DIE ÖSTERREICHISCHE BIBLIOTHEKENVERBUND
UND SERVICE GMBH

obv sg

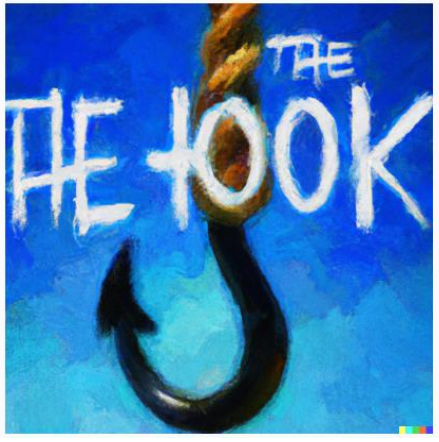
OFF THE HOOK
USING WEBHOOKS FOR PROVIDING FEEDBACK ON BIBLIOGRAPHIC DATA

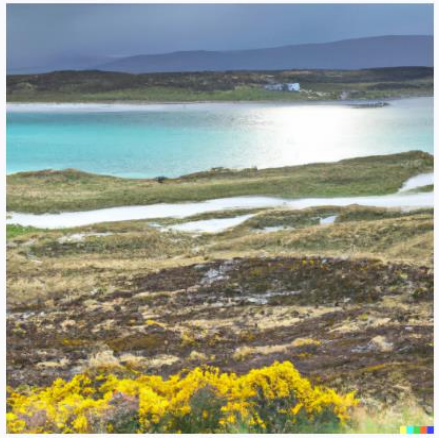
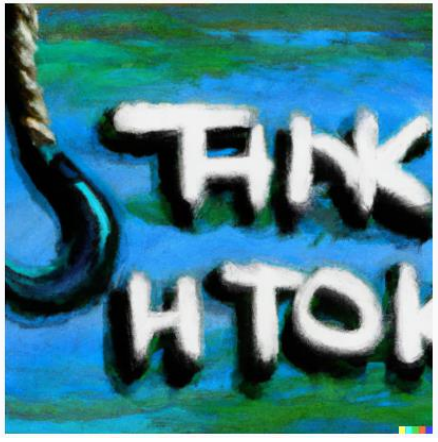
IGELU DEVELOPER DAY LEUVEN 2023-09-14

STEFAN MAJEWSKI

CONTENTS

- Description of our use-case
- Webhooks in their use
- What kind of components do we need?
- Abstraction of checking errors
- Fishing for errors
- The trawler in action
- Weaknesses



WHO WE ARE

AUSTRIAN LIBRARY NETWORK (AND SERVICE LTD)

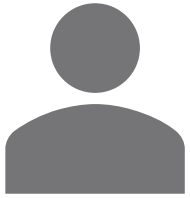
- Network of Scientific, Humanities and Administrative Libraries in Austria
- More than 90 Institutions, still growing
- Institutions range from large libraries like the Austrian National Library, large university libraries to smaller foundations
- Hosting a shared catalogue with network zones and institutional zones

DESCRIPTION OF OUR USE-CASE

THE USE-CASE

- Bibliographic data are at the heart of any library catalogue
- Nobody wants errors!
- Errors happen, no matter what
- We want to reduce the amount of user-facing errors
- We want to test records to find errors
- Whenever a record is changed it shall be checked
- The feedback shall be immediate
- The feedback shall be as actionable as possible

Librarians



Patrons

Institution

THE USE-CASE

THE MOST EFFICIENT TIME TO FIX AN ERROR

- We are talking about catalogue issues
- When the staff user is still working with the data
- As soon as possible, ideally when the user is still in context

Solution: Webhooks!

WEBHOOKS IN USE

WHY WEBHOOKS
WHAT THEY PROMISE

- Almost immediate notification on catalogue changes
- No dependency on publishing runs
- No dependency on rules when or what needs to be changed to be triggered.

- Which promises are kept?
  - The feedback is in near real-time
  - Whenever a person saves the record in metadata-record, an update event is triggered
  - Records are redelivered an hour later if the service is down for a short period of time

- Which hopes have been in vain?
  - Normalization Jobs (even smaller ones 50k records) do not trigger Webhook events.

WEBHOOKS
WHAT THEY CONTRIBUTE

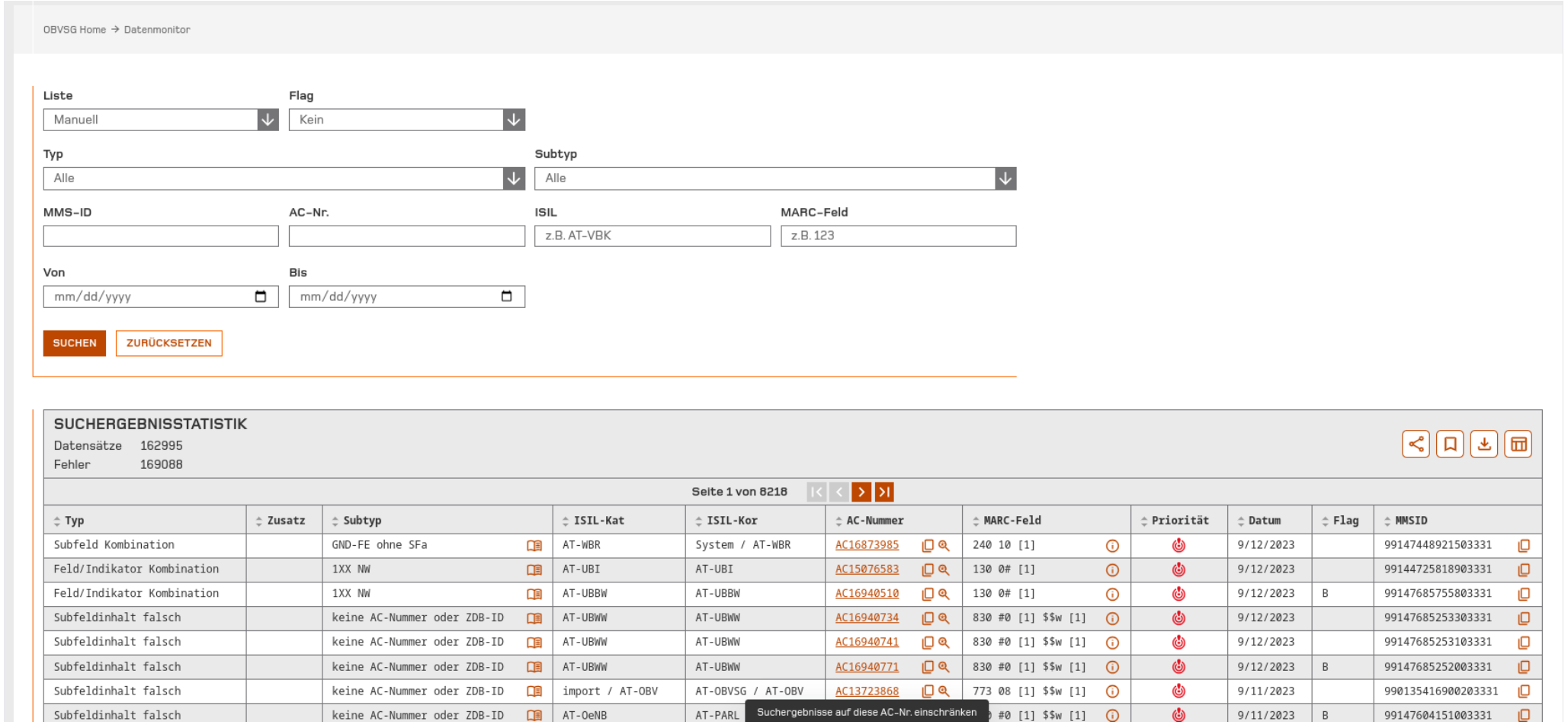
- Immediate update on bibliographic updates due to user-interaction
- NOT: any possible update; therefore, updates are missing
- NOT: the full set of available data

- Mitigation:
  - Add a second channel
  - Fill the initial set of data with from a different source
  - Resolve ambiguities between the different sources

REQUIRED COMPONENTS

COMPONENTS FOR THE BASIC TASKS

DELIVERING THE IDENTIFIED PROBLEMS TO THE USERS

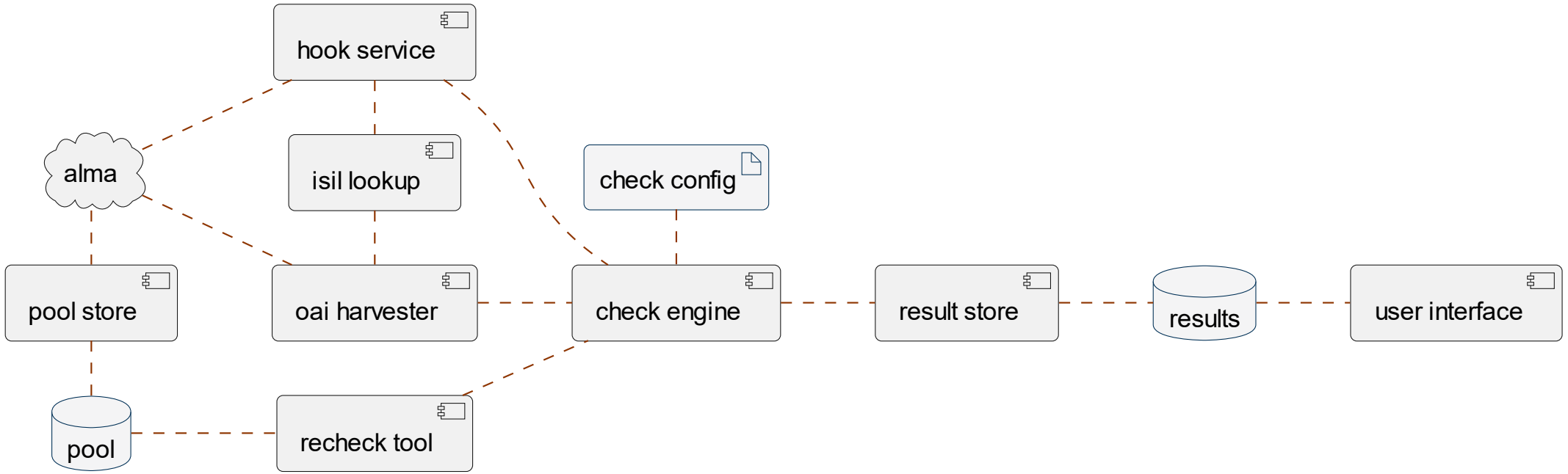


COMPONENTS FOR THE BASIC TASKS

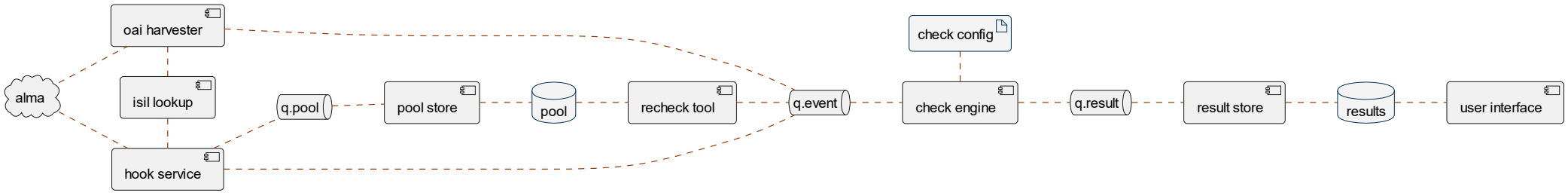
HANDLING RECORDS, IDENTIFYING ERRORS, MAKE ERRORS SEARCHABLE AND IDENTIFIABLE

COMPONENTS FOR THE BASIC TASKS

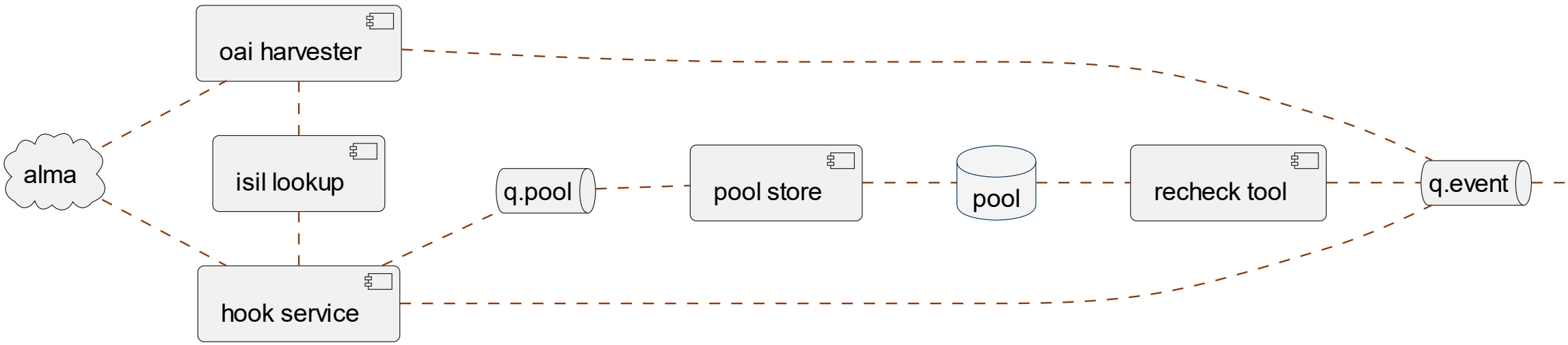
HANDLING RECORDS, IDENTIFYING ERRORS, MAKE ERRORS SEARCHABLE AND IDENTIFIABLE



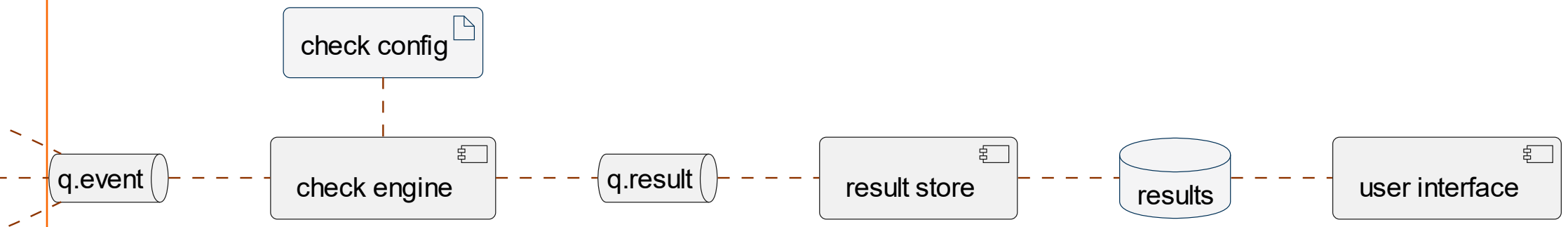
COMPONENTS TO CONNECT & FILL GAPS



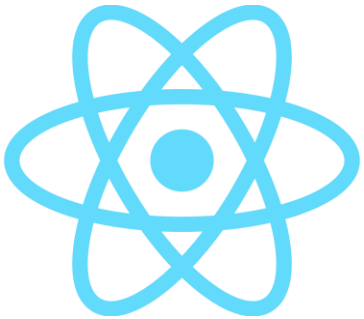
COMPONENTS TO CONNECT & FILL GAPS

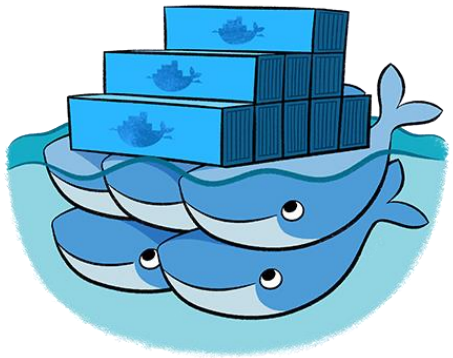


COMPONENTS TO CONNECT & FILL GAPS



OUR STACK





ExLibris Alma

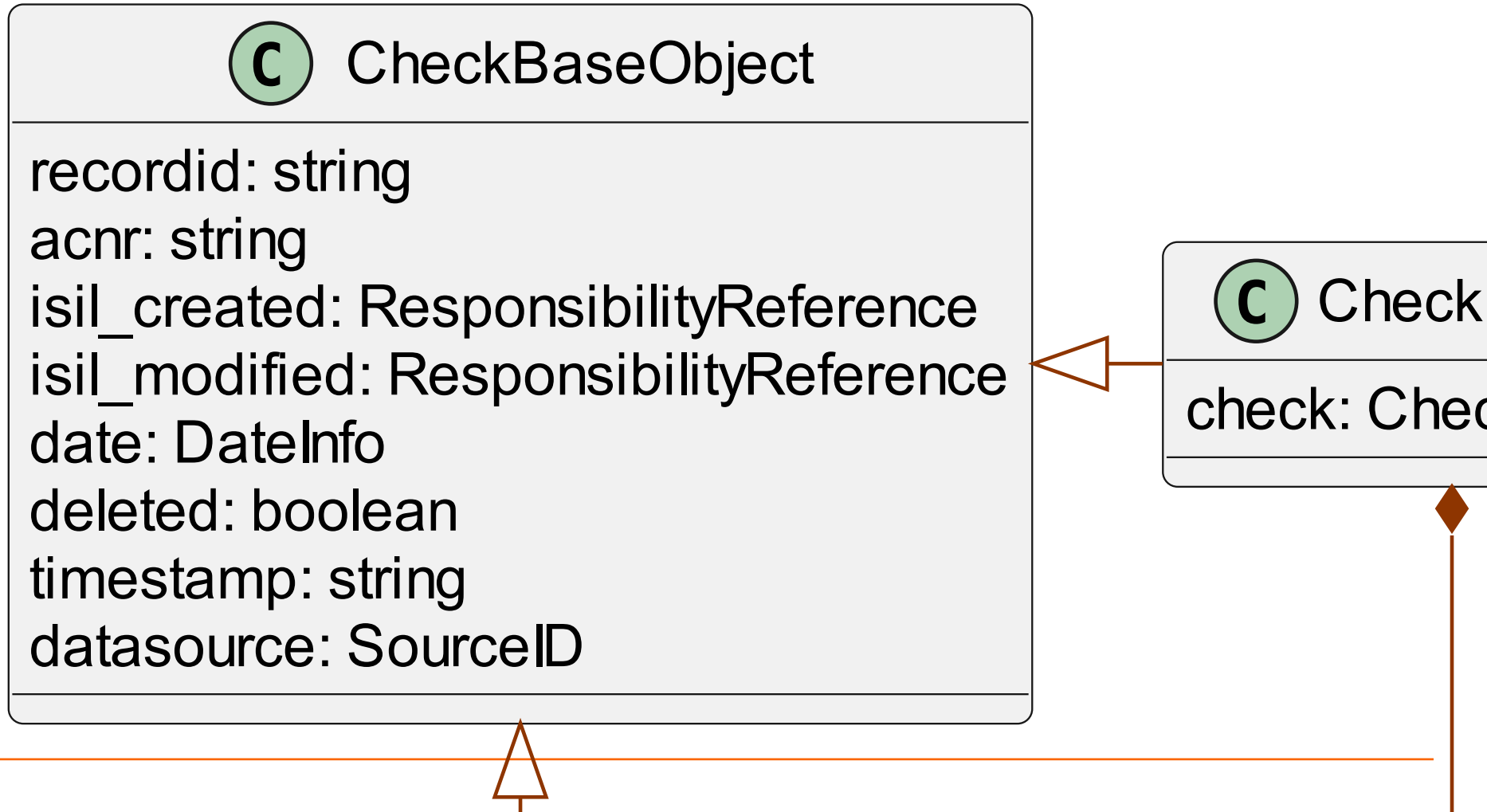
RabbitMQ

mongoDB.

ABSTRACTION CHECKING RECORDS

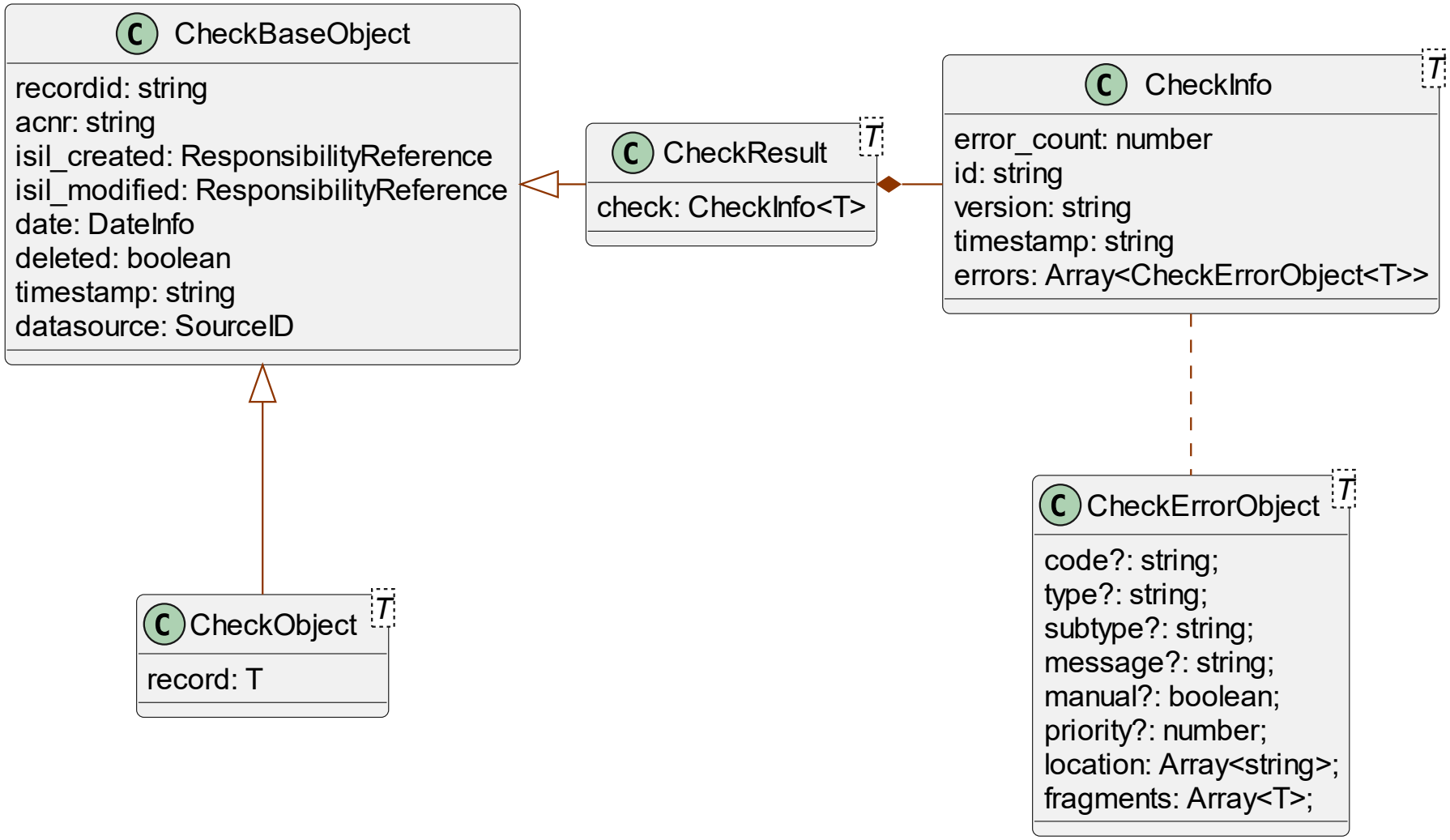
BASE CLASSES FOR CHECKING AND REPORTING ERRORS

MESSAGES, INPUT, OUTPUT

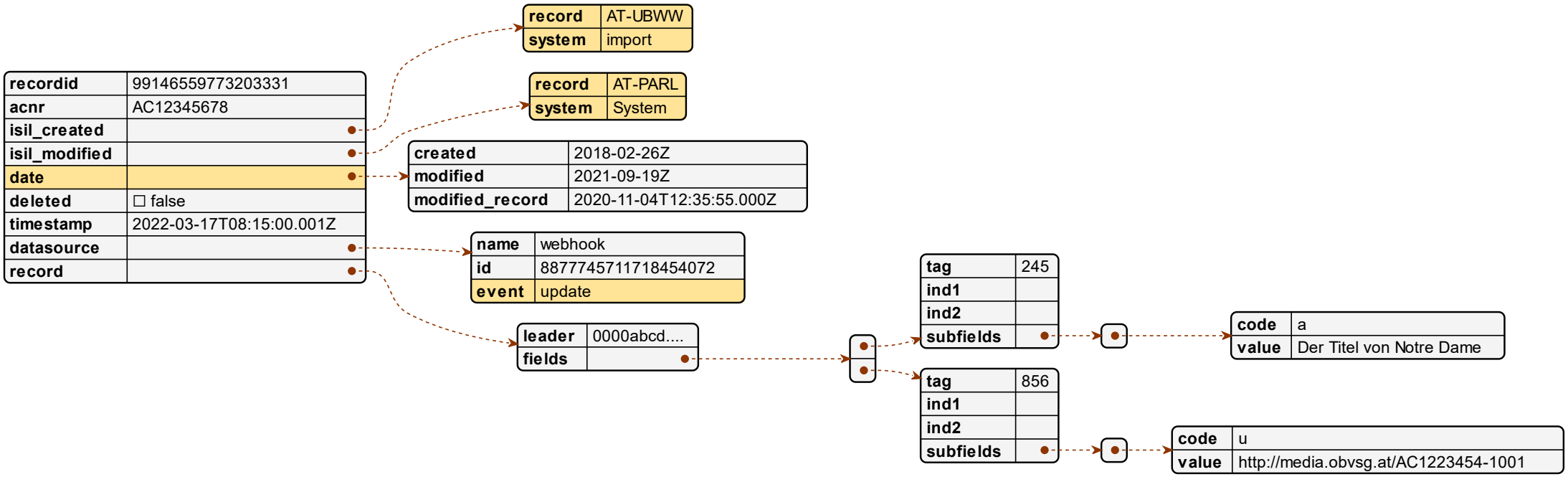


BASE CLASSES FOR CHECKING AND REPORTING ERRORS

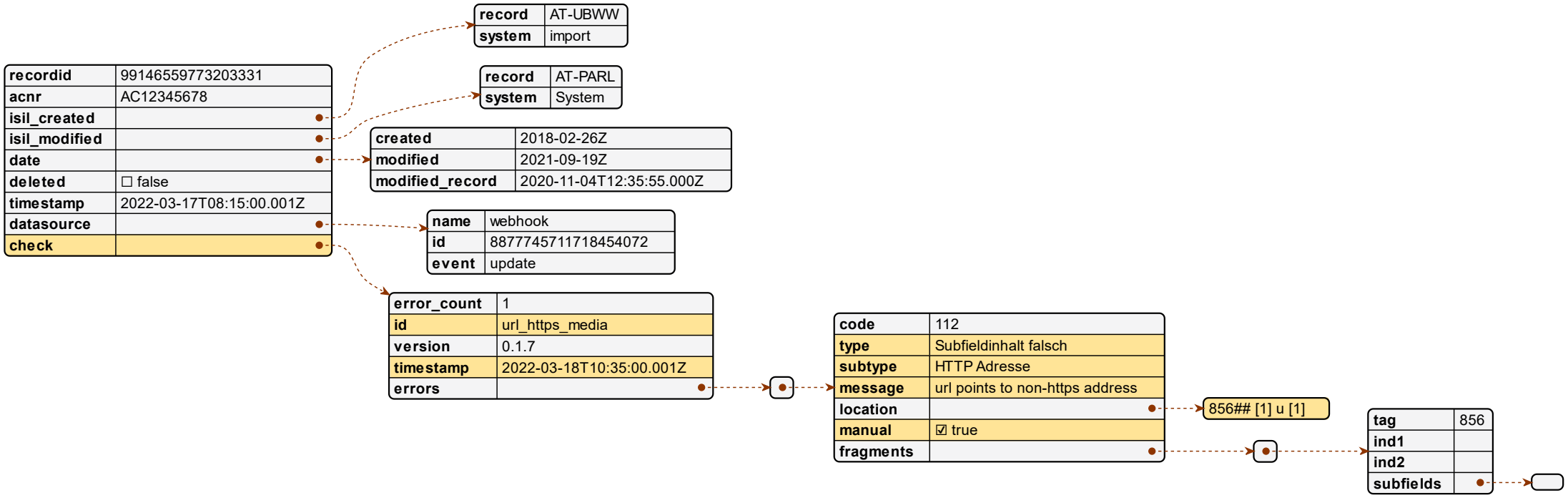
MESSAGES, INPUT, OUTPUT



CHECKOBJECT WITH DATA

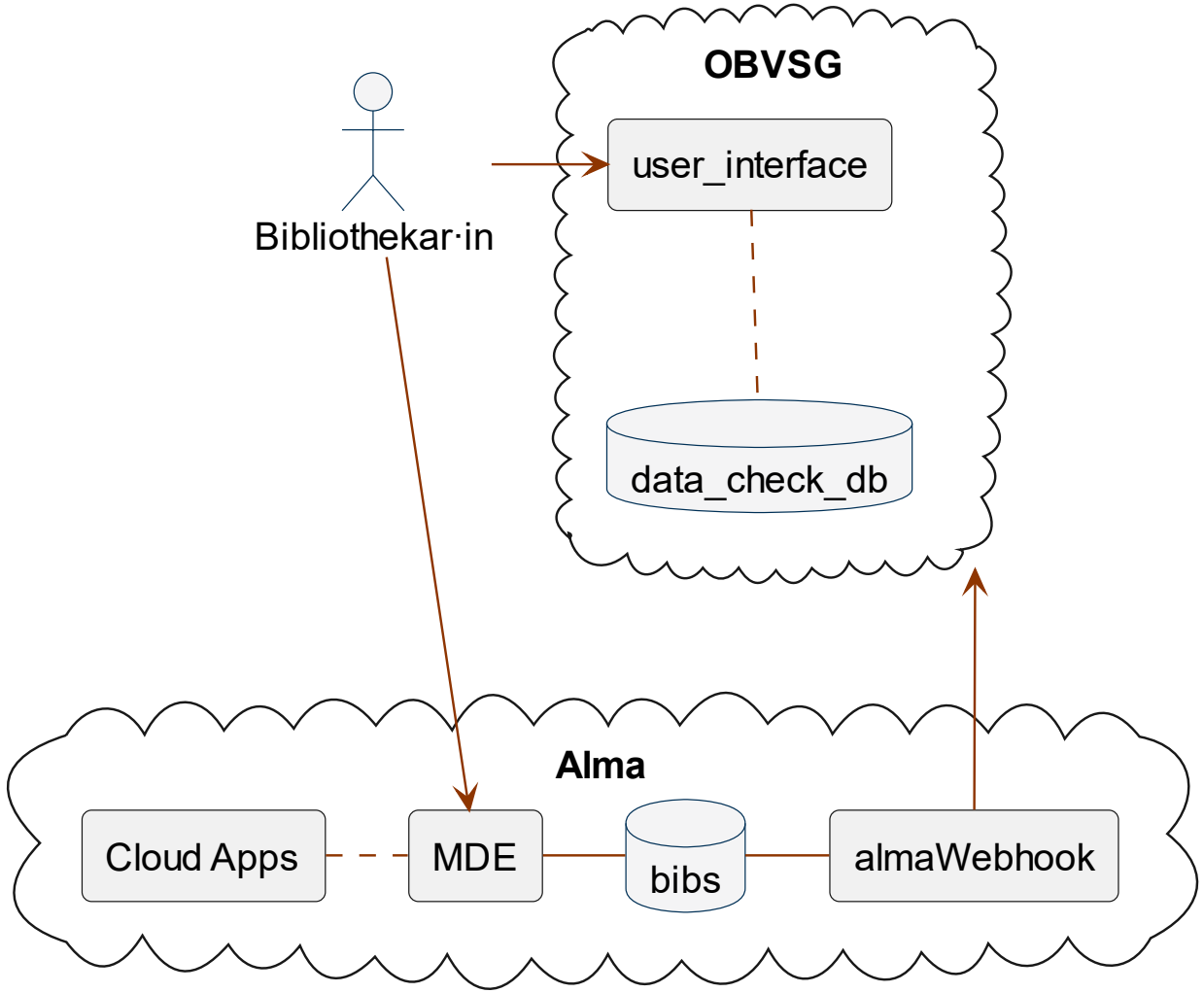


CHECKRESULT



THE TRAWLER IN ACTION

THE USERS' PLACE IN THE USE-CASE



LIVE EXAMPLE

- We want to look for an error and do something with it

Typ	Zusatz	Subtyp	ISIL-Kat	ISIL-Kor	AC-Nummer
Subfeld Kombination		GND-SE ohne SFa	AT-UBL	AT-UBL	AC16799123
Feld/Indikator Kombination		1XX NW	import / AT-OeNB	AT-OeNB	AC13961363
Subfeldinhalt falsch		keine AC-Nummer oder ZDB-ID	import / FIBA	AT-UBI	AC00514880
Subfeldinhalt falsch		keine AC-Nummer oder ZDB-ID	import / AT-UBI	AT-UBI	AC07652275
Subfeldinhalt falsch		keine AC-Nummer oder ZDB-ID	import / FIBA	AT-UBI	AC00518330
Subfeldinhalt falsch		keine AC-Nummer oder ZDB-ID	import / FIBA	AT-UBI	AC00518158
Subfeldinhalt falsch		AC-Nummer nicht gültig	import / AT-UBW	AT-UBS	AC11923269
Subfeldinhalt falsch		keine AC-Nummer oder ZDB-ID	import / AT-UBW	AT-UBS	AC11923269
Subfeldinhalt falsch		keine AC-Nummer oder ZDB-ID	import / FIBA	AT-UBI	AC00236556
Subfeldinhalt falsch		keine AC-Nummer oder ZDB-ID	import / FIBA	AT-UBI	AC00508963
Subfeldinhalt falsch		keine AC-Nummer oder ZDB-ID	import / FIBA	AT-UBI	AC00513726
Subfeldinhalt falsch		keine AC-Nummer oder ZDB-ID	import / FIBA	AT-UBI	AC00512751
Subfeldinhalt falsch		keine AC-Nummer oder ZDB-ID	import / FIBA	AT-UBI	AC00063838
Feld/Indikator Kombination		1XX NW	AT-ABPU	AT-ABPU	AC16941532
Subfeldinhalt falsch		AC-Nummer nicht gültig	import / AT-UBS	AT-UBI	AC08427024
Subfeld Kombination		GND-FE ohne SFa	import / AT-UBMW	AT-UBMW	AC09352078
Subfeld Kombination		GND-FE ohne SFa	import / AT-UBMW	AT-UBMW	AC09352090
Subfeld Kombination		GND-FE ohne SFa	import / AT-UBMW	AT-UBMW	AC09352106
Subfeldinhalt falsch		AC-Nummer nicht gültig	AT-OeNB	AT-OeNB	AC16941809
Subfeldinhalt falsch		keine AC-Nummer oder ZDB-ID	AT-UBG	AT-UBG	AC16941827

```
alma-webhook-event    |   },
alma-webhook-event    |   deleted: false,
alma-webhook-event    |   timestamp: '2023-09-13T09:14:50.588Z',
alma-webhook-event    |   datasource: { name: 'webhook', id: '1909963824663047841
alma-webhook-event    |   record: MarcRecord {
alma-webhook-event    |     _validationOptions: {},
alma-webhook-event    |     leader: '06672nam a2200397 c 4500',
alma-webhook-event    |     fields: [
alma-webhook-event    |       [Object], [Object], [Object],
alma-webhook-event    |       [Object], [Object], [Object],
alma-webhook-event    |       [Object], [Object], [Object],
alma-webhook-event    |       [Object], [Object], [Object],
alma-webhook-event    |       [Object], [Object], [Object],
alma-webhook-event    |       [Object], [Object], [Object],
alma-webhook-event    |       [Object], [Object], [Object],
alma-webhook-event    |       [Object], [Object], [Object],
alma-webhook-event    |       [Object], [Object], [Object],
alma-webhook-event    |       [Object]
alma-webhook-event    |     ]
alma-webhook-event    |   }
alma-webhook-event    | }
record-check-engine   | {"level":"info","message":"_consumeFromQueue recordid: 99
alma-webhook-event    | {"name":"Alma Webhook Event Service","hostname":"3b400ebc
rocessed: 1909963824663047841"},"msg":"","time":"2023-09-13T09:14:50.621Z","v":0
```

STRENGTHS, WEAKNESSES & OUTLOOK

WEAKNESSES

- Debugging can be challenging if things go wrong
- The overall architecture needs to be very carefully crafted; changes to fundamental message properties require a very careful approach
- Alma does not reliably create events for every change
- Data normalization jobs do not trigger updates (but may fix or introduce errors)
- No possibility to queue sets of data into alma webhooks
- Different sources of data are aggregated from different sources. Unfortunately, the time of change is not 100% safe to be determined.
- Inconsistencies in data from two sources will happen.

STRENGTHS

- Event-driven approach pays off, swift feedback leads to user-activity
- Developers can focus on smaller concerns
- Alma-webhooks are coming in close to real-time
- Webhook retries in case of service interruptions work well
- Messaging-based architecture separates concerns cleanly
- Individual services are rather simple to replace

OUTLOOK

- Adding more data checks
- Adding other types of errors/data to the monitored data source
- Bothering ExLibris to add an option to receive all changes to data as a notification on the bib webhook
- Bringing the service to regular production
- Exploring on using an Alma CloudApp to report check results

THANK YOU

Questions?