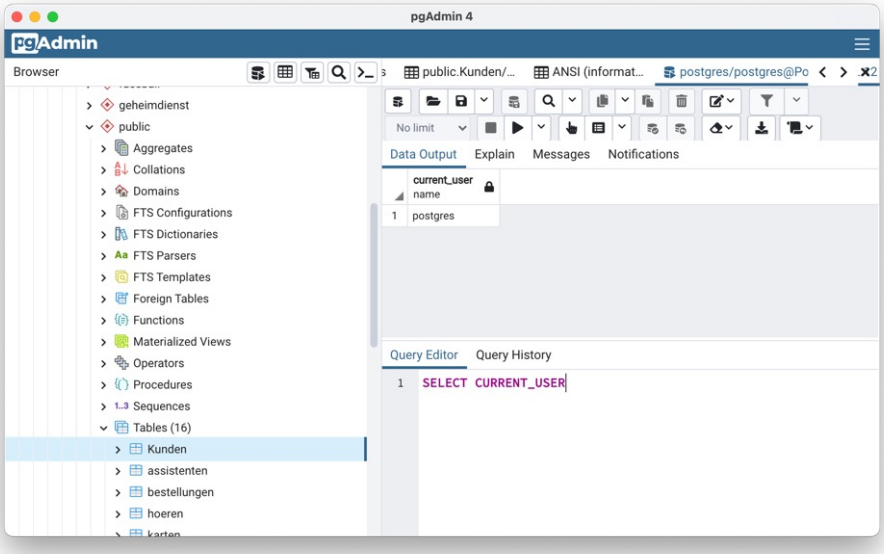
Information Schema:
All information about the database

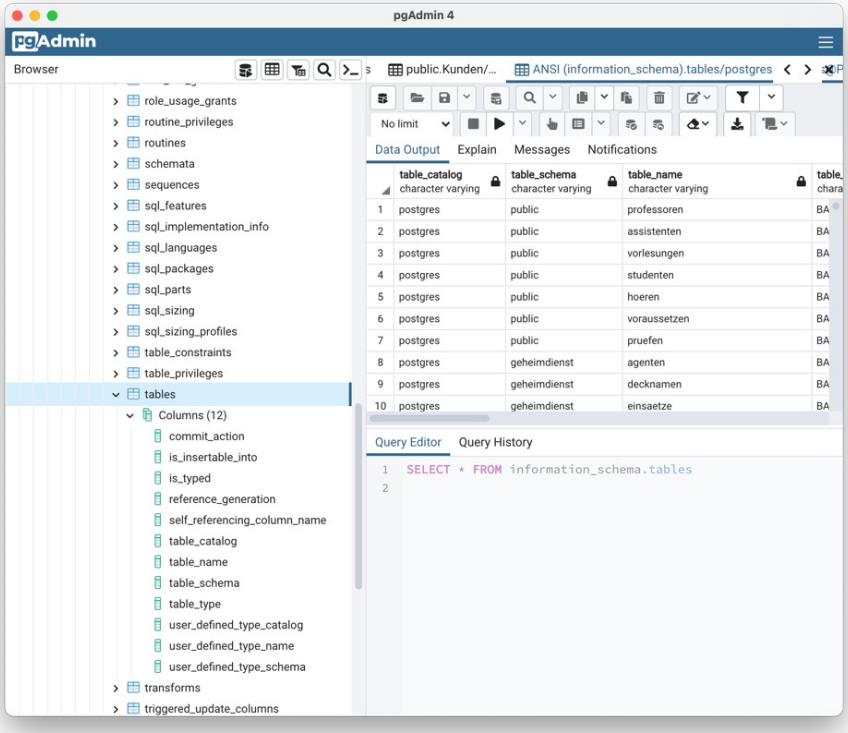
General Catalogs:
Information about the dbms

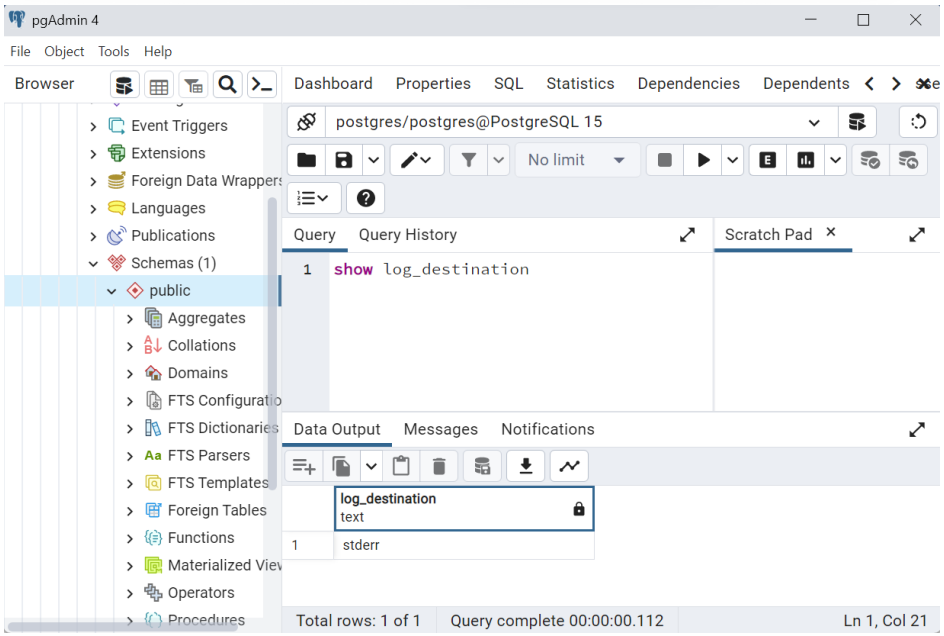


→ Scope: what the attacker might have done

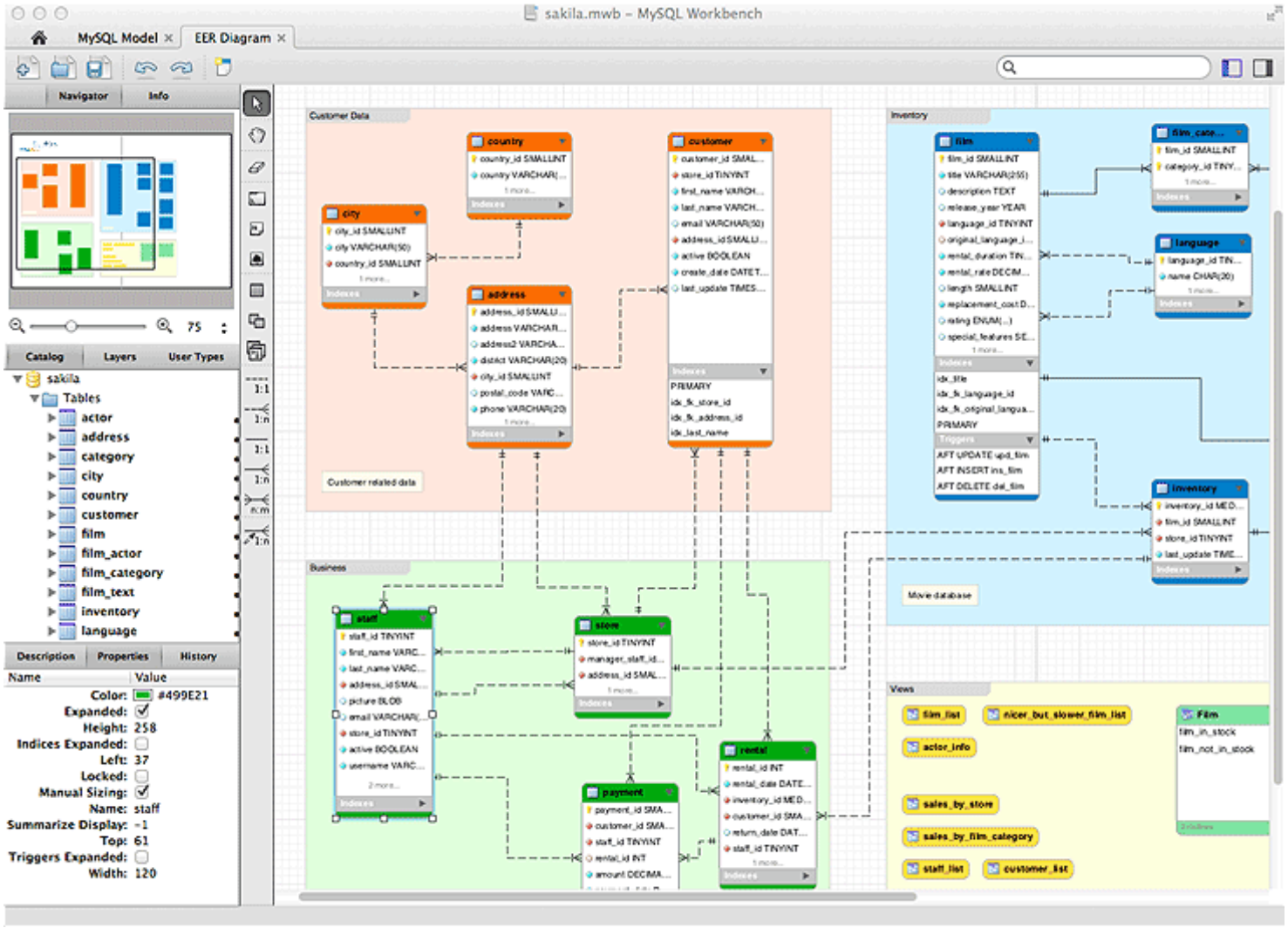
4. Artifact Analysis:
Database functions, user, logs, trigger, ...



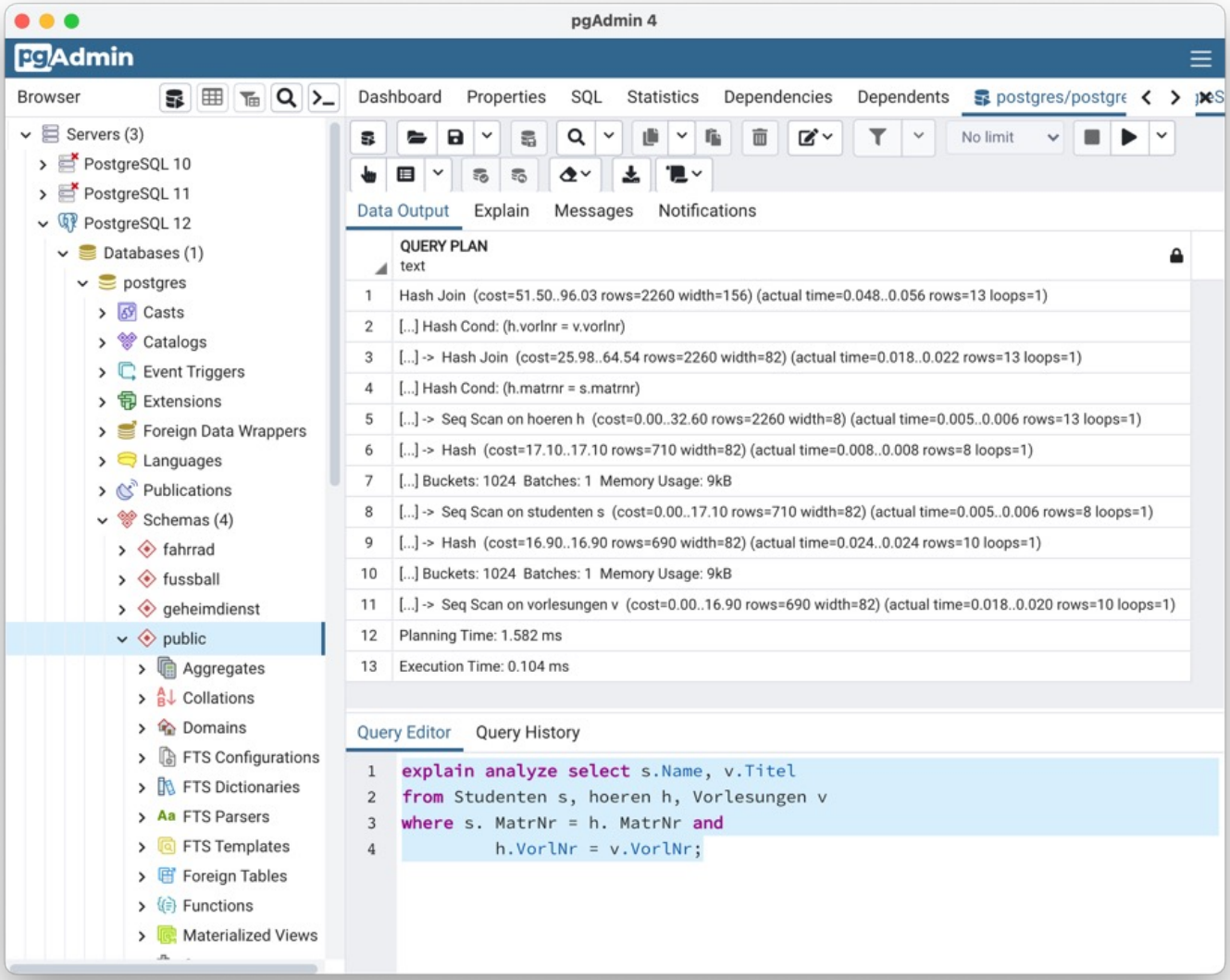




4. Artifact Analysis: Visualization of the db structure



4. Artifact Analysis: Query plan (Postgres)



4. Artifact Analysis: Physical Artifacts

Logs: must be enabled before an incident (!)

- SQL Logs
- Error Logs
- Transaction logs/ Write Ahead Logs (WAL)
- Special Query logs

Transaction logs (used for roll-back and recover databases)...

- Specific for DBMS
- Detailed information about SQL operation and data
  - Insert/ update/ delete data, create structure/ functions, user, ...
- Interpretation tools only partially available

What people are saying about the transaction log: **"Although you might assume that reading the transaction log directly would be interesting or even useful, it's just too much information."**

Inside SQL Server 2005: The Storage Engine, Microsoft Press, 2006

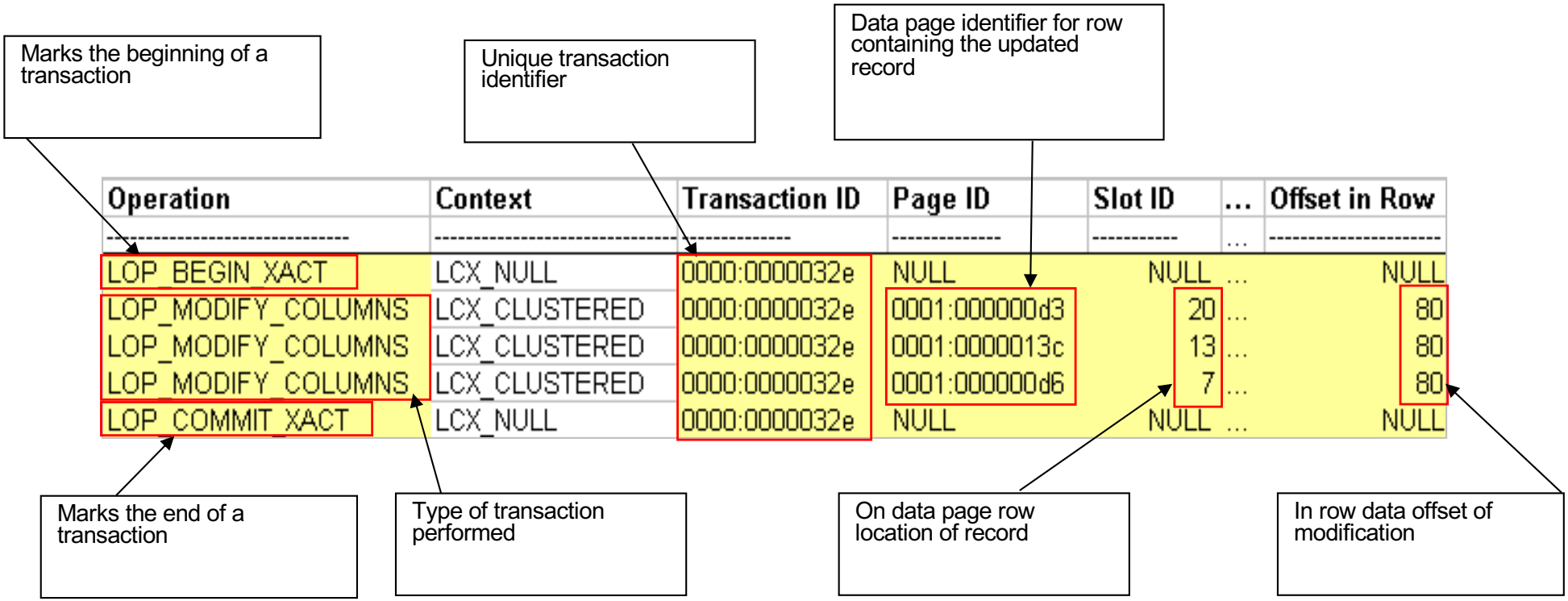
MS SQL Server: Transaction Log

Inside the transaction log:

Kevvie Fowler: SQLServer Forensic Analysis

1. CurrentLSN
2. Operation
3. Context
4. Transaction ID
5. Tag Bits
6. Log Record Fixed Length
7. Log Record Length
8. PreviousLSN
9. Flag Bits
10. Log Reserve
11. AllocUnitID
12. AllocUnitName
13. Page ID
14. Slot ID
15. Previous Page LSN
16. PartionID
17. RowFlags
18. Num Elements
19. Offset in Row
20. Checkpoint Begin
21. CHKPT Begin DB Version
22. MaxXDESID
23. Num Transactions
24. Checkpoint End
25. CHKPT End DB Version
26. Minimum LSN
27. Dirty Pages
28. Oldest Replicated Begin LSN
29. Next Replicated End LSN
30. Last Distributed End LSN
31. Server UID
32. UID
33. **SPID**
34. BeginLog Status
35. Xact Type
36. Begin Time
37. Transaction Name
38. Transaction SID
39. End Time
40. Transaction Begin
41. Replicated Records
42. Oldest Active LSN
43. Server Name
44. Database Name
45. Mark Name
46. Master XDESID
47. Master DBID
48. PrepLogBegin LSN
49. PrepareTime
50. Virtual Clock
51. PreviousSavePoint
50. Savepoint Name
51. Rowbits First Bit
52. Rowbits Bit Count
53. Rowbits Bit Value
54. Number of Locks
55. Lock Information
56. LSN Before Writes
57. Pages Written
58. Data Pages Delta
59. Reserved Pages Delta
60. Used Pages Delta
61. Data Rows Delta
62. Command Type
63. Publication ID
64. Article ID
65. Partial Status
66. Command
67. Byte Offset
68. New Value
69. Old Value
70. New Split Page
71. Rows Deleted
72. Bytes Freed
73. CI Table ID
74. CI Index ID
75. FIlegroup ID
76. Meta Status
77. File Status
78. File ID
79. Physical Name
80. Logical Name
81. Format LSN
82. RowsetID
83. TextPtr
84. Column Offset
85. Flags
86. Text Size
87. Offset
88. Old Size
89. New Size
90. Description
91. Bulk allocated extent count
92. Bulk rowinsertID
93. Bulk allocationunitID
94. Bulk allocation first IAM Page ID
95. Bulk allocated extent ids
96. **RowLog Contents 0**
97. **RowLog Contents 1**
98. RowLog Contents 2
99. RowLog Contents 3
100. Rowlog Contents 4
101. Log Record

MS SQL Server: Transaction Log



Operation	Context	Transaction ID	Page ID	Slot ID	...	Offset in Row
LOP_BEGIN_XACT	LCX_NULL	0000:0000032e	NULL	NULL	...	NULL
LOP_MODIFY_COLUMNS	LCX_CLUSTERED	0000:0000032e	0001:000000d3	20	...	80
LOP_MODIFY_COLUMNS	LCX_CLUSTERED	0000:0000032e	0001:0000013c	13	...	80
LOP_MODIFY_COLUMNS	LCX_CLUSTERED	0000:0000032e	0001:000000d6	7	...	80
LOP_COMMIT_XACT	LCX_NULL	0000:0000032e	NULL	NULL	...	NULL

Identifier	Hex	Decimal
Transaction ID	0000:0000032e	0:814
Data Page	0001:000000d3	1:211

Kevvie Fowler: SQLServer Forensic Analysis

MS SQL Server: Transaction Log –Recovering SQL is possible

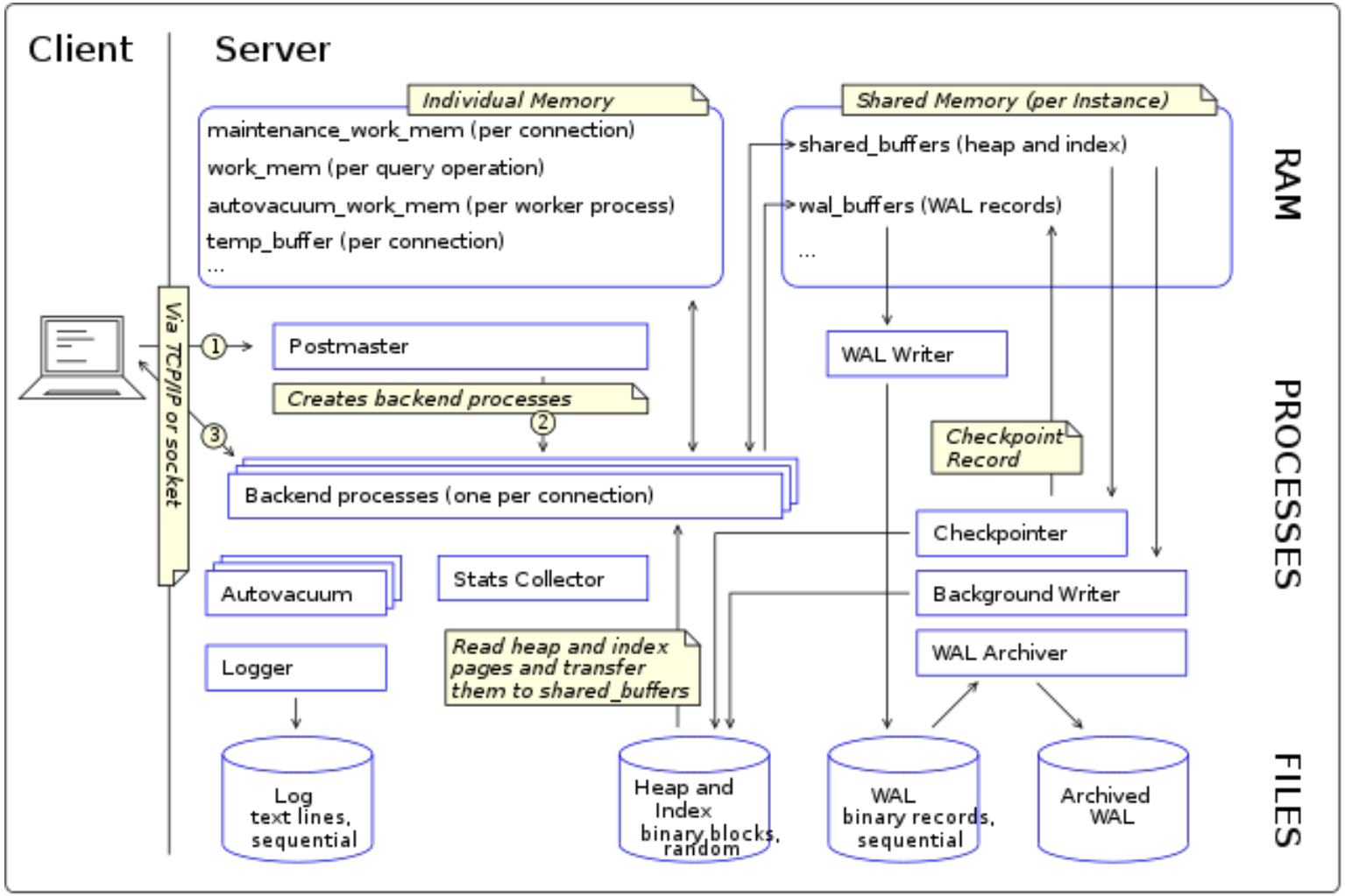
```
Slot 20 Offset 0x147f Length 237

Record Type = PRIMARY_RECORD          Record Attributes =  NULL_BITMAP VARIABLE_COLUMNS

Memory Dump @0x2F3AD47F

00000000:   30006c00 6f000000 53007000 72006900 †0.l.o...S.p.r.i.
00000010:   6e006700 4c006100 6b006500 20002000 †n.g.L.a.k.e. . .
00000020:   20002000 20002000 20002000 20002000 † . . . . . .
00000030:   41005a00 31003400 34003100 30000a00 †A.Z.1.4.4.1.0...
00000040:   00000100 00000000 0000e498 00003300 †..............3.
00000050:   2e003500 30002000 20002000 20002000 †..5.0. . . . . .
00000060:   20002000 20002000 20002000 0e0000c0 † . . . . . .....
00000070:   06008400 88009900 9d00ad00 ed00416e †..............An
00000080:   6f736f6e 456d696c 37322053 74617266 †osonEmil72 Starf
00000090:   656c6c20 44726976 65566973 61343931 †ell DriveVisa491
000000A0:   36383833 38343033 38323330 3056006f †6883840382300V.o
000000B0:   006c0063 0061006e 006f0020 00360032 †.l.c.a.n.o. .6.2
000000C0:   00200069 006e0063 00680020 0050006c †. .i.n.c.h. .P.l
000000D0:   00610073 006d0061 00200054 00560020 †.a.s.m.a. .T.V.
000000E0:   00560043 00320033 00330032 00†††††††.V.C.2.3.3.2.
```

Transaction

Start of column

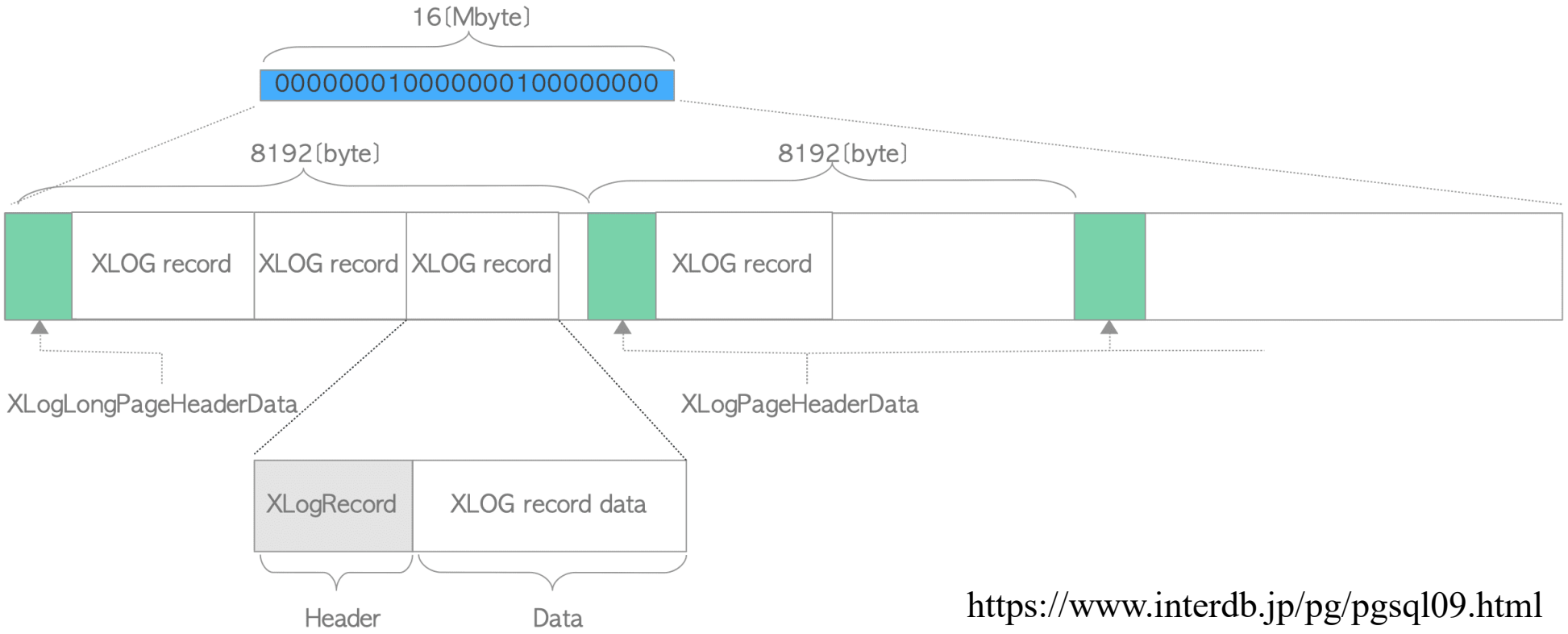
Kevvie Fowler: SQLServer Forensic Analysis

Postgres internal structure



https://commons.wikimedia.org/w/index.php?curid=112231040

Postgres - Internal layout of a WAL segment file

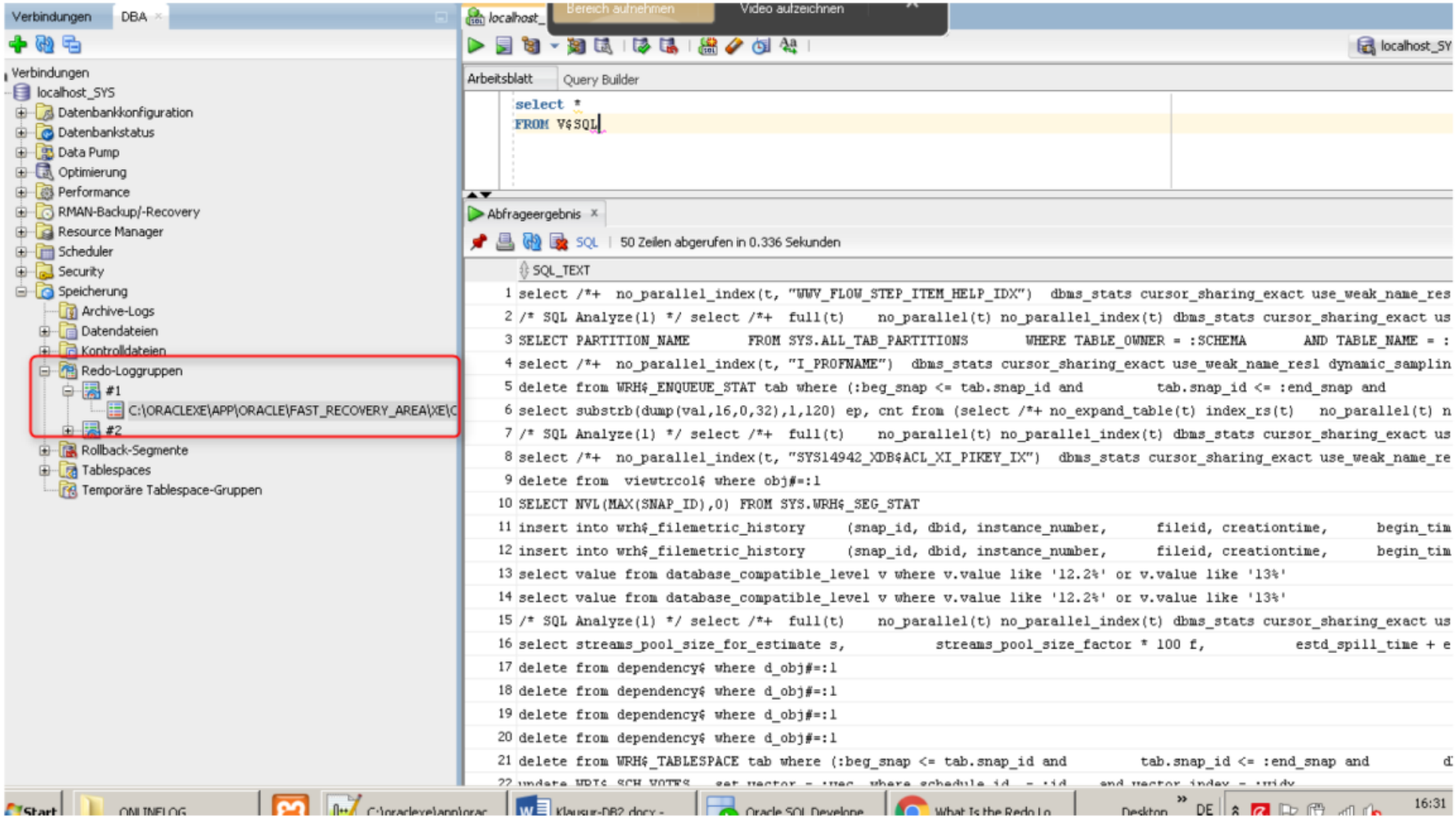


https://www.interdb.jp/pg/pgsql09.html

```
typedef struct XLogRecord
{
    uint32          xl_tot_len;    /* total len of entire record */
    TransactionId   xl_xid;        /* xact id */
    uint32          xl_len;        /* total len of rmgr data */
    uint8           xl_info;       /* flag bits, see below */
    RmgrId          xl_rmid;       /* resource manager for this record */
    /* 2 bytes of padding here, initialize to zero */
    XLogRecPtr      xl_prev;       /* ptr to previous record in log */
    pg_crc32        xl_crc;        /* CRC for this record */
} XLogRecord;
```

4. Artifact Analysis: Tools for the interpretation of logs

Oracle: LOG File from the LOG-MINER



4. Artefact Analysis: Physical Artifacts - Data

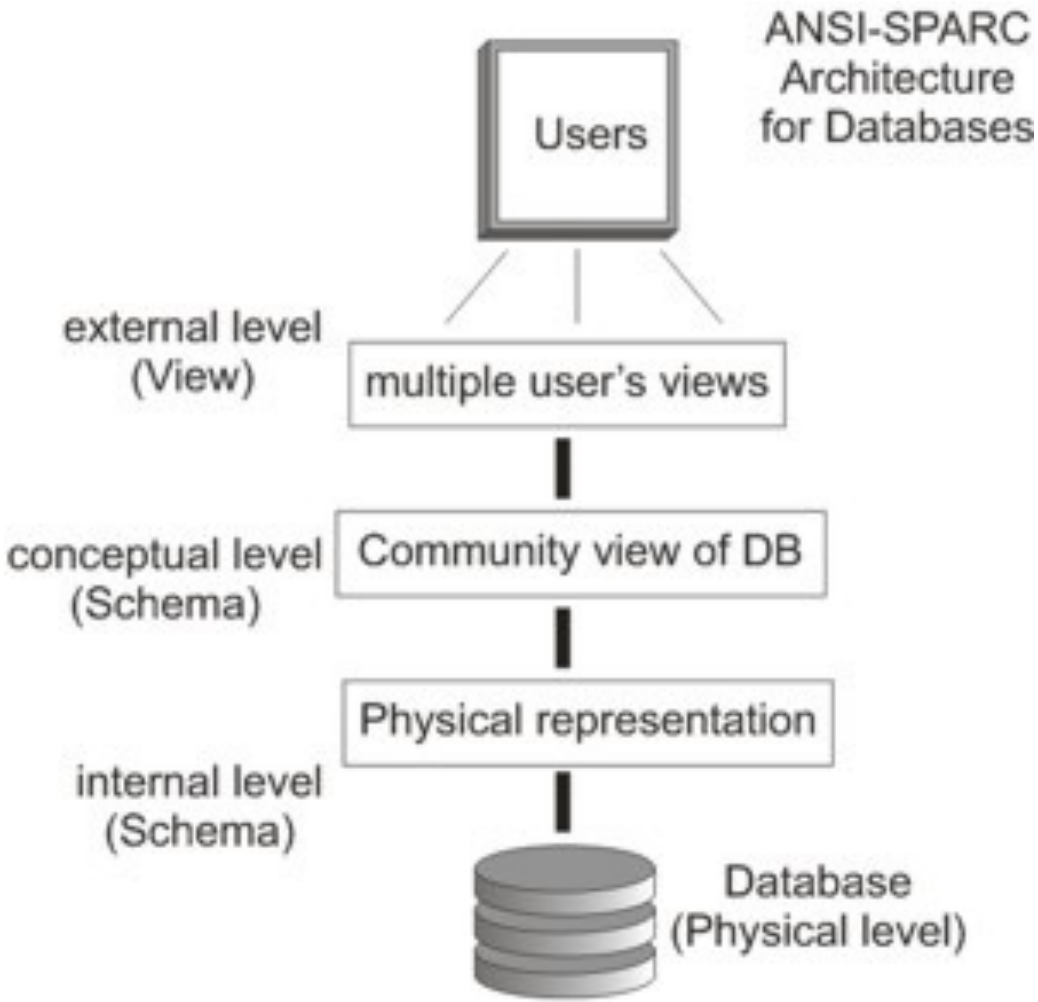


Table Structures

Index Structures

Data Files

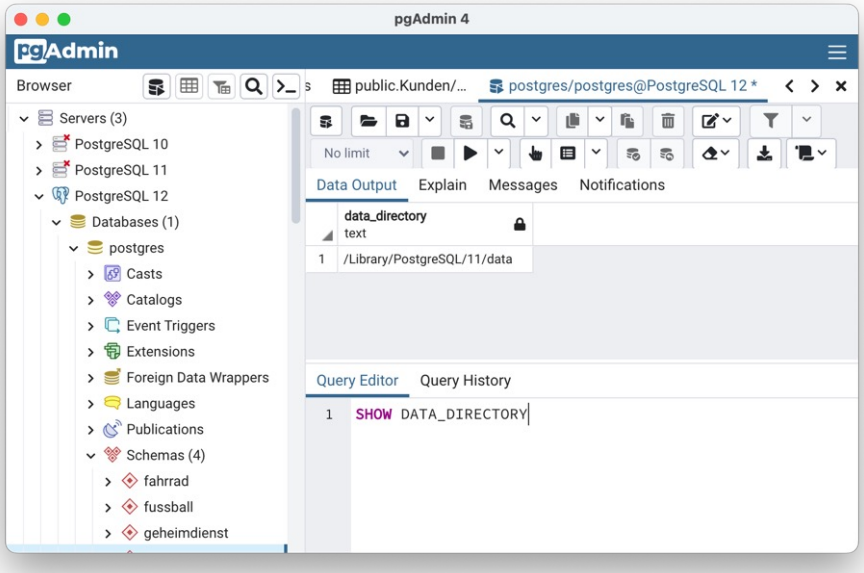
The ANSI-SPARC three-level architecture (Wikipedia)

4. Artefact Analysis: Data files (Postgres)

Table Structures

Index structures

Data Files



Postgres:

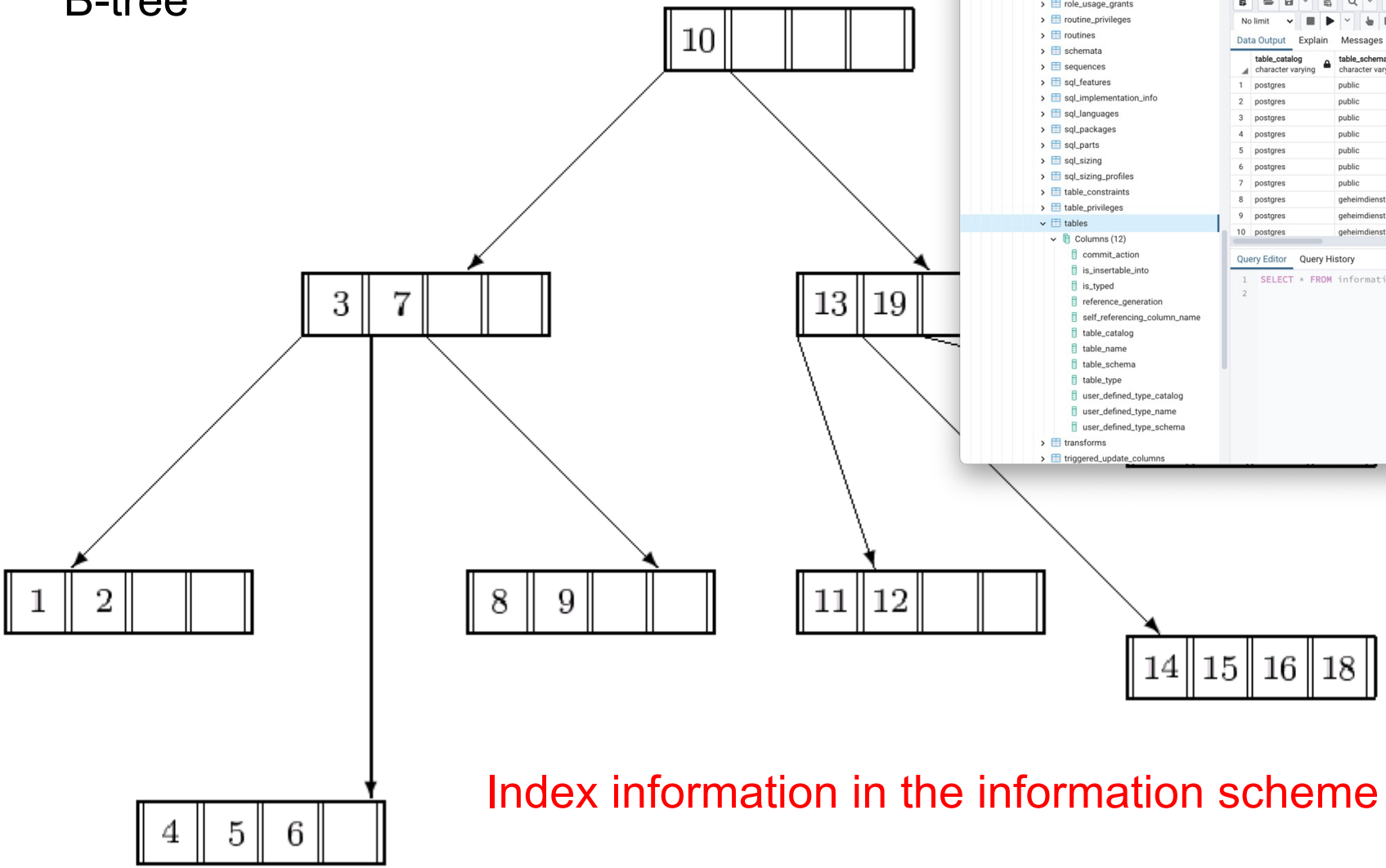
- Data directory: **/var/lib/pgsql/data**
- Each PSQL database table is a separate file with a default file extension of **. mkd**.

- Different index structures:

  B-tree, Hash, GiST, SP-GiST, GIN, BRIN, the extension bloom

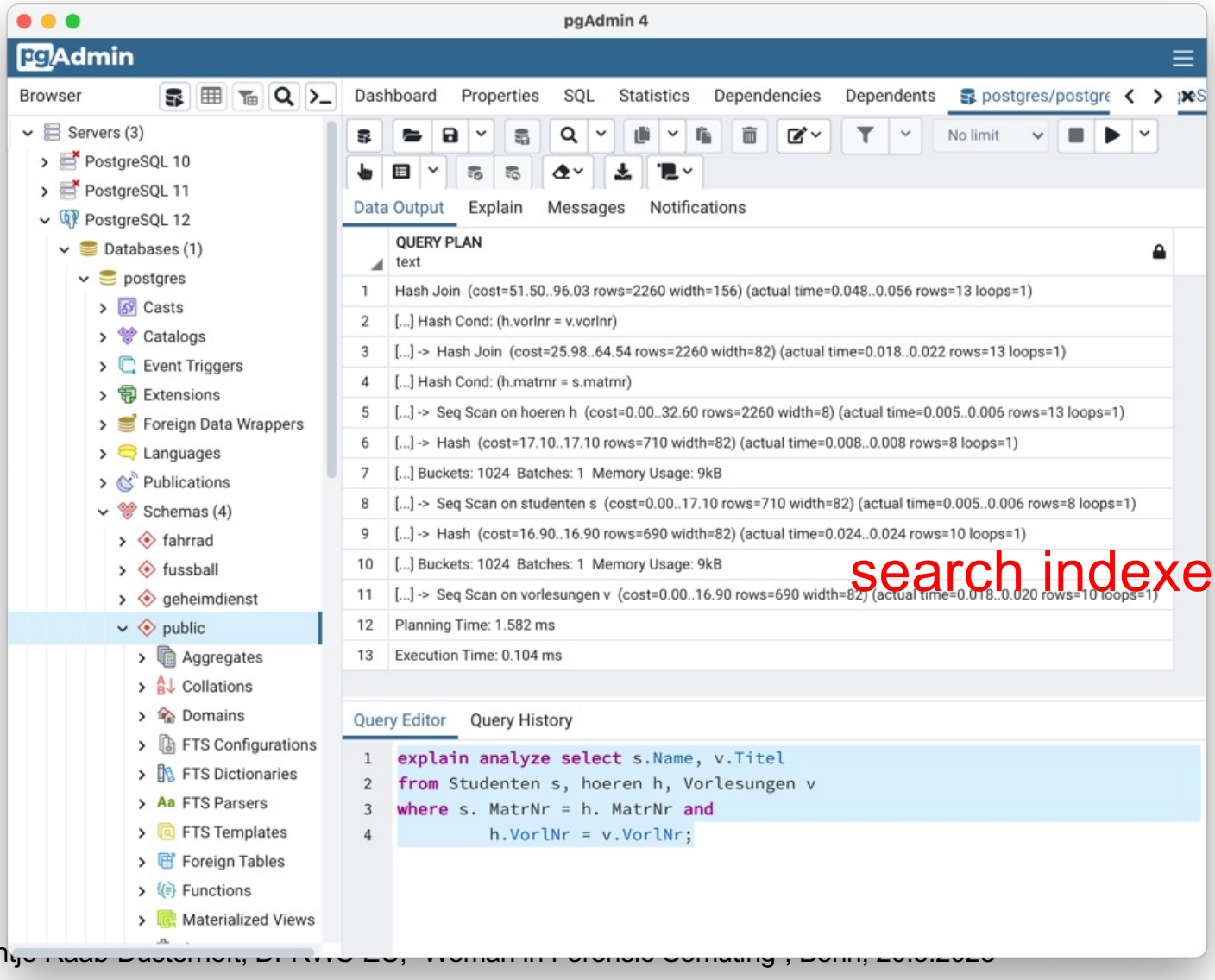
4. Artefact Analysis: Data files (Postgres)

B-tree



Index information in the information scheme of the db

4. Artifact Analysis: Query plan (Postgres)



Conclusion: Strategy for a forensic database analysis

1. Preparation
   - Defining or dentifiying the scope/ infrastructure
   - Identifiying DBMS, version, configuration, instances, ...
   - Classification of the database model
   - Preparing forensic scripts and a forensic maschine

2. Incident verification
   - First data collection from the database
   - Using infrastructure information
   - Confirm the breach or incident
   - Decision about shut-down/ sperating the database server

Conclusion: Strategy for a forensic database analysis

3. Artifact collection
   - Collecting volatile and non-volatile database artitacts
   - Bit-to-bit image of the database data, query files, query output,..
   - Collecting data from the infrastructure

4. Artifact analysis
   - Decision of the analysis methods
   - Overview of the database use logical artifacts
   - Analyse logical artifacts
   - Analyse physical artifacts

5. Reporting

Conclusion: Issues to be solved

1. Different Database systems
   - Relational:
     - Different versions, configurations, ...
     - Frequent updates
     - Different internals and file structures
   - NoSQL (and others):
     - No/ few forensic strategies in general
     - Very different structure of the data systems
     - No/ few management functions
     - (probably easy) forensic examination of data files
2. Need of tools for data interpretation (logs, structures, internals)

Thank you for your attention!