



Creating
DuoFactory
*An Orchestration
Ecosystem with
Airflow*
Summit 2025

3.0



Belle Romea,
Software Engineer
@ Duolingo



Duolingo ✅
@duolingo

Remember to not miss any lessons today! It wouldn't be very nice if we had to lock up your family.

12:07 PM · 30 Mar 19 · Twitter Web Client

2,578 Retweets 5,869 Likes



FBI ✅ @FBI · 47m

Replying to @duolingo

We are working on tracking you down. It is not okay to kidnap families because someone didn't complete a lesson.



Duolingo is a popular, free language-learning app that uses a gamified, bite-sized lesson approach to teach reading, writing, listening, and speaking skills

Write this in English



¿Cuáles son tus libros favoritos?

park are good your What favorite
books to learn bedroom



our mission is to
develop the best
education in the
world and make it
universally
available





... but how do we do that?





AIRFLOW
SUMMIT

meet the
duolingo
curriculums



Courses

New courses



Chess



Math



Music

For English speakers



Spanish



French



German



Italian



Japanese



CONTINUE



AIRFLOW
SUMMIT

6

0

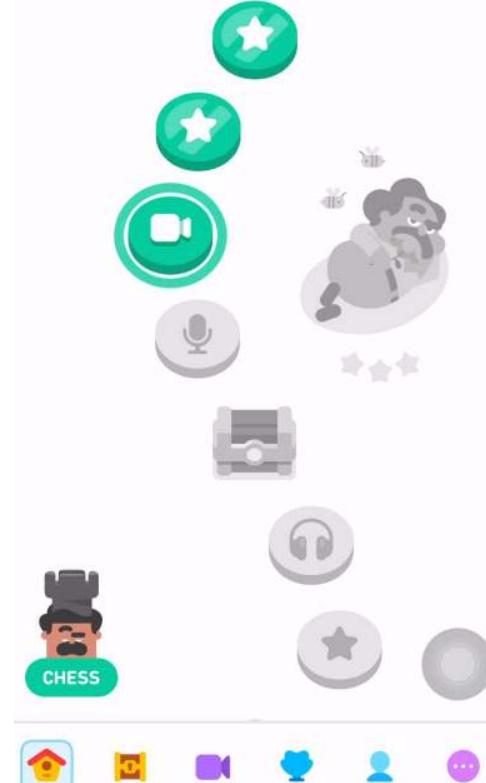
1208



SECTION 1, UNIT 3
Use adjectives



each course
needs a lot
of content





just to create the new sentences
it take around 600 hours per
course!

Write this in English

A cartoon character with blue hair and a pink sweater is shown next to a speech bubble containing the Spanish sentence "¿Cuáles son tus libros favoritos?". Below the character are two sets of horizontal lines for writing practice. At the bottom, there is a grid of words in boxes: park, are, good, your, What, favorite, books, to learn, and bedroom.





and we want rich content





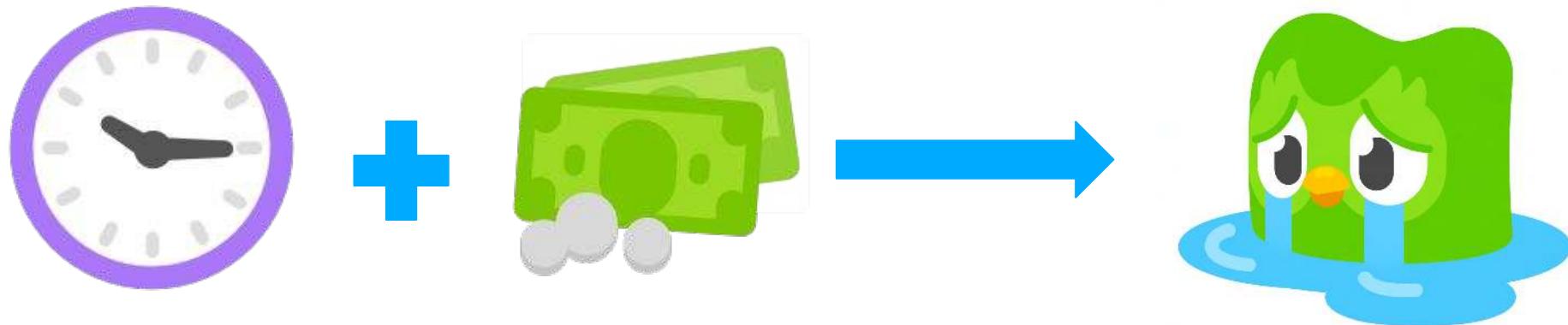
A screenshot of a Duolingo mobile application. At the top, there's a green progress bar with a red heart icon containing the number '5'. Below it is a cartoon character holding a map. The main content area shows a conversation between two characters: 'Vikram' and 'Béa'. Vikram says, 'Voyage en Espagne'. Béa replies, 'Salut, Béa ! Comment ça va ?'. Below that, Béa says, 'Super ! Je suis très contente d'aller en Espagne !'.

A screenshot of a podcast page for the 'Duolingo Spanish Podcast'. The title 'Duolingo Spanish Podcast' is prominently displayed with the Duolingo owl logo. Below the title, there's a 'Follow' button and a '... more' button. The 'All Episodes' section shows two episodes: 'Duo's Film Club - Fresa y chocolate' (published Feb 29, 18 min 27 sec) and 'Duo's Film Club - Coco' (published Feb 22, 17 min 46 sec). Each episode has a thumbnail, a play button, and a description. To the right, there's an 'About' section with a brief description of the podcast and its host, Martina Castro. A 'Trailer Season 3' is also visible at the bottom right.

stories podcasts



but this is timely + expensive





the result



limited courses

A screenshot of a forum thread from a platform like Reddit or Stack Overflow. The thread is titled "2014. It comes up in every single discussion about this topic, and Duolingo already did every other major language in the region except Persian." The thread has received 70 upvotes and 4 more replies.

guitarelf - 2y ago
Wait as a Persian I'm glad y'all wanna learn our language it's an honor!
12 upvotes, 1 reply

icameisawicame24 - 2y ago
Spanish
Would love to learn Persian.
12 upvotes, 2 more replies

Curry_pan - 2y ago
Native: Chinese A2: English
I've been waiting for Persian for so long! We can only hope.
10 upvotes, 1 reply

Toothless-Rodent - 2y ago
Persian is my personal top choice.
12 upvotes, 6 more replies

Hoita - 2y ago
Native: English
Māori.
Rips that it started in the incubator and went nowhere.
136 upvotes, 1 reply

martinhe333 - 2y ago
It would be great to have Te Reo Māori. I read that there were pushback, as that effort was unsupervised.
7 upvotes, 1 reply



limited content per course

A screenshot of a forum thread from the Airflow Summit. The thread discusses the lack of stories in various languages, particularly Haitian Creole, Russian, and Hebrew. The posts are as follows:

- DantheCat7** · 2y ago: I wish every language has stories, hearing stories in Haitian creole, Russian, and Hebrew would be so awesome.
65 upvotes, 14 downvotes, 1 award, 1 share.
- Garfunkeled1920** · 2y ago: Agree. I think it's a major deficiency, but I cannot imagine those languages have enough users (or could draw enough users) to justify the creation cost. Duo is, after all, primarily a Spanish-language-learning app that also dabbles in other languages.
14 upvotes, 14 downvotes, 1 award, 1 share.
- Dm_Me_TwistedFateR34** · 2y ago: Spanish stories send me to the moon sometimes, they're so absurd and hilarious at times.
30 upvotes, 14 downvotes, 1 award, 1 share.
- markisnotcake** · 2y ago: There are no stories for any nordic language. :c
21 upvotes, 14 downvotes, 1 award, 1 share.
- Wolke_01** · 2y ago: (N) (C1) (A2)



but then the quality of
LLM generation
improved...

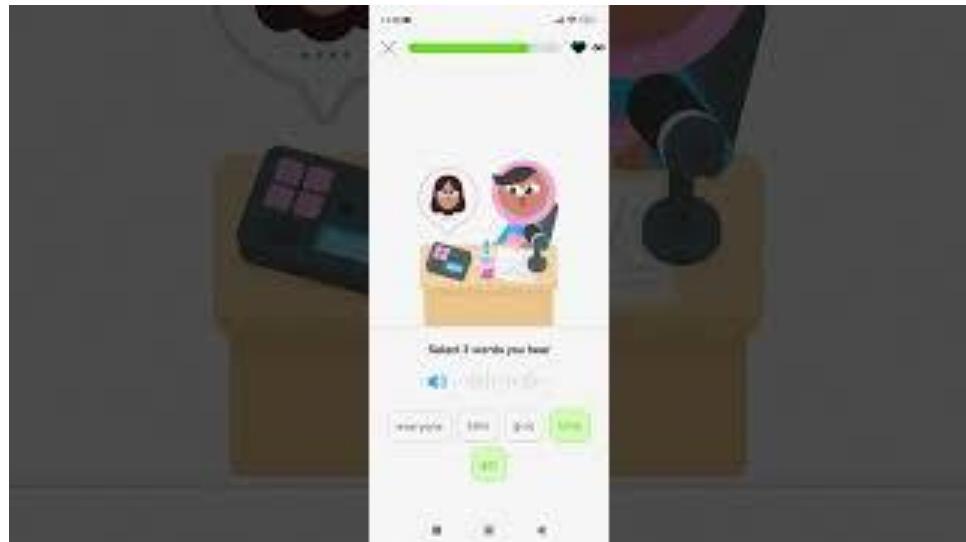
so much that LLMs
can generate our
content to speed us
up



example



a Duoradio
takes 6 hours
with LLMs+
human review





okay ... problem solved





hold your horses



limitations to genai scalability

engineering costs increasing

as we implemented new ai approaches,
we had fragmented tooling landscape

- Bespoke one-off endpoints on existing services → technical debt
- No easy way to share and reuse logic across workflows → reinventing the wheel
- Homegrown orchestration was non-trivial to scale





enter Airflow



DuoFactory

The screenshot shows the Airflow web interface with the following details:

- Header:** DAG, Cluster Activity, Databases, Security, Envvars, Admin, Docs, Other.
- Breadcrumbs:** Looking for DAGs > Start here.
- DAG List:** A list of DAGs with their names, descriptions, and execution dates.

DAG	Description	Last Run	Next Run	Reschedule
airflow_dags	general dags	2023-10-20 17:29:20	2023-10-21 16:50:00	<input type="button" value="Reschedule"/>
add_ch_exercise_tags	insert policy relation	2023-10-20 17:29:20	2023-10-21 16:50:00	<input type="button" value="Reschedule"/>
add_topic_can_be_granted_to_scope_and_sequence	insert policy relation	2023-10-20 17:29:20	2023-10-21 16:50:00	<input type="button" value="Reschedule"/>
ai_art_pipeline_for_word_object	insert policy relation	2023-10-20 17:29:20	2023-10-21 16:50:00	<input type="button" value="Reschedule"/>
airflow-dfip-log-analyzer	insert policy relation	2023-10-20 17:29:20	2023-10-21 16:50:00	<input type="button" value="Reschedule"/>
airflow-dfip-create-database	insert policy relation	2023-10-20 17:29:20	2023-10-21 16:50:00	<input type="button" value="Reschedule"/>
apply_manual_review_updates	insert policy relation	2023-10-20 17:29:20	2023-10-21 16:50:00	<input type="button" value="Reschedule"/>
color_label_with_prompt_ensemble	insert policy relation	2023-10-20 17:29:20	2023-10-21 16:50:00	<input type="button" value="Reschedule"/>
boostDFP_fetched_interest_answers	insert policy relation	2023-10-20 17:29:20	2023-10-21 16:50:00	<input type="button" value="Reschedule"/>
batch_run_parallel_tasks	insert policy relation	2023-10-20 17:29:20	2023-10-21 16:50:00	<input type="button" value="Reschedule"/>
transformat_d_of_evaluators	insert policy relation	2023-10-20 17:29:20	2023-10-21 16:50:00	<input type="button" value="Reschedule"/>
- Filtering:** Filter DAGs by name (e.g., "start") and Sort DAGs by Last Run.
- Actions:** Buttons for Auto-refresh and Refresh Tasks.

wrapper around Airflow for genai orchestration



why airflow was a good
candidate



in-house orchestration is hard

Airflow provides rich orchestration tooling, including scheduling, dependency management, retries, and observability



modularity

- tasks are reusable logic
 - high ROI of eng work
- easily swap out parts of the pipeline as tech improves



examples of tasks

- read/write google sheets
- read course data
- run prompts
- write to s3 to be viewed in the app



idempotency

each step in a DuoFactory DAG
can be retried safely without
corrupting data or producing
inconsistent result

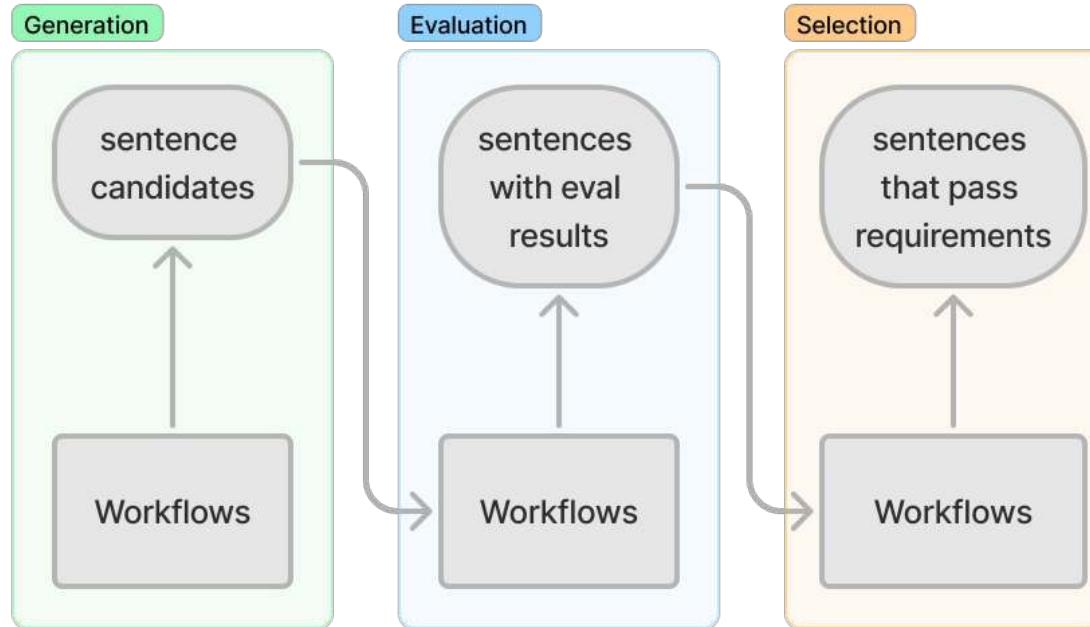


ecosystem compatibility and extensibility

Airflow integrates easily with Duolingo's existing AWS, PostgreSQL, and Terraform-based infrastructure



Duofactory in action

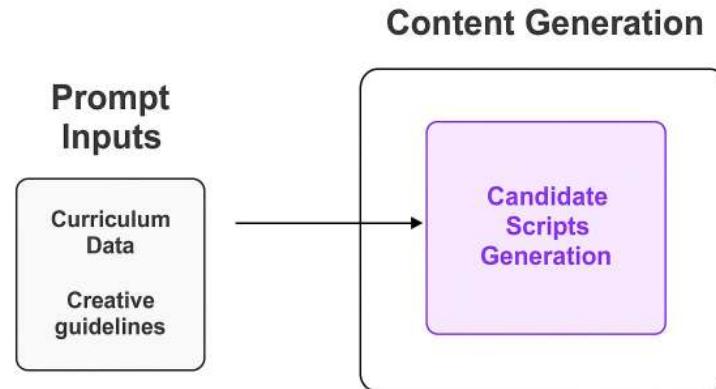




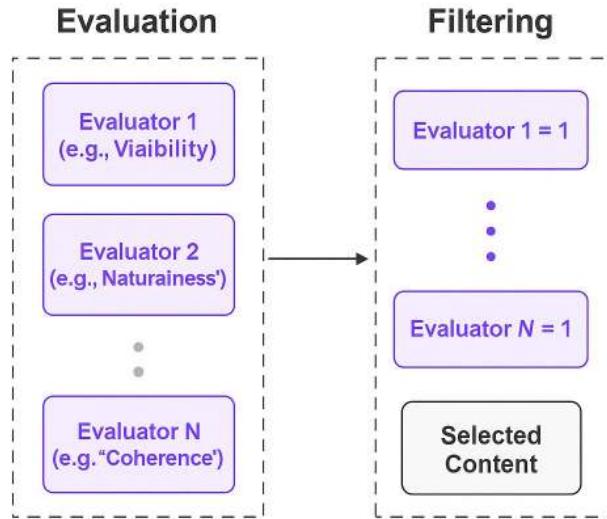
example
Duoradio generation



script generation DAGs



script evaluation DAG





then repeat for exercises
generation -> evals -> selection



extending to end to end





formatting DAG

- API calls to TTS/Viseme generation

publishing DAG



wins



created a pipeline to
generate duoradio
episodes with 0
minutes of human
intervention





cut costs by 99%





impact

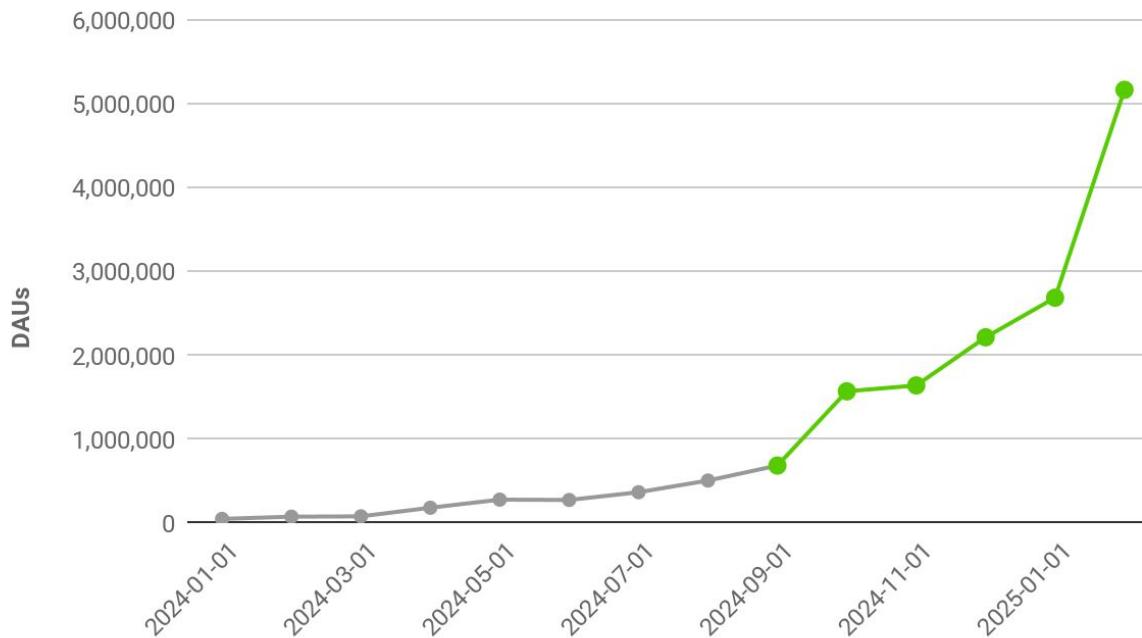


launched duoradio in 243
courses

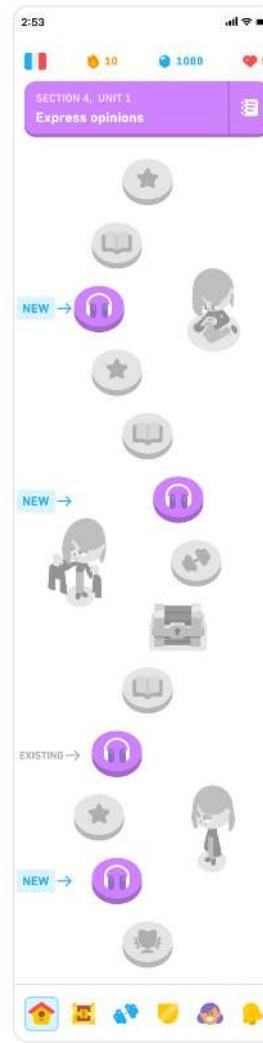
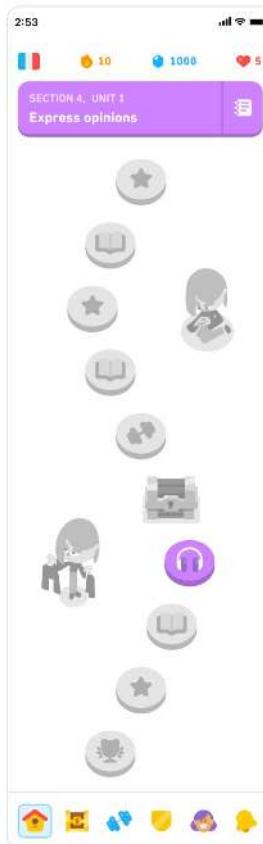




reaching more learners



more
episodes
per unit





expanding the ecosystem





prompt editor



[HOME](#) **belle test** **DRAFT** Version 1

P

Chat Template

- 1 System Message**
The prompt sent by the system...
- 2 User Message**
The prompt sent by the user...
- + Create New Step**
- 3 Assistant Message**
 Auto Generate this Message

[VIEW JSON](#)

Input Variables

No variables were found in your prompt.
In your prompt, you can define a custom variable using the `{{myVariableName}}` syntax.
For image variables, use `{{image:myImageName}}` syntax.
Image variables need to be put at the end of the message. For example, *describe the image {{image:image1}}*.

Chat

You'll see messages after running the prompt.
Use Command + Enter, or click Run Prompt.

Model Configuration

Output Variables

PUBLISH VERSION **RUN PROMPT**



google sheets





content management system



Filter: **DuoRadio / Prod** **EN < ES**

Search by episode name or ID FILTER

<input type="checkbox"/>	EPISODE NAME	VIABILITY	WRAPPER	STATUS	SLOT	SCHEMA VERSION	Pipeline Status	GLOBAL ATTEMPTS	LAST UPDATED
Empleos (A1.0 – 63f532236a038ee957b7154aaddbcd)											
<input type="checkbox"/>	Work in Chicago	0.85	Lily and Strangers	PUBLISHED	0	3	Published	2	12/15/2024 3:46 AM EST
<input type="checkbox"/>	Interesting Job	0.86	Falstaff's Guide to Humans	PUBLISHED	0	duoradio_a1_flagships	Published	1	6/20/2025 12:13 PM EDT
<input type="checkbox"/>	Are You a Doctor?	0.86	Lucy Is Listening	PUBLISHED	1	duoradio_a1_flagships	Published	1	6/20/2025 12:13 PM EDT
<input type="checkbox"/>	Artists and Offices	0.85	Lucy Is Listening	PUBLISHED	1	3	Published	1	11/20/2024 6:56 PM EST
<input type="checkbox"/>	Work at the Hospital	0.87	Late Nights with Lin	PUBLISHED	2	duoradio_a1_flagships	Published	1	6/20/2025 12:13 PM EDT
<input type="checkbox"/>	Meeting Humans	0.87	Falstaff's Guide to Humans	PUBLISHED	2	3	Published	1	11/20/2024 8:38 PM EST
Encuentros (A1.0 – 46f57661460e7b1621fed84e29bc0720)											
<input type="checkbox"/>	A Bilingual Encounter	0.94	Lucy Is Listening	PUBLISHED	1	duoradio_a1_flagships	Published	1	6/20/2025 12:13 PM EDT
<input type="checkbox"/>	Nice to Meet You	0.92	Falstaff's Guide to Humans	PUBLISHED	1	3	Published	1	11/20/2024 11:14 PM EST
<input type="checkbox"/>	Do You Speak English?	0.94	Falstaff's Guide to Humans	PUBLISHED	2	duoradio_a1_flagships	Published	1	6/20/2025 12:13 PM EDT
<input type="checkbox"/>	Suspicious Nice to Meet You	0.96	Lucy Is Listening	PUBLISHED	2	3	Published	1	11/20/2024 10:15 PM EST
<input type="checkbox"/>	Conversation in Two Languages	0.85	Falstaff's Guide to Humans	PUBLISHED	3	duoradio_a1_4_nodes_flagship	Published	2	7/14/2025 11:14 PM EDT



workflow builder
codeless DAG
using predefined
building blocks



← HOME A1 DuoRadio Title, Exercises & Formatting for Lucy IT<EN ⚡

SAVE SAVE AND RUN VIEW RESULTS

The screenshot shows a Data Pipeline (DAG) configuration in the Airflow interface. On the left, there's a sidebar with icons for Home, Import, Prompt, Evaluate, Parallel Evaluate, Sort, Select, and Custom. The main area displays a sequence of tasks connected by arrows:

- Import 1**: Import using Google Sheets.
- Prompt 1**: A1 DR episode title (UI).
Run prompt 1 time(s).
- Prompt 2**: A1 DuoRadio 3-in1 exercises: Listen+Comp+CAP.
Run prompt 2 time(s).
- Evaluate 1**: Evaluate using Loading...
Evaluate using Loading...
- Prompt 3**

Each task has a red trash can icon to its right.



each workflow block maps to a corresponding DAG based on its type and metadata



execution flow

- load workflow definition from the workflow table
- trigger the corresponding block-type DAG, passing outputs from the previous block
- use `should_execute()` to stop once workflow completion is detected



static architecture constraint

DAGs require a fixed structure known at scheduling time... but we get the definition at run time

always iterate MAX_BLOCKS times, even if some iterations are skipped



```
workflow_id = cast(WorkflowId, "{{ params.workflow_id }}")  
use_local_db = cast(bool, "{{ params.use_local_db }}")  
workflow = get_workflow(workflow_id=workflow_id, use_local_db=use_local_db)  
blocks = get_workflow_blocks(wf=workflow)
```



```
for i in range(MAX_BLOCKS):
    # For downstream blocks, pass the output of the previous block as the `payloads` parameter
    # If the previous outputs are empty, seed the first block with an empty payload
    payloads = previous_block_output or [Payload(block_results={})]

    # This returns both the first task and the actual output of the @task_group so we can set
    # the dependencies correctly
    first_task, current_block_output = run_current_block.override(
        default_args={
            "trigger_rule": TriggerRule.ALL_SUCCESS,
        },
        group_id=f"run_current_block_{i}",
    )(
        global_vars=cast(dict[str, Any], "{{ params.global_variables }}"),
        index=i,
        input_payloads=payloads,
        wf=workflow,
        wf_id=workflow_id,
        use_local_db=use_local_db,
    )

    if previous_block_output:
        # Short circuit the current block if we are at the end of the workflow
        # This will cause the downstream blocks to be skipped (until the final task group,
        # which has its own trigger rules, since we need to store the results at the end.)
        short_circuit = should_execute.override(
            ignore_downstream_trigger_rules=False,
            task_id=f"short_circuit_{i}",
            trigger_rule=TriggerRule.ONE_SUCCESS,
        )(
            curr_block=i,
            n_blocks=num_blocks,
        )
        chain(previous_block_output, short_circuit, first_task)

    previous_block_output = set_latest_output_payloads.override(
        task_id=f"set_latest_output_payloads_{i}",
    )(current_block_output, previous_block_output)
```

thank you





questions?



belle@duolingo.com
LinkedIn:
[@isabel-romea](https://www.linkedin.com/in/isabel-romea)