

# A DWH in the Oracle Cloud

**From the bucket to the dashboard**

Moody: the DWH for Moodle

Dr. Andrea Kennel, fhnw



# Dr. Andrea Kennel

**.SYM** 



Consultant

Lecturer for Databases

Coach for Project Management

University of Applied Sciences  
Northwestern Switzerland

Brugg/Windisch, Switzerland



[andrea.kennel@fhnw.ch](mailto:andrea.kennel@fhnw.ch)  
[andrea@infokennel.ch](mailto:andrea@infokennel.ch)  
[www.infokennel.ch](http://www.infokennel.ch)

# Motivation

- Students write exams in Moodle LMS (Learning Management System)
- We want to improve the quality of examinations and examination questions
- We want to improve the quality of education
- The analysis options in Moodle are limited
  
- Building an exemplary DWH

# Moodle: Time analysis

## Antworten-Rückblick

Schritt	Zeit	Aktion	What	When	Time
1	3. März 2025, 13:23:04	Antworten-Rückblick	Start		13:23:04
2	3. März 2025, 13:25:25	Schritt Zeit	Leave question	1	13:25:25
3	3. März 2025, 13:27:43	1 3. März 2025, 13:23:04	Leave question	2	13:27:12
4	3. März 2025, 13:34:53	2 3. März 2025, 13:27:12	Leave question	1	00:01:47
		3 3. März 2025, 13:34:53	Leave question	3	00:00:31
			Leave question	2	00:00:31
				1	00:03:04
				3	00:02:52
				2	00:01:47
				1	00:03:04

# Moodle DWH: Time analysis

Moodle-DWH

Prüfung wählen

MSP FS23

Hier klicken für übersichtlichere Chart-Darstellung

Durchschnittliche Bearbeitungszeit pro Frage

Alle anzeigen Balkendiagramm **Box-Plot**

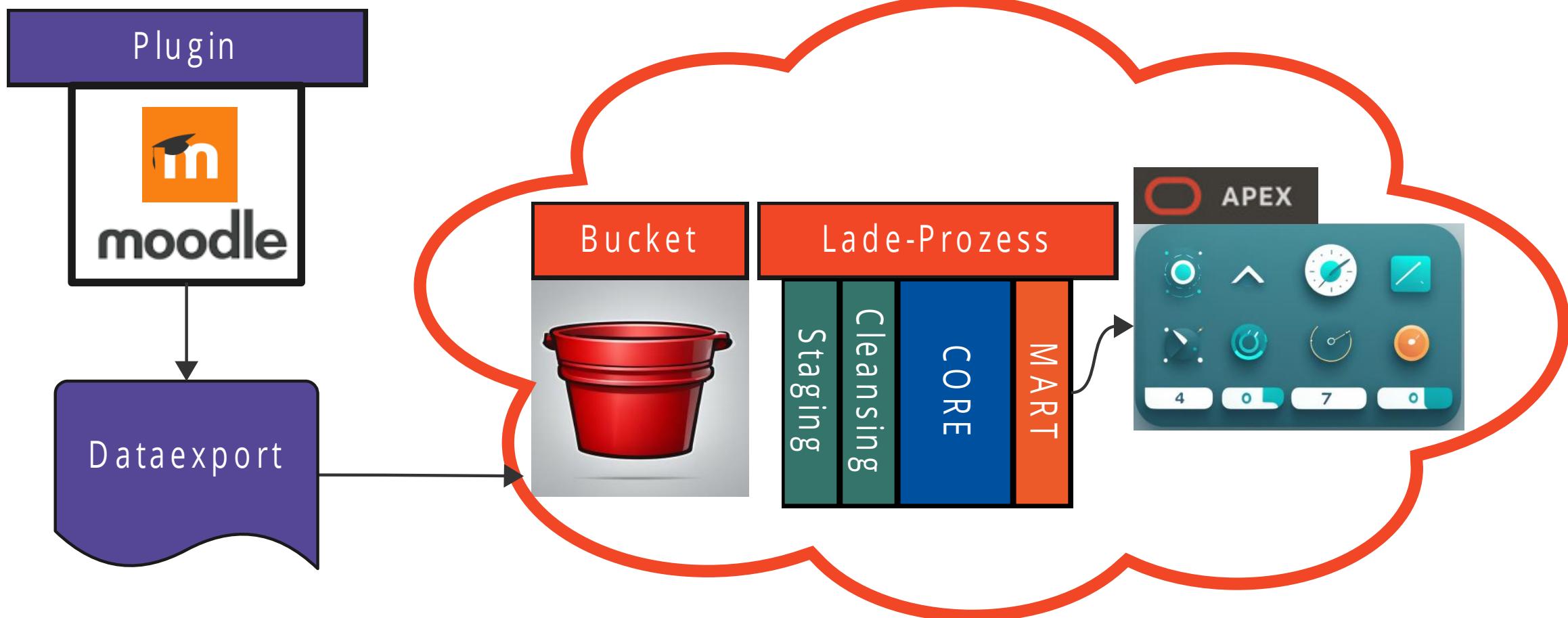
Box-Plot

The figure displays five horizontal box plots, each representing a different task. The tasks are labeled on the left: '01 Aufgabe eines DBMS 2', '02 Aufgaben eines DBMS 1', '03 ERD erstellen', '04 GROUP - Variante 1', and '05 GROUP - Variante 2'. Each box plot shows the distribution of editing times for that specific task. The boxes represent the interquartile range (IQR), the median is indicated by a horizontal line inside the box, and whiskers extend to the minimum and maximum values. Outliers are shown as individual blue dots.

Aufgabe	Median (approx.)	Q1 (approx.)	Q3 (approx.)	Min (approx.)	Max (approx.)
01 Aufgabe eines DBMS 2	10	8	12	5	20
02 Aufgaben eines DBMS 1	10	8	12	5	25
03 ERD erstellen	15	10	20	5	40
04 GROUP - Variante 1	15	10	20	5	30
05 GROUP - Variante 2	10	8	12	5	25

<https://gdbb147251565c3-moodykea.adb.eu-frankfurt-1.oraclecloudapps.com/ords/#p4-box-chart>

# Overview



# From Moodle to the bucket

Diagram illustrating the flow from Moodle to the Bucket:

```
graph TD; Moodle[Plugin moodle] --> Dataexport[Dataexport]; Dataexport --> Bucket[Bucket]
```

The screenshot shows the creation of a Pre-Authenticated Request (PAR) in an Object Storage interface. The 'Pre-Authenticated Requests' section is active.

**Create Pre-Authenticated Request**

**Pre-Authenticated Request Details**

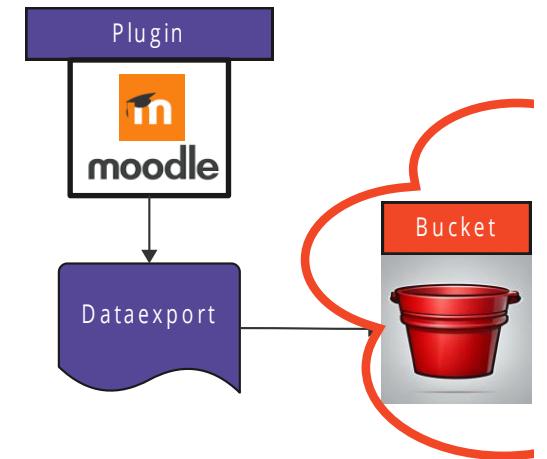
- Backend:** DWH\_24 (selected)
- Description:** Bucketzugriff bis Dez 24
- URL:** https://objectstorage.eu
- Username:** [empty]
- Password:** [empty]
- Enabled:** Ja
- Allowed user:** [empty]

**Bucket Configuration:**

- Name:** par-bucket-20231121-1131
- Access Type:** Permit object read (selected)
- Expiration:** Nov 28, 2023 10:31 UTC

**Note:** The current URL is deprecated and will no longer be supported in a future release of the c below. [Learn more](#)

**Buttons:** Create Pre-Authenticated Request, Cancel, Änderungen speichern, Abbrechen



# From Moodle to the bucket with PHP-function cURL\_exec

```
// PUT to a Pre-Authenticated Requests enabled Oracle Object Storage Bucket.  
$url = $DB->get_field('report_datawarehouse_bkends', 'url', ['id' => $backendid]);  
// Initiate cURL object.  
$curl = curl_init();  
// Set your URL.  
curl_setopt($curl, CURLOPT_URL, $url . $filename);  
// Indicate your protocol.  
curl_setopt($curl, CURLOPT_PROTOCOLS, CURLPROTO_HTTPS);  
// Set HTTP method to PUT.  
curl_setopt($curl, CURLOPT_PUT, 1);  
// Indicate the file you want to upload.  
curl_setopt($curl, CURLOPT_INFILE, fopen($tempfolder . '/' . $filename, 'rb'));  
...  
// Execute.  
curl_exec($curl);
```

# Data protection

User is double anonymised

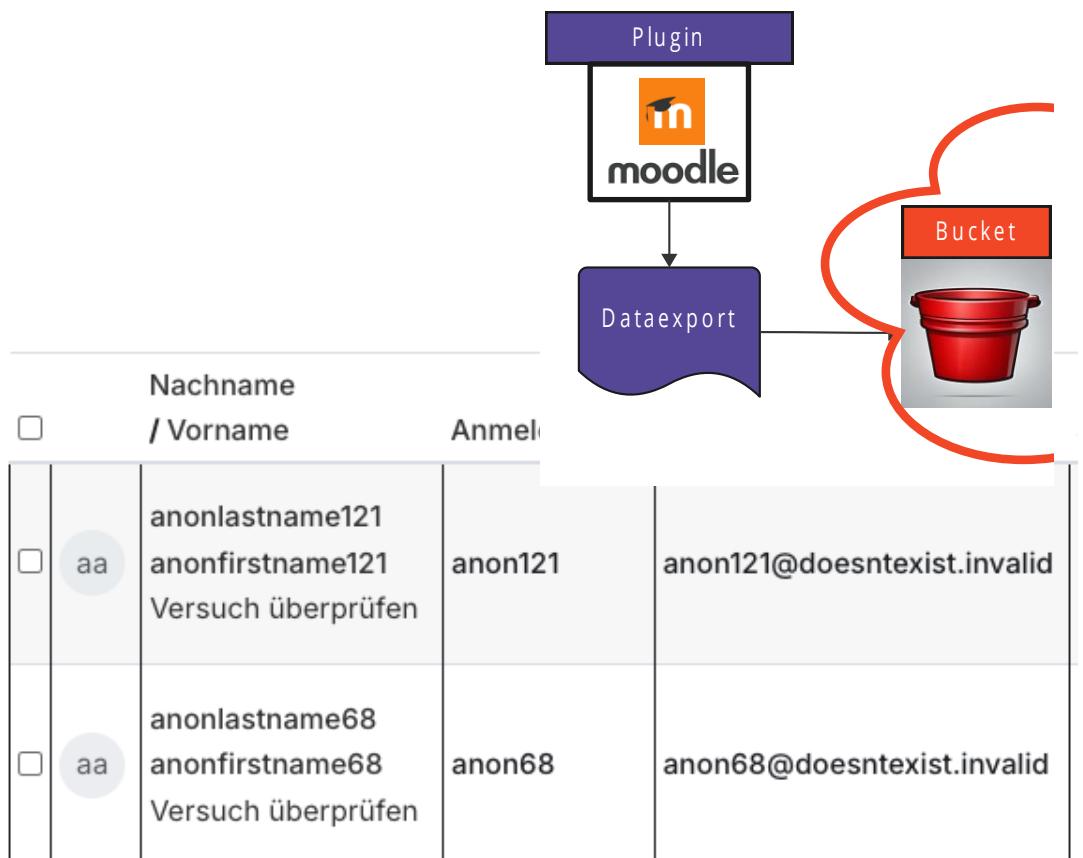
1. Anonymise users in Moodle

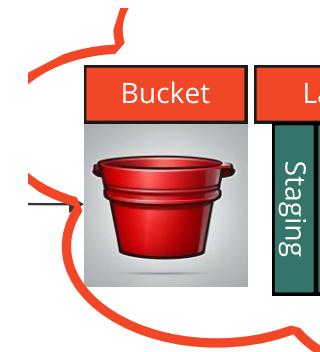
2. Export hash only

SELECT

```
...  
  (SELECT MD5(username) FROM mdl_user WHERE id = qzat.userid) AS "qzatuserid",  
...
```

3. Data ‘knocked flat’ as csv into the bucket





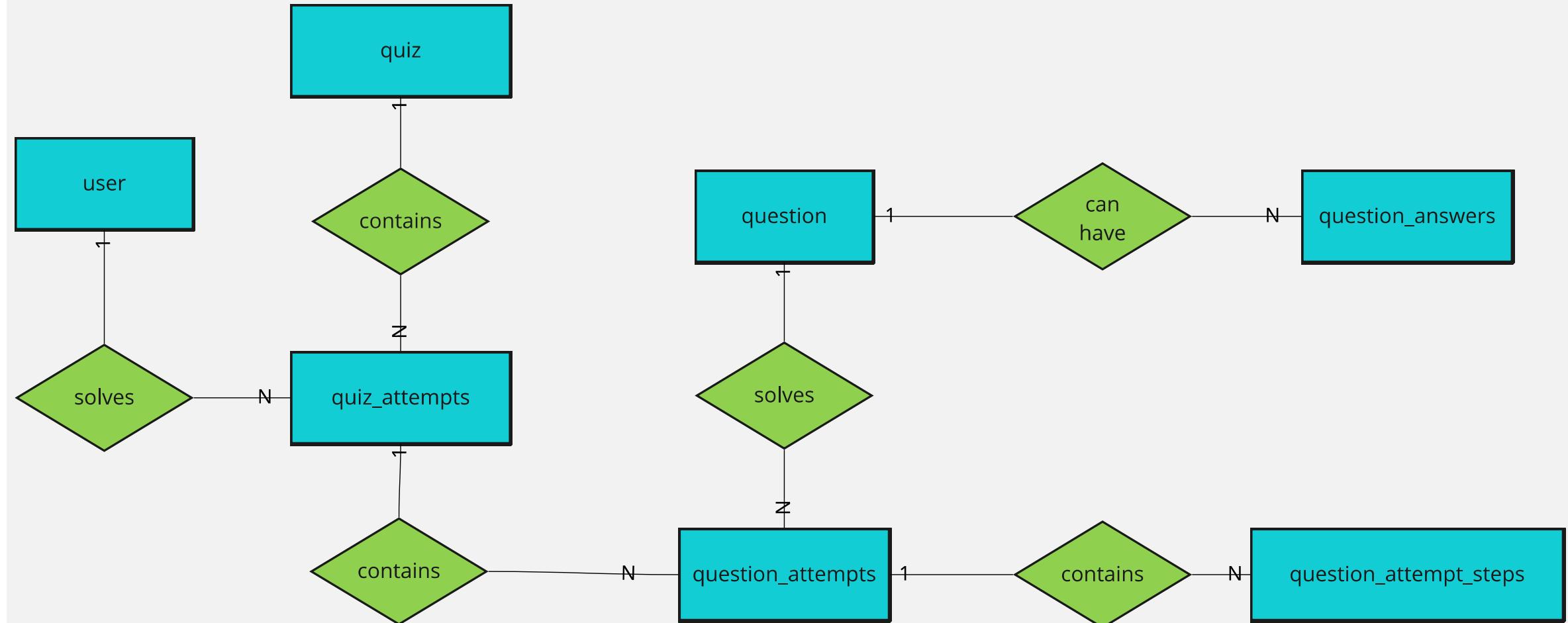
# From the bucket to the staging with DBMS\_CLOUD.COPY\_DATA

```
"C##CLOUD$SERVICE"."DBMS_CLOUD"."COPY_DATA"  
( TABLE_NAME      => l_TABLE_NAME  
,CREDENTIAL_NAME => l_CREDENTIAL_NAME  
,FILE_URI_LIST   => l_FILE_URI_LIST  
,FIELD_LIST       => l_FIELD_LIST  
,FORMAT          => l_FORMAT  
,OPERATION_ID     => l_OPERATION_ID  
);
```

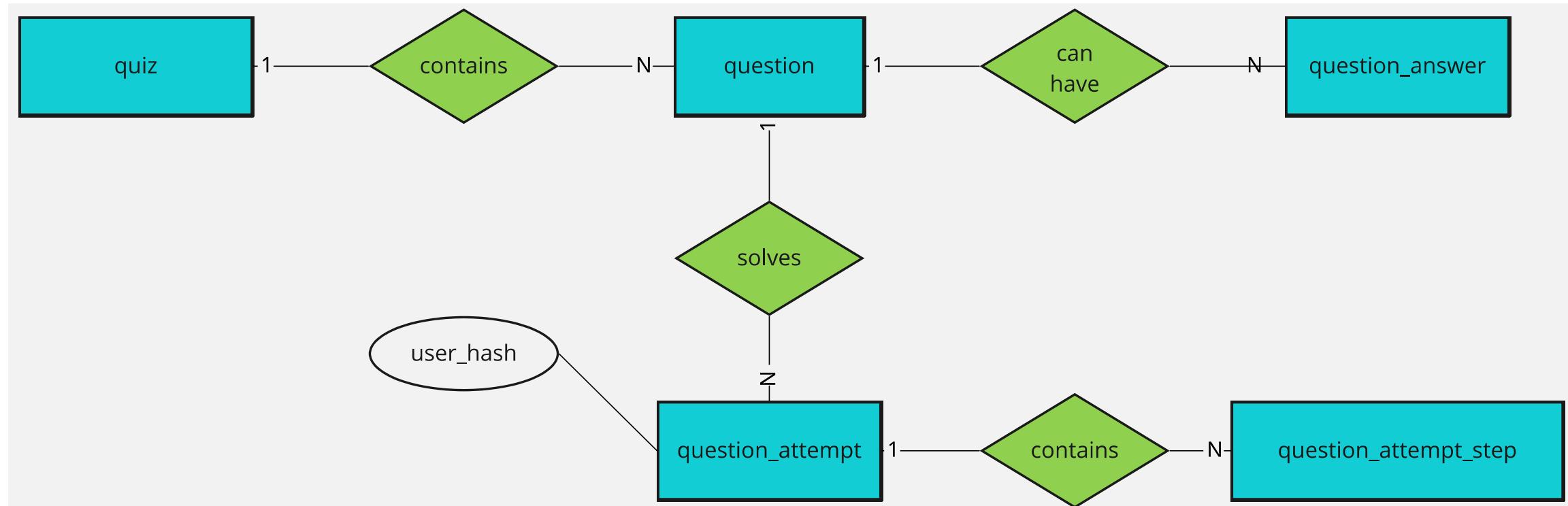
```
--  
v_full_path VARCHAR2(2000) := p_bucket_path || p_file_name;  
  
-- DECLARE  
l_TABLE_NAME      DBMS_QUOTED_ID := '"DATA_VERSION_BACH"';  
l_CREDENTIAL_NAME DBMS_QUOTED_ID := '"OBJ_STORE_CRED"';  
l_FILE_URI_LIST   CLOB := v_full_path;  
l_FIELD_LIST      CLOB :=  
q'[  
    "SOURCESYSTEM"           CHAR(4000)  
    , "AUTHORIZEDUSERS"        CHAR(4000)  
    , "QUESTIONID"             CHAR(4000)  
];
```

Data from CSV 1:1 into a table

# The source data model (extract)



# The model in the cleansing

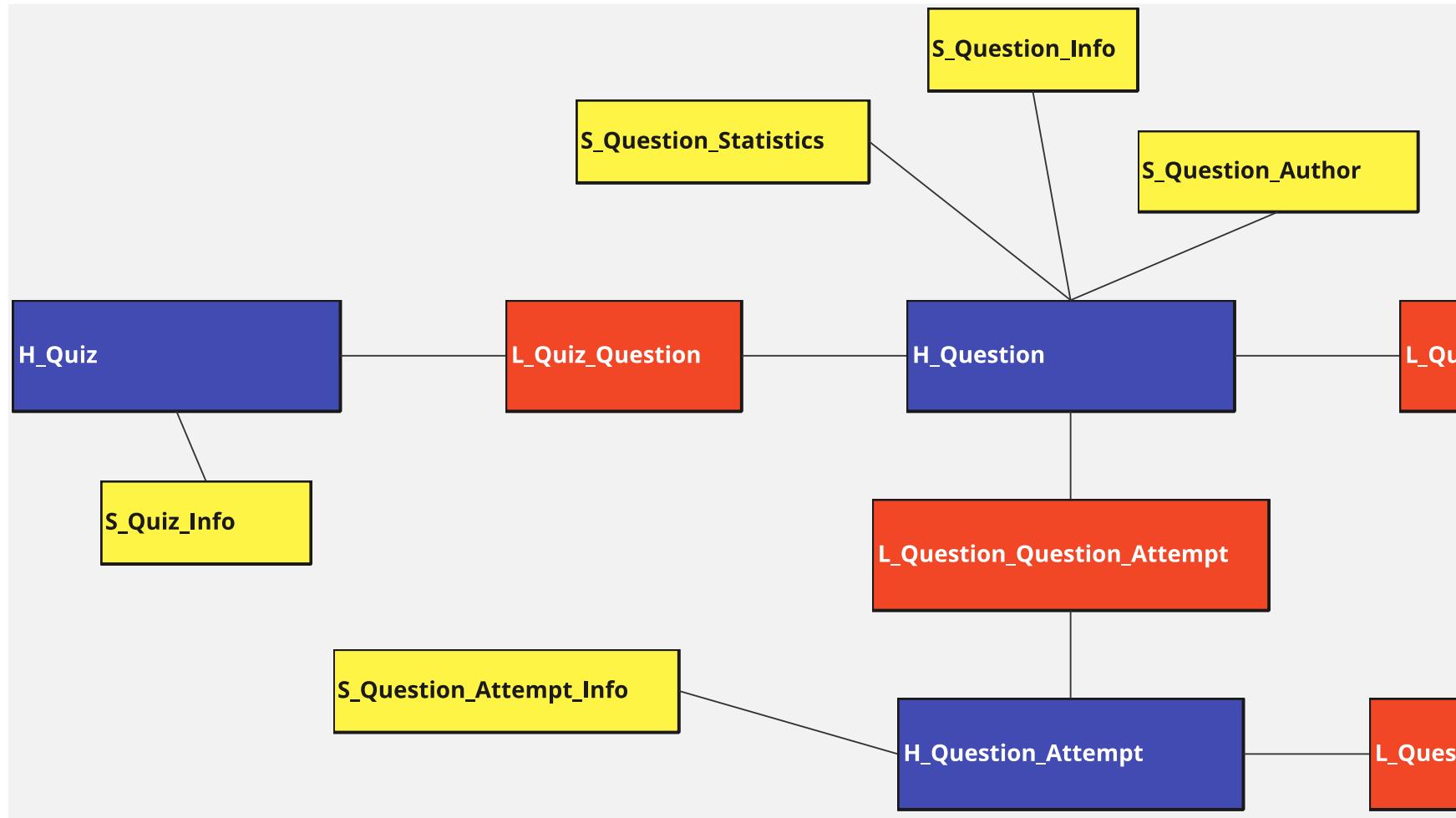


# From the staging to the cleansing

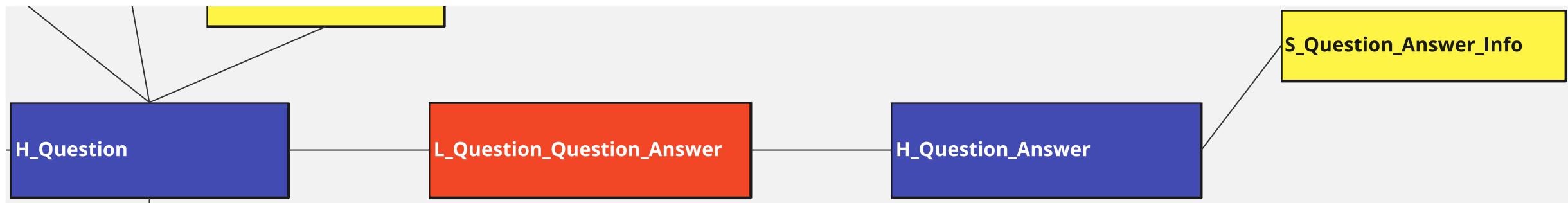
```
INSERT INTO MD_CLEANSING.quiz (
    load_id,
    sourcesystem,
    quiz_id,
    authorizedusers,
    quiz_name,
    quiz_grade)
SELECT DISTINCT
    p_load_id load_id,
    sourcesystem,
    quizid quiz_id,
    authorizedusers,
    quizname quiz_name,
    quizgrade quiz_grade
FROM MD_STAGING.data_view d;
```

```
INSERT INTO MD_CLEANSING.question_attempt_step_data (
    load_id,
    sourcesystem,
    authorizedusers,
    question_attempt_step_data_id,
    question_attempt_step_id,
    step_data_name,
    step_data_value)
SELECT DISTINCT
    p_load_id load_id,
    sourcesystem sourcesystem,
    authorizedusers authorizedusers,
    questionattemptstepdataid question_attempt_step_data_id,
    qastepid question_attempt_step_id,
    questionattemptstepdataname step_data_name,
    questionattemptstepdatavalue step_data_value
FROM MD_STAGING.data_view
WHERE questionattemptstepdataid IS NOT NULL;
```

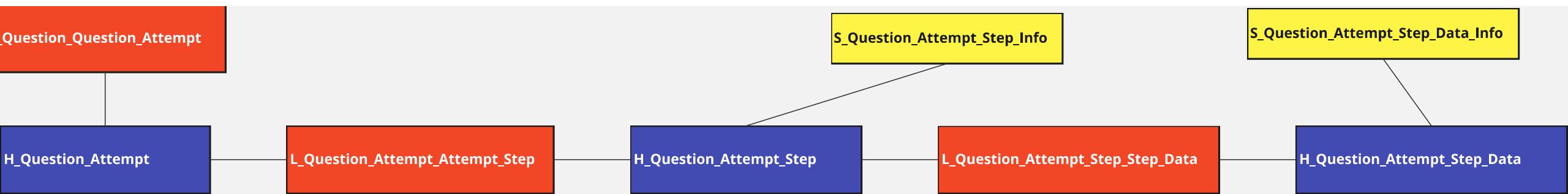
# The model in the core



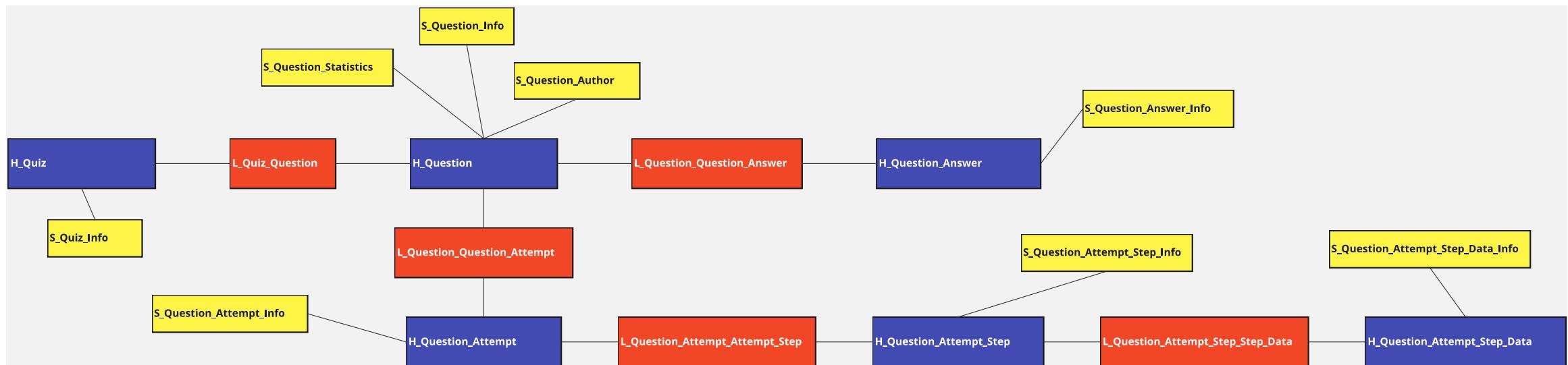
# The model in the core

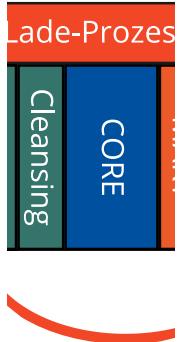


# The model in the core



# The model in the core





# From the cleansing to the core

1. Fill HUB

H\_Question

Surrogate key for HUB

h\_question\_sid

Source keys

question\_id  
load\_id

# From the cleansing to the core

2. Fill Satellites for HUB

S\_Question\_info

Find foreign key to HUB via  
source key

h\_question\_sid

```
SELECT
    h.h_question_sid,
    cl.question_type, cl.question_name,
    cl.question_text
FROM MD_CLEANSING.question cl INNER JOIN
    H_Question h ON (cl.question_id = h.question_id
                    AND cl.load_id = h.load_id);
```

# From the cleansing to the core

3. Fill Link between HUB

L\_Question\_Question\_Answer

Surrogate key for Link

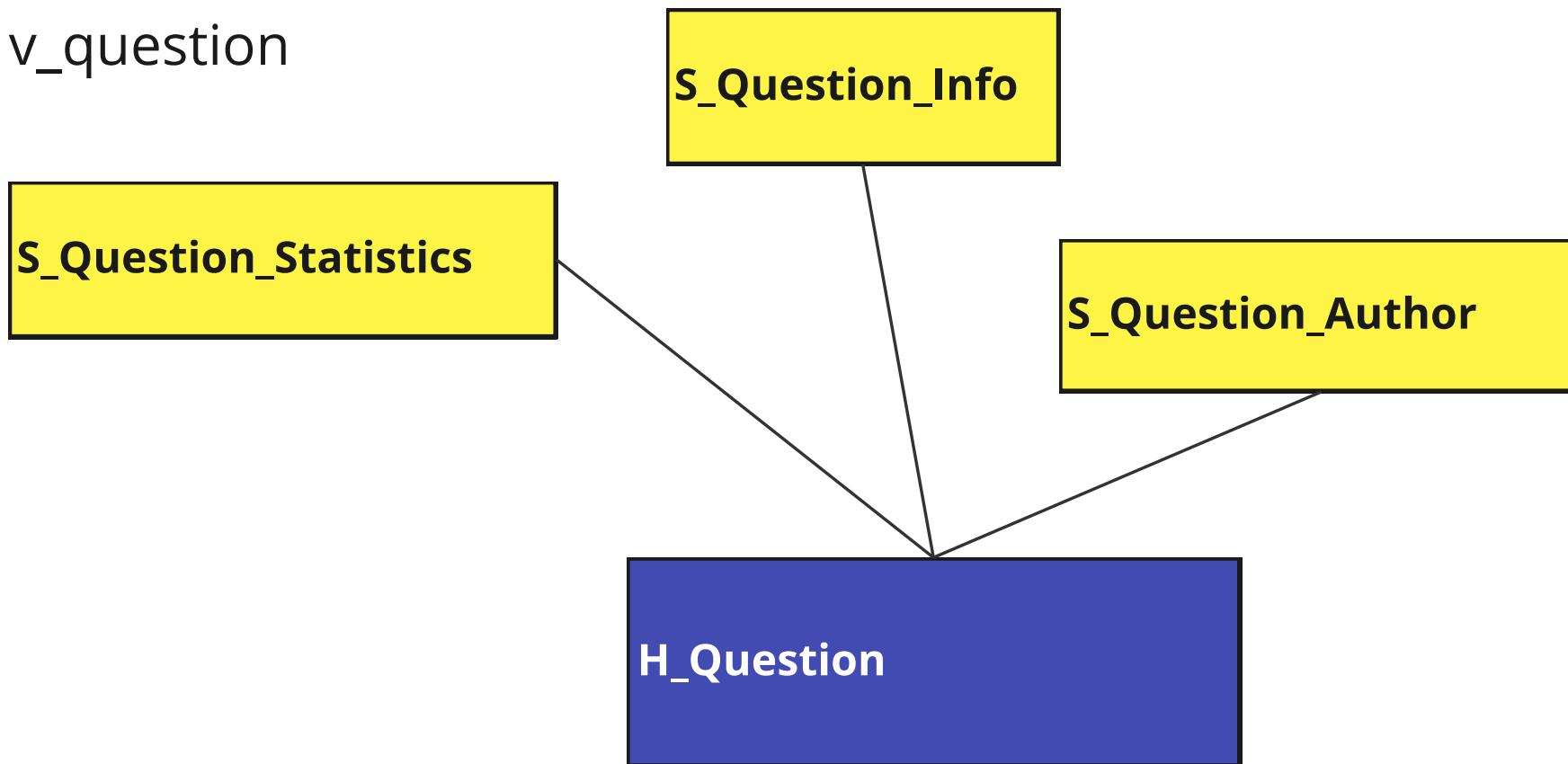
l\_question\_question\_answer\_sid

Find both foreign keys to HUB  
via source key

h\_question\_sid  
h\_question\_answer\_sid

# Views for easier access to core

v\_question



# Views for easier access to core

v\_question\_answer



S\_Question\_Answer\_Info

# Views for easier access to core

v\_question\_attempt

L\_Question\_Question\_Attempt

S\_Question\_Attempt\_Info

H\_Question\_Attempt

# MART Layer

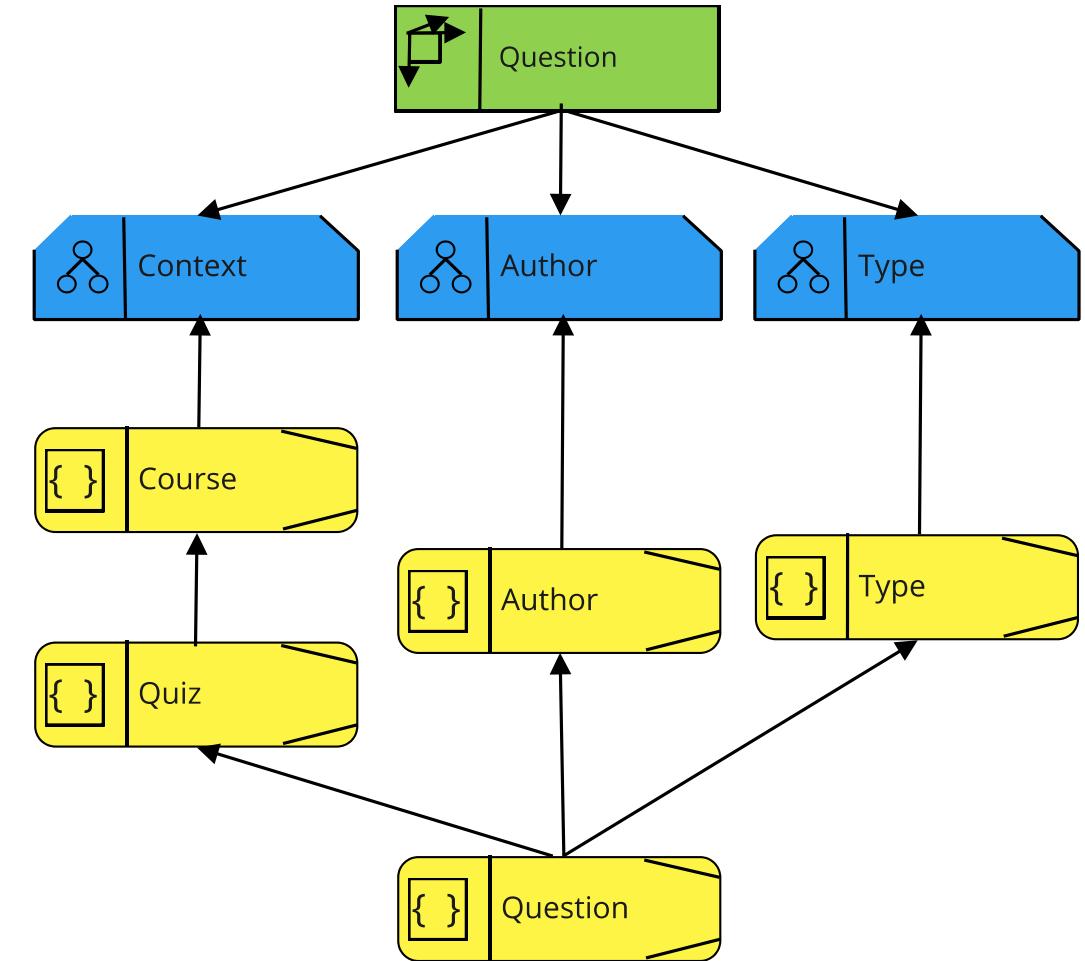
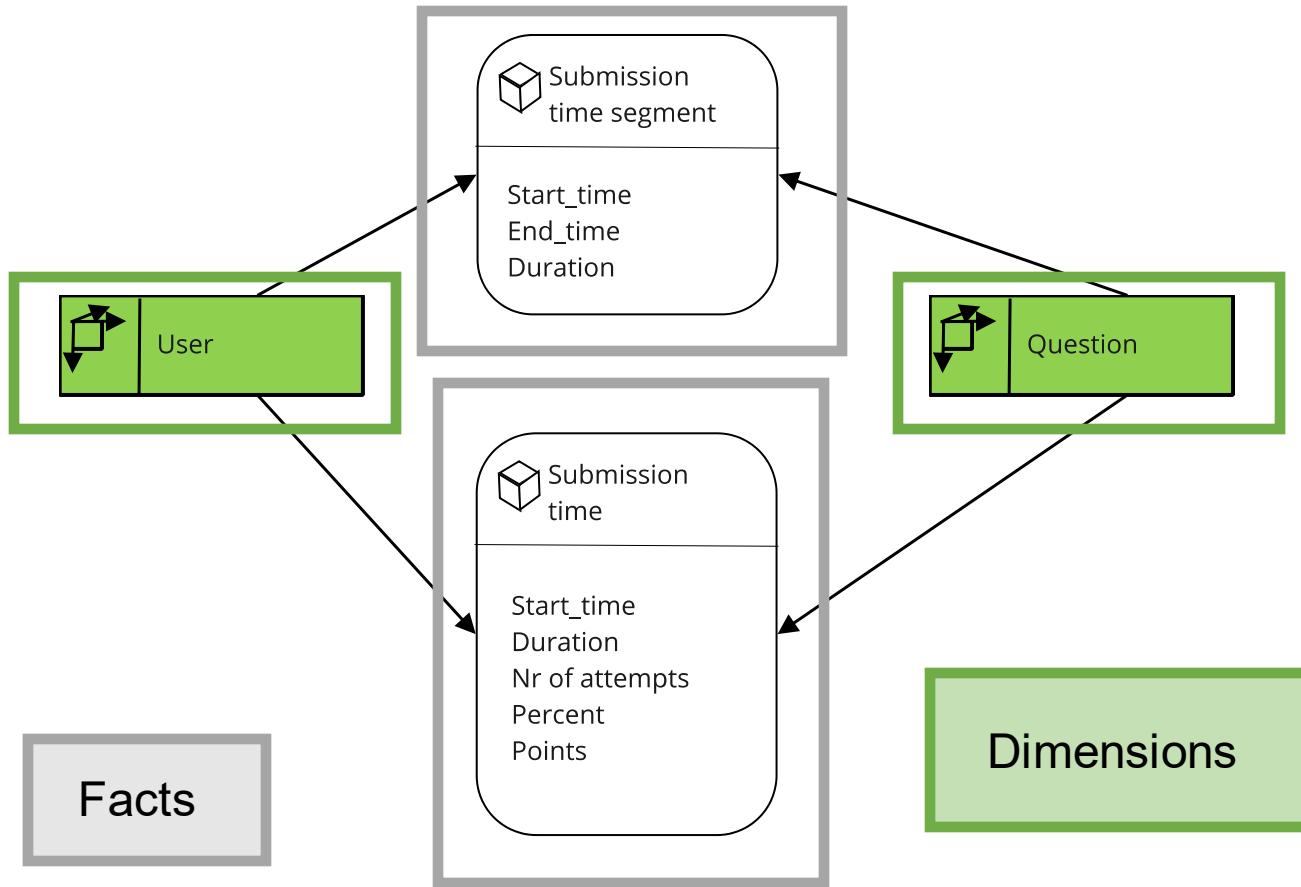


Purpose of the mart: Provide data for analyses

- Different approach than in the core: Tailored to the needs of end users.
- Different concepts and conventions for modelling

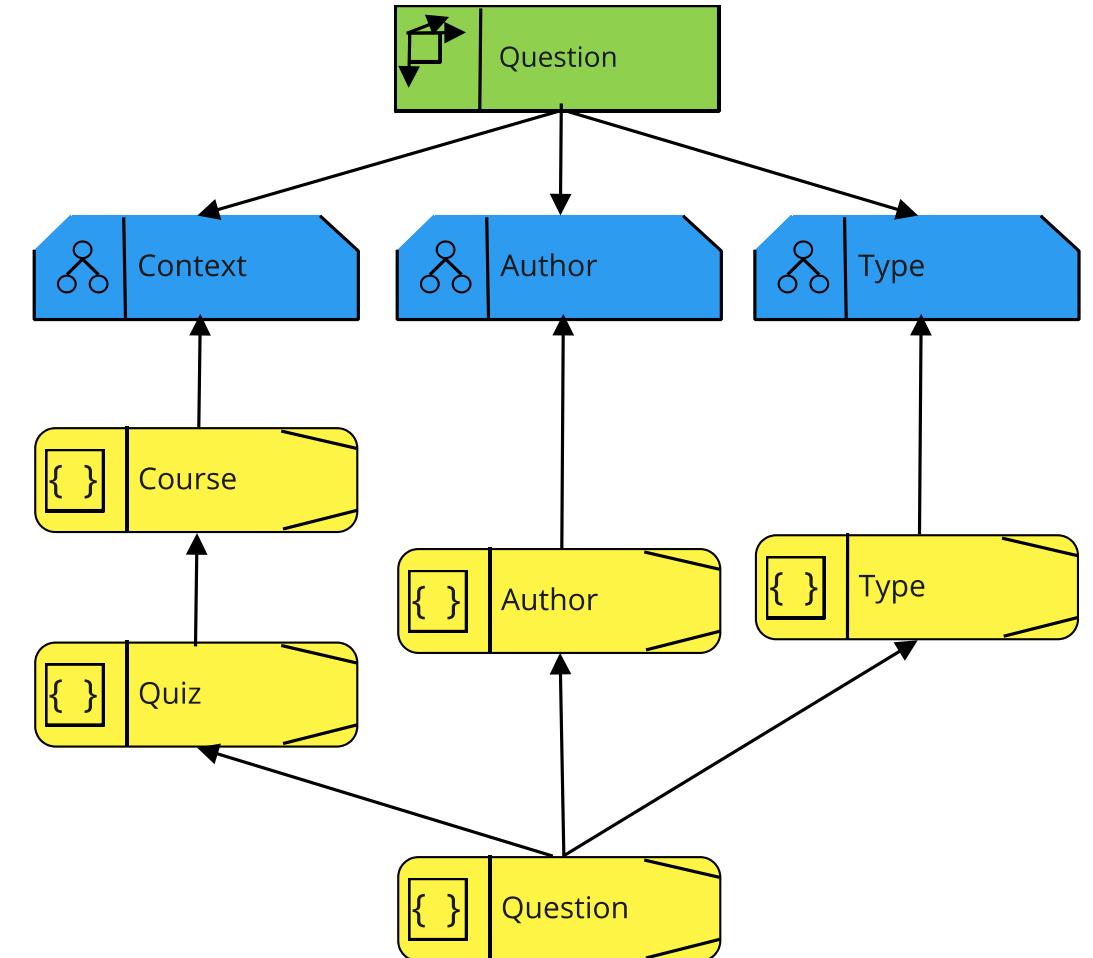
Bilder: Adobe Stock

# Time analysis: ADAPT modelling

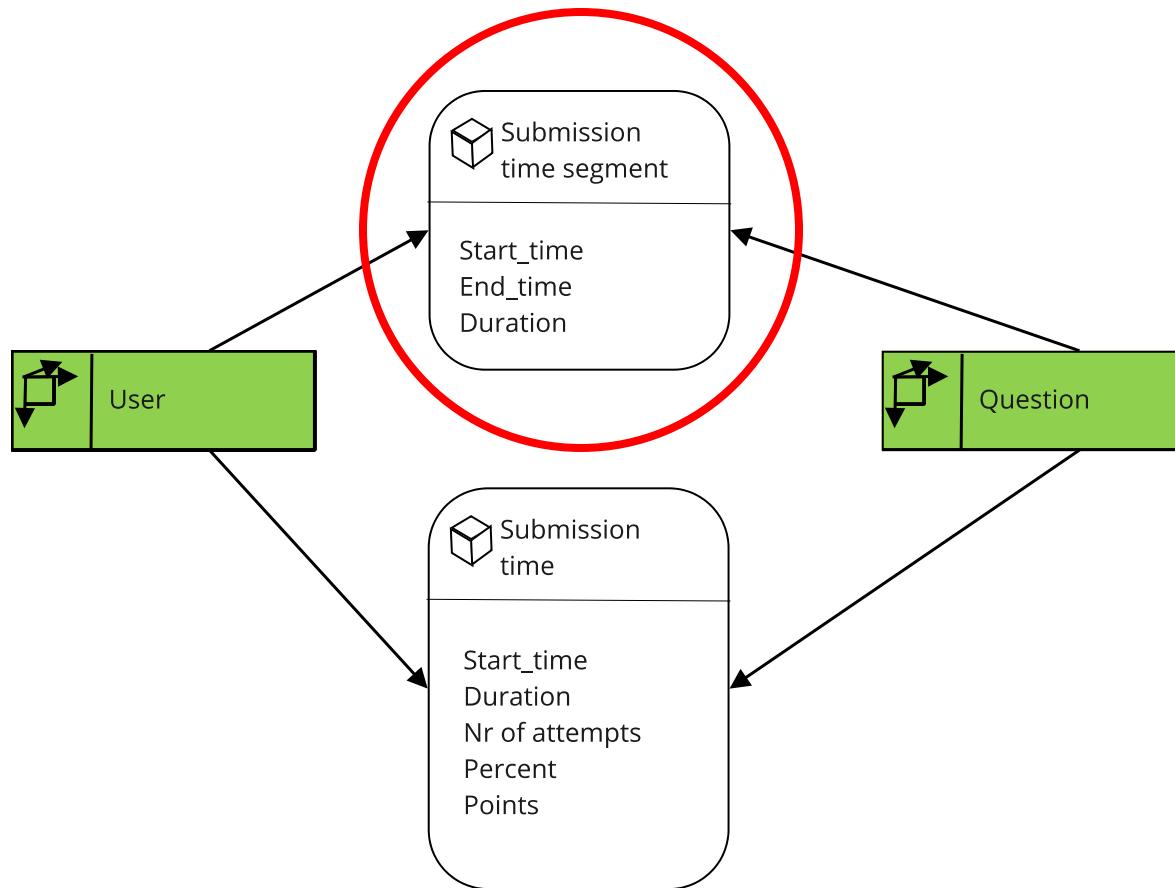


# From the core to the mart – Question

```
INSERT INTO MD_MART.DM_D_Question
SELECT
    seq_md_id_question.nextval MD_ID_Question,
    q.question_id QU_MD_BK_question_id,
    q.load_id QZ_MD_BK_load_id,
    q.question_name QU_question_name,
    q.question_text QU_question_text,
    ...
    q.question_type TY_question_type
FROM MD_CORE.v_question q
WHERE q.load_id = this_load_id;
```



# From the core to the mart – Submission time segment



# Moodle: Time analysis

## Antworten-Rückblick

Schritt	Zeit	Aktion	What	When	Time
1	3. März 2025, 13:23:04	Antworten-Rückblick	Start		13:23:04
2	3. März 2025, 13:25:25	Schritt Zeit	Leave question	1	13:25:25
3	3. März 2025, 13:27:43	1 3. März 2025, 13:23:04	Leave question	2	13:27:12
4	3. März 2025, 13:34:53	2 3. März 2025, 13:27:12	Leave question	1	00:01:47
		3 3. März 2025, 13:34:53	Leave question	3	00:00:31
			Leave question	2	00:00:31
				1	00:03:04
				3	00:02:52
				2	00:01:47
				1	00:03:04

# From the core to the mart – Submission time segment

```
INSERT INTO MD_MART.DM_F_Question_Submission_Time_Segment
SELECT
    MD_ID_Question,
    quiz_attempt_user_hash MD_ID_Candidate,
    time_prev_start_time,
    createdunixtime end_time, time_spent duration, load_id
FROM (
    SELECT
        d_q.MD_ID_Question,
        qa.quiz_attempt_user_hash,
        LAG(step.createdunixtime, 1, 0) OVER (PARTITION BY qa.QUIZ_ATTEMPT_USER_HASH
            ORDER BY step.createdunixtime) AS time_prev,
        step.createdunixtime -
        LAG(step.createdunixtime, 1, 0) OVER (PARTITION BY qa.QUIZ_ATTEMPT_USER_HASH
            ORDER BY step.createdunixtime) AS time_spent
    FROM md_core.V_QUESTION_ATTEMPT qa JOIN
        MD_CORE.V_QUESTION_ATTEMPT_STEP step
        ON qa.H_QUESTION_ATTEMPT_SID = step.H_QUESTION_ATTEMPT_SID JOIN
        ...
        MD_MART.DM_D_Question d_q ON d_q.QU_MD_BK_question_id = q.question_id AND d_q.QZ_MD_BK_load_id = q.load_id
    WHERE qa.load_id = this_load_id)
```

# From the core to the mart – Submission time segment

Important: Test data!

Specially created and timed Moodle exam  
→ Protocol

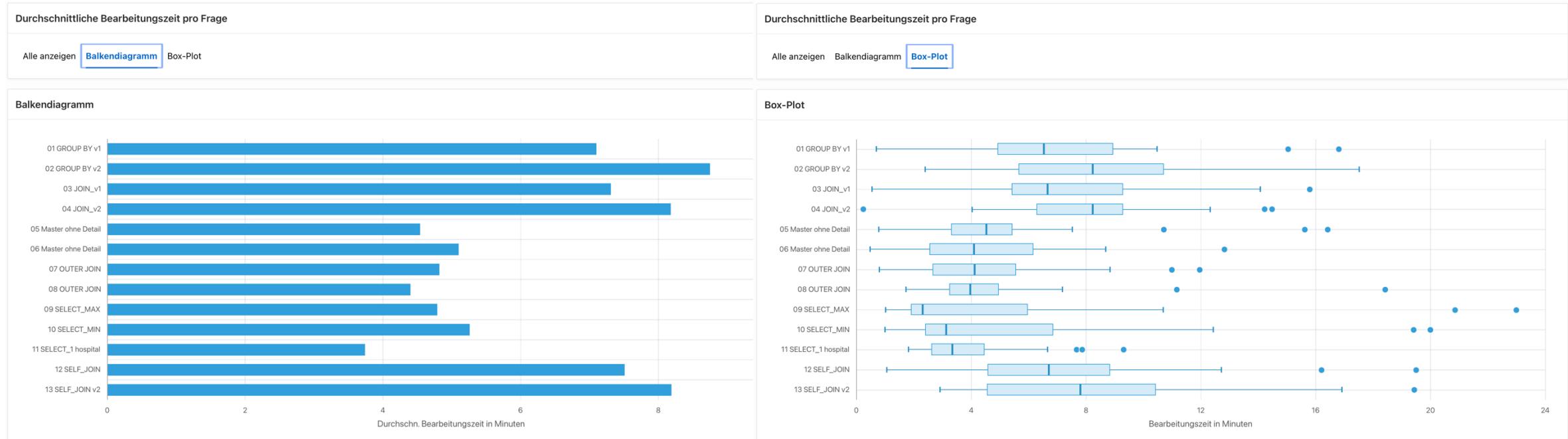


# Interface for End User: Oracle APEX

The screenshot shows the Moodle-DWH application interface. At the top, there is a blue header bar with the title "Moodle-DWH". Below the header is a navigation sidebar on the left containing three items: "Home" (with a house icon), "Bearbeitungsdauer" (with a document icon), and "MC, MTF, KPrime" (with a document icon). The main content area features a blue icon of a database and the text "Moodle-DWH". Below this, a large white box contains the heading "Willkommen bei Moodle-DWH" and two links: "Was ist Moodle-DWH?" and "Anleitung für Dozierende".



# Interface for End User: Oracle APEX





# From the core to the mart – MC answer option

Choise of answer option

STEP_DATA_NAME	STEP_DATA_VALUE
choice0	1
choice1	1
choice2	0
choice3	0
choice4	0
_order	680,684,681,683,682

???

ChatGPT!

Answer option

QUESTIONANSWERSID	ANSWER
683	<p dir="ltr"></p><p dir="ltr">SELECT gruppen_name</p><p c
681	<p dir="ltr" style="text-align: left;"></p><p dir="ltr"><
684	<p></p><p dir="ltr">SELECT ig.gruppen_name</p><p dir="ltr
682	<p dir="ltr"></p><p dir="ltr">SELECT gruppen_name</p><p c
680	<p dir="ltr"><span style="font-size: 0.9375rem;">SELECT i



# From the core to the mart – MC answer option

...

SELECT

```
c.step_data_name,  
c.original_step_data_value,
```

```
REGEXP_SUBSTR(o.order_values, '[^,]+', 1, _order
```

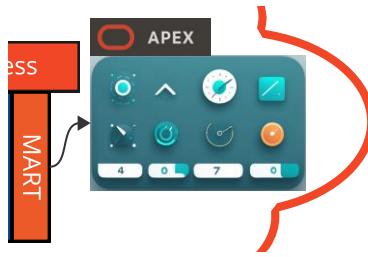
```
CAST(SUBSTR(c.step_data_name, 7) AS INT) + 1)
```

```
AS questionanswersid,
```

...

STEP_DATA_NAME	STEP_DATA_VALUE
choice0	1
choice1	1
choice2	0
choice3	0
choice4	0
_order	680,684,681,683,682

STEP_DATA_NAME	ORIGINAL_STEP_DATA_VALUE	QUESTIONANSWERSID
choice0	1	680
choice1	1	684
choice2	0	681
choice3	0	683
choice4	0	682



# From the core to the mart – MC answer option

## Hints for using ChatGPT:

- Provide context about the required part of the data model
- Precise wording (this requires that I know exactly what I want)
- Understanding of the data model
- Understanding how the data is related
- Test cases → Response received can be checked
- Step-by-step approach



# MC answer option in APEX

Die Anzahl der Datensätze der Ergebnismenge von 4. ist grösser als die der Ergebnismenge von 1.

48

55

103

Die Anzahl der Datensätze von 2. stimmt mit der Anzahl der Datensätze von 3. überein.

81

22

103

Jeder Datensatz der Ergebnismenge von 1. ist in der Ergebnismenge von 2. enthalten.

53

50

103

Jeder Datensatz der Ergebnismenge von 1. ist in der Ergebnismenge von 3. enthalten.

51

52

103

korrekt  
inkorrekt

# First extension: kprime and mtf

Markieren Sie den gewünschten Fragetyp ×

-  Kprim (ETH)
-  Lückentext (Cloze)
-  Lückentextauswahl
-  MTF (ETH)

Wählen Sie einen Fragetyp, um  
seine Beschreibung zu sehen.

Problem:  
Require other data from other source tables

# First extension: kprime and mtf

- Analysing the data model in the source
- Export of the data, analysis in Excel
- Adaptations in all layers of the DWH

# Pitfalls

1. Several versions of data extracts
2. Staging: new table
3. Cleansing: tables with more attributes
4. Core: new satellites, hubs and links
5. Mart: more marts

What is the easiest way to tackle this?

# Pitfalls

- 1. Several versions of data extracts
  - 2. Staging: new table
  - 3. Cleansing: tables with more attributes
  - 4. Core: new satellites, hubs and links
  - 5. Mart: more marts
- 1. Attributes can be NULL

# Pitfalls

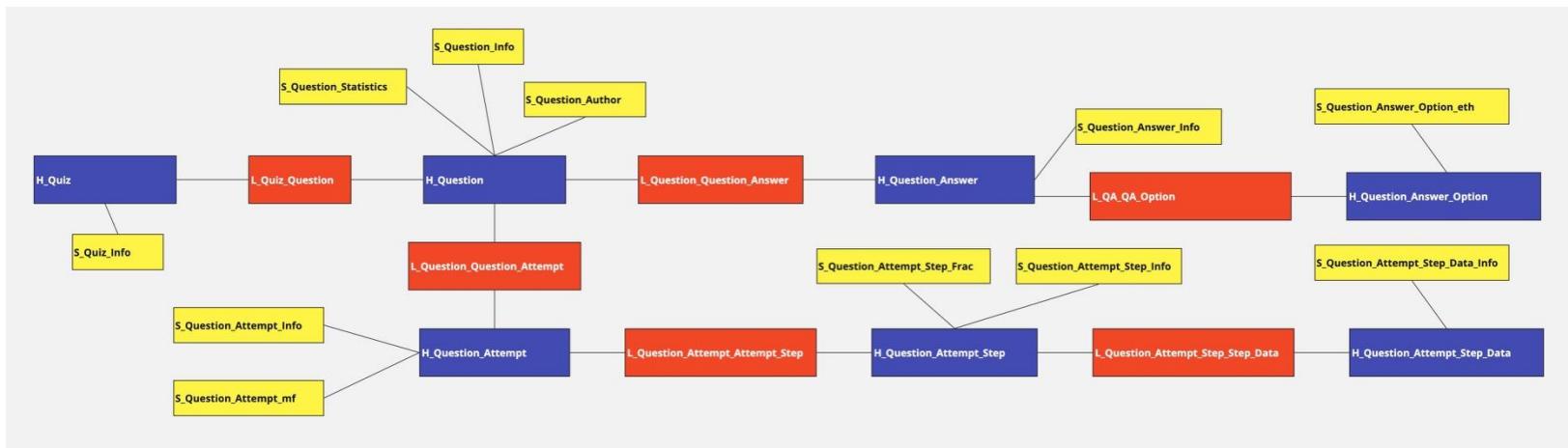
1. Several versions of data extracts
  2. Staging: new table
  3. Cleansing: tables with more attributes
  4. Core: new satellites, hubs and links
  5. Mart: more marts
2. Staging always only contains data from the current load
- One table per version
- View that merges all tables with UNION
- Code version-dependent

# Pitfalls

1. Several versions of data extracts
  2. Staging: new table
  3. Cleansing: tables with more attributes
  4. Core: new satellites, hubs and links
  5. Mart: more marts
3. Cleansing always only contains data from the current load
- Tables get more attributes
- New tables
- Code version dependent

# Pitfalls

1. Several versions of data extracts
2. Staging: new table
3. Cleansing: tables with more attributes
4. Core: new satellites, hubs and links
5. Mart: more marts



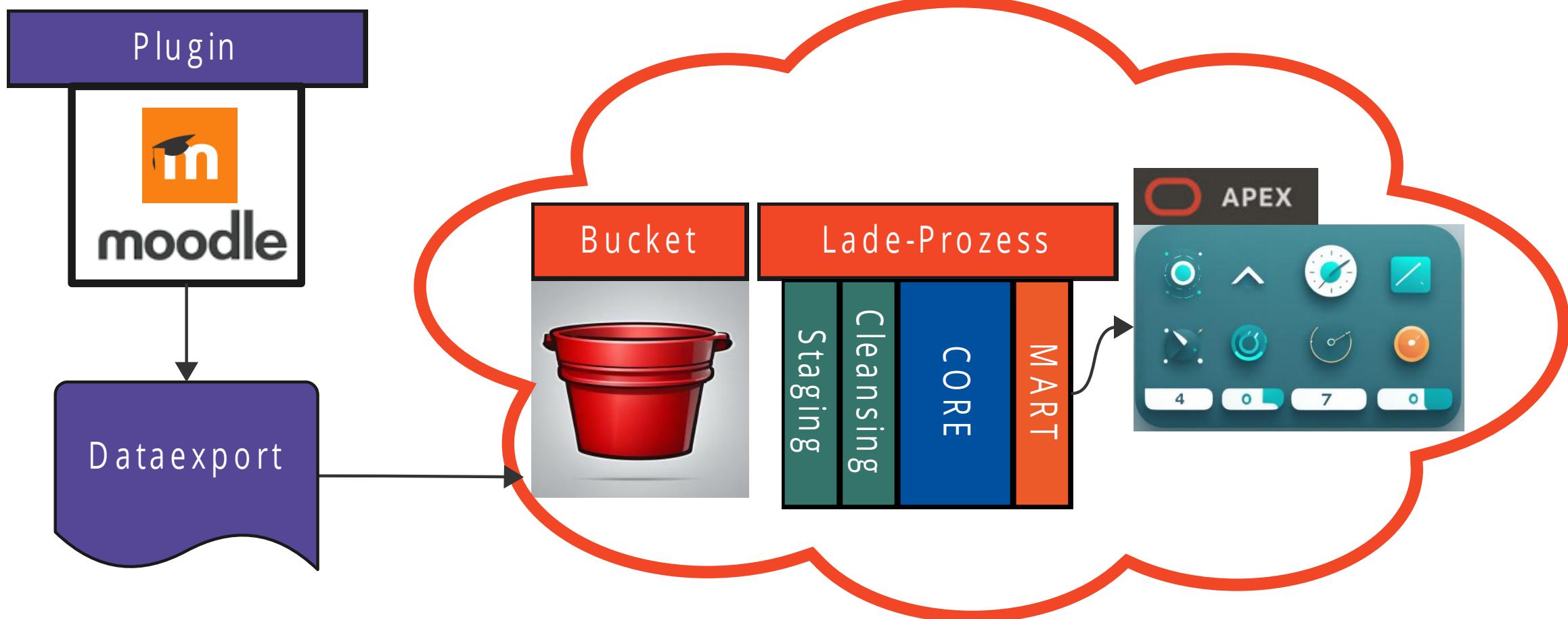
# Flexible solutions

1. View for staging
2. Always delete data in cleansing
3. Core as Data Vault
4. Views on core
5. IF (NOT) EXISTS in scripts

# What could come next?

1. Evaluation of Multiple Choice ✓
2. Points achieved
3. Points achieved in correlation to duration
4. Export more than one test at a time
5. User Management

# Overview



# Conclusion

1. Pilot brings added value
2. Easy to expand
3. Flexible thanks to clear concepts

# Dr. Andrea Kennel

**.SYM** 



Consultant

Lecturer for Databases

Coach for Project Management

University of Applied Sciences  
Northwestern Switzerland

Brugg/Windisch, Switzerland



[andrea.kennel@fhnw.ch](mailto:andrea.kennel@fhnw.ch)  
[andrea@infokennel.ch](mailto:andrea@infokennel.ch)  
[www.infokennel.ch](http://www.infokennel.ch)