AIQA

Ready your data

Why did we create this tool?
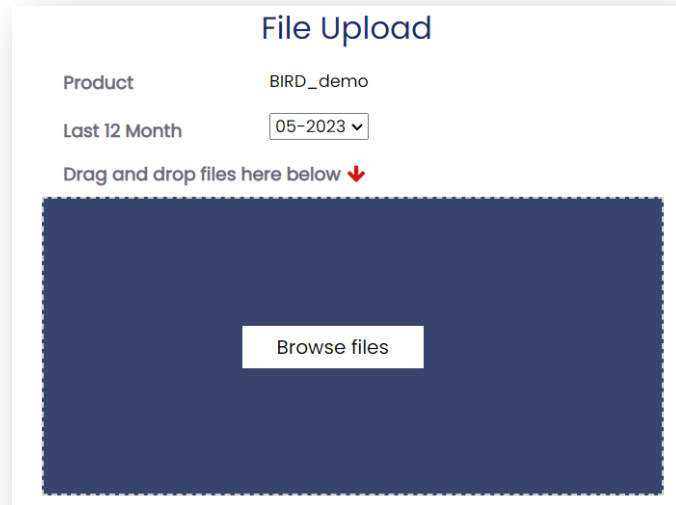
- We built this tool to make your data conform with the BIRD data model and improve overall data quality

- For this purpose, we already gathered the available information and integrated the BIRD LDM into the machine for use on actual BIRD related data

- In order to demonstrate this, we have created a sample dataset to show the different functionalities of the machine

- Additional information or an on-site presentation is possible on request

Sample dataset (redacted format)

AIQA

|	A	B	G	I	J
1	Customer_ID	BIRD_Legal_entity_identifier	BIRD_Type_of_interest_rate	BIRD_Inception_date	BIRD_Settlement_date
2	17955461	549300S4KLFTLO7GSQ80#	7	01/01/2023	01/03/2023
3	46492773	335800A16O7QD1TDQO62%	10	02/03/2022	07/10/2023
4	36670399	HWUPKR0MPOU8FGXBT394	2	12/12/2022	07/10/2022
5	39577943	54930043XZGB27CTOV49	1	01/07/2022	01/02/2022
6	13849117	787RXPR0UX0O0XUXPZ81?	2	01/02/2022	01/01/2010
7					

- The first errors we can spot are in the LEI, there cannot be any special characters

- Next would be the type of interest rate, since the value of this attribute is defined to range from 1-3

- We also created a rule that says that the inception date always has to come before the settlement date

22.04.24 Deployment Solutions for AiQA

Upload and cleaning of the testfile
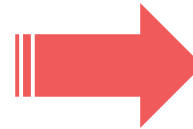
- The machine will now analyze the file based on the BIRD LDM

- We are making sure that the data is suitable for upload and will not crash the system, for example due to duplicate primary keys or deviations of the defined data format

- There is also a Regex Validation with which we are making sure that all the fields comply with their intended purpose and no invisible characters, like invisible leading spaces, are found

- After that comes the rule validation. This step is executing a variety of complex rules to guarantee integrity and completeness among others
  - Also the design is made in such a way that the process time does not increase linearly in relation to the amount of uploaded datasets
  - So that, when uploading 200000 datasets it will not take twice as long as 100000
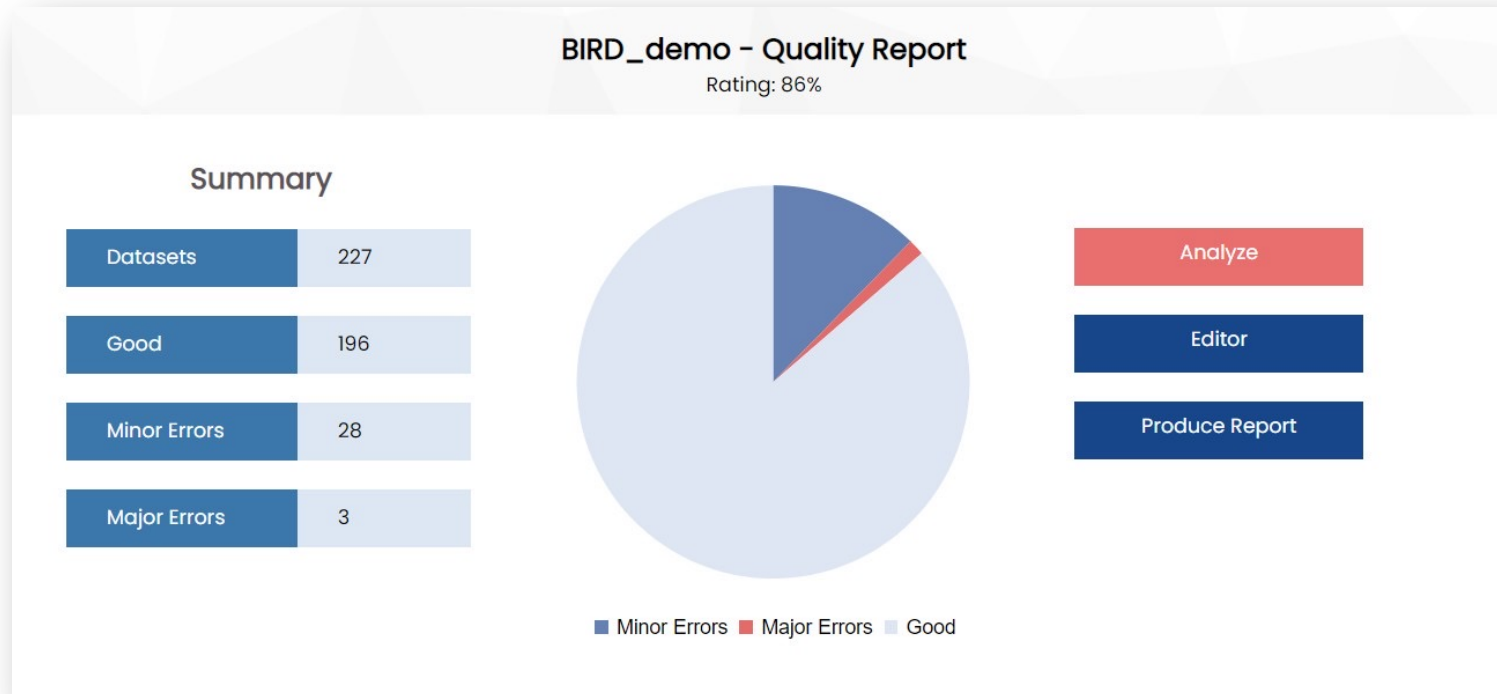
Overview tab

- Here we have a high-level Overview showing the amount of datasets and the overall rating and health of the data

Analysis tab
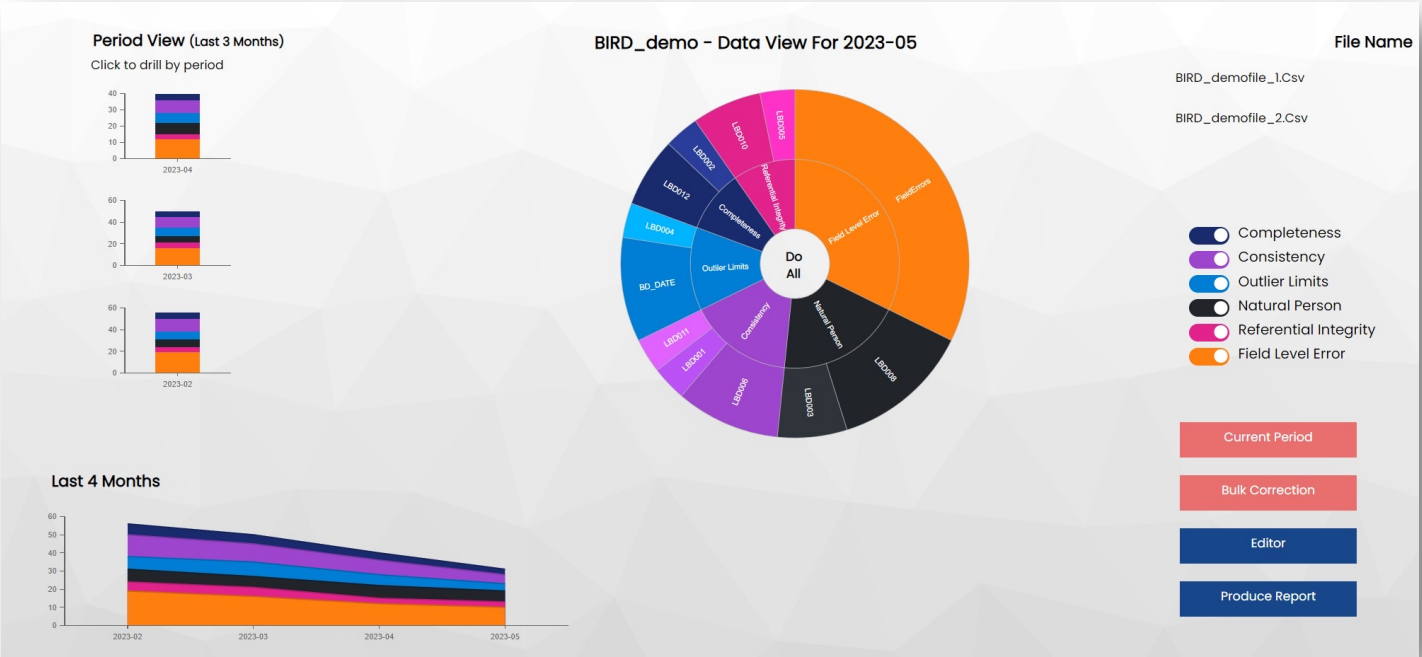
- Here we can see the structural composition of the issues in form of a pie chart in the middle

- We can also follow the data quality development over the last four months on the lefthand side

Correction-tab

- The first error we can see is the special character in the LEI

- The next error is the rule violation, the inception date has to be less than the settlement date

- We can also change the wrong value for the interest rate, this can easily be corrected via clicking on the corresponding data field, giving you all the possible values, the field can take as well as their corresponding descriptions (see snippet on the right)

- When clicking on accept All the machine will analyze every correction that was made. It has to ascertain that the corrected values do not go against the set constraints, as to make sure that it cannot accept wrong data

LEI Mapping For 2023-05

Search: _____

Entity/Instr ID	Name	LEI ID	LEI Name	Corrected
13849117	NIKE, INC.	636700IQ7G6F5LCDSN70	ANNIKEN GJELSETH HOLDING AS	✕
17955461	NVIDIA Corporation			✕
36670399	APPLE INC.			✕
39577943	TESLA, INC.			✕
46492773	SAMSUNG ELECTRONICS COMPANY LIMITED			✕

Showing 1 to 5 of 5 entries

Previous 1 Next

Manual correction for field

Search _____ 🔍

LEI Name	LEI Id	LEI Details
SASU INVIDIA INVEST	9695007IHKM3MXE3EY58	1 RUE LORD BYRON, PARIS 8, FR, glief
NVIDIA GRAPHICS PRIVATE LIMITED	335800TGZ6N1BWAN7Q35	BAGMANE GOLDSTONE BUILDING, BANGALORE, IN, glief
INVIDIA-PRIVATSTIFTUNG	529900FBLI7TJBJPAZ37	Wuhrstrasse 6, Vaduz, LI, glief
PURAENVIDIA FILMS SL	529900ZONZ3QJFXD5V94	CL SANTA ANA 3, MARBELLA, ES, glief
Nvidia Land Development, LLC	54930030GWBYMACGJH60	C/O Corporation Service Company, Wilmington, US, glief

Showing 1 to 5 of 11 entries

Previous 1 2 3 Next

- The machine also integrates Information from external databases, in this case we have a direct connection to the GLEIF database

- When clicking on a field, the machine will collect the suggested data from the 2.2 million available entries in the database. This finds the corresponding data needed for every LEI, for example the address or organization and recommends appropriate entries for further processing

The library tab

Edit Field Type Library

Field Type Name * BIRD_demo   Deactivate

Regular Expression * BIRD_all_possible_dates

Regex Description

Data Type String   Category Consistency

Is Lookup

Min Value Enter Min   Default Value test

Max Value Enter Max   Error Category Minor

Is Ignore

Description Field Type Description

- The functionality is not just limited to finding and correcting errors but also allows for transformations through a customization-function that defines datatypes as well as default values and lets you transform the input into the given standard

Reloading and AI integration

When loading the data again, the machine will do the following steps:

- In the first stage it will rely on already established data models

- In the second stage it will integrate external datasources (such as the GLEIF database)

- And in the third and final phase it will apply an AI based on machine learning, which can remember past corrections which have passed all validation checks

Correction after reload

BIRD_demo Bulk Correction For 2023-05

Entity/Instr ID	Validation Code	Field	Value	Error Type	Corrected
17955461	FieldErrors	bird_legal_entity_identifier	549300S4KLFTLO7GSQ80#	Regex Pattern Mismatch	✗
13849117	FieldErrors	bird_legal_entity_identifier	787RXPR0UX0O0XUXPZ81? → 636700IQ7G6F5LCDSN70	Regex Pattern Mismatch	✓
46492773	FieldErrors	bird_legal_entity_identifier	335800A16O7QD1TDQO62%	Regex Pattern Mismatch	✗
17955461	FieldErrors	bird_type_of_interest_rate	7 → 3	Lookup value not found	✓
46492773	FieldErrors	bird_type_of_interest_rate	10	Regex Pattern Mismatch and Lookup value not found	✗
54930043XZGB27CTOV49	BD_DATE	bird_inception_date	01/07/2022 → 10/10/2001	Rule Validation Failure	✓
787RXPR0UX0O0XUXPZ81	BD_DATE	bird_inception_date	01/02/2022	Rule Validation Failure	✗
HWUPKR0MPOU8FGXBT394	BD_DATE	bird_inception_date	12/12/2022	Rule Validation Failure	✗
984500CDAD7ABDA80D33	LBD012	bird_linked_enterprise_party_identifier	8417512	Rule Validation Failure	✗
967600HFR8T06MNBMC57	LBD012	bird_linked_enterprise_party_identifier	1608651	Rule Validation Failure	✗

Showing 1 to 10 of 31 entries

Previous 1 2 3 4 Next

Edit | Accept | Next | Accept All | Produce Report | Corrected Data

- Now you can see that the machine learned from the earlier input and makes suggestions based on what it learned

  ➢ You can now either accept the suggestions or manually make new corrections

API integration

Product List

Search: [   ]

Edit	Product	Description	Domain	Is LEI	Status	Api
✎	BIRD_demo	Product to show different functionalities of...	BIRD_demo_1, BIRD_demo_2	No	Active	
✎	DZ_API_ANNACREDIT	AP Product	ENTITY_API	No	Active	⬇
✎	DZ_LEI	LEI Dataset	LEI_Entity	Yes	Active	
✎	DZ_ANNACREDIT	Annacredit Dataset	Entity	No	Active	

- The system can also be integrated into your data flow via an API:

  - For this each product generates its own API which can interact with your ETL
  - This can also be adapted to various systems
  - The same works when feeding the corrected data back into the source systems (Correction   Corrected data)

Reports

I	J	N	O
Old Value	Suggested Value	Change User	Change Timestamp
9978211			
Jack			
01/02/2022			
US6541061031			
Lee			
787RXPR0UX0O0XUXPZ81?	549300JPLJF5XUZV6M76	Lucas.Christmann@de.ey.com	10/05/2023 15:15
NIKE, INC.			
22			
01/01/2010			
30000			
2			
13849117			
4033622			
Huang			
01/01/2023			
US67066G1040ï¿½			
Jen-Hsun			
549300S4KLFTLO7GSQ80#	549300S4KLFTLO7GSQ80	Lucas.Christmann@de.ey.com	10/05/2023 15:16
NVIDIA Corporation			
8			
01/03/2023			
1000000			
7	3	Lucas.Christmann@de.ey.com	10/05/2023 15:16
17955461			
02/03/2022			

- You will receive a detailed report about all the corrections carried out, the time and about who made the correction

- An audit proof protocol in form of a pdf will also be generated, documenting all changes made by the system or individual users including timestamps