[PRD] Contributor Insights V2 - Cohorts & Action Automation

Problem Statement

Project Managers currently have to manually gather contributor data and derive insights to take actions such as disabling low performers, promoting high performers, or sending reminders to inactive users.

- This is a daily, highly repetitive workflow that becomes increasingly difficult as projects scale and contributor counts grow. Its a very time taking process currently.
- Without a central view of contributor groups (high-performing, low-performing, inactive), PMs must repeatedly piece together information and apply actions manually.
- This process is slow, scattered, and prone to human error, consuming significant PM time.

As a result, project management on FTS remains reactive: PMs often react late to performance issues because identifying the right group of contributors itself takes too long, which leads to slower intervention and operational inefficiencies.

Issue Detected	Action Require
Inactive or low-activity users, high AHT, excessive skips, low submissions	Send reminder / engagement email
High reviewer rejection rate, low QC pass rate, poor eval score	Disable user from workflow
High QC pass rate, consistently low rejection rate	Promote contributor to reviewer step

Without a way to easily group contributors based on metrics and thresholds, PMs stay in a **reactive mode**, making decisions based on fragmented insights rather than a structured, repeatable way.

Value on Solving

Enabling PMs to create Contributor Cohorts gives them a fast, centralised, and reliable way to segment contributors based on performance metrics (quality, throughput, engagement, AHT, etc.) and thresholds.

With cohorts:

- PMs can **immediately see filtered groups** such as “Worst Performers”, “New Users”, or “Inactive Contributors” without manually rebuilding lists daily/weekly, improving the speed and accuracy of operational decisions.
- Cohorts become a **reusable building block** across FTS: these filtered lists can plug directly into future action-based features (task limits, QC sampling configs, nudges, promotions).
- It allows us to set the foundation for **automation**, where the same cohort definitions can be used to trigger actions automatically.

The feature reduces PM effort, eliminates human error, and enables more **proactive performance management** at scale.

So, cohorts introduce a **systematic, scalable way** to group contributors, making FTS a more powerful project management platform and unlocking the next layer of automated operations.

Solutioning

What is a Cohort?

A **cohort** is a dynamically generated, filtered set of contributors based on selected performance metrics and threshold conditions.

Rather than manually piecing together contributor insights, PMs can apply multiple filters (e.g., QC Pass %, AHT, Activity, Rejections) to the existing metrics table and instantly view the contributors who meet those criteria.

Why Cohorts?

- Cohorts exist to make **targeted action** possible.

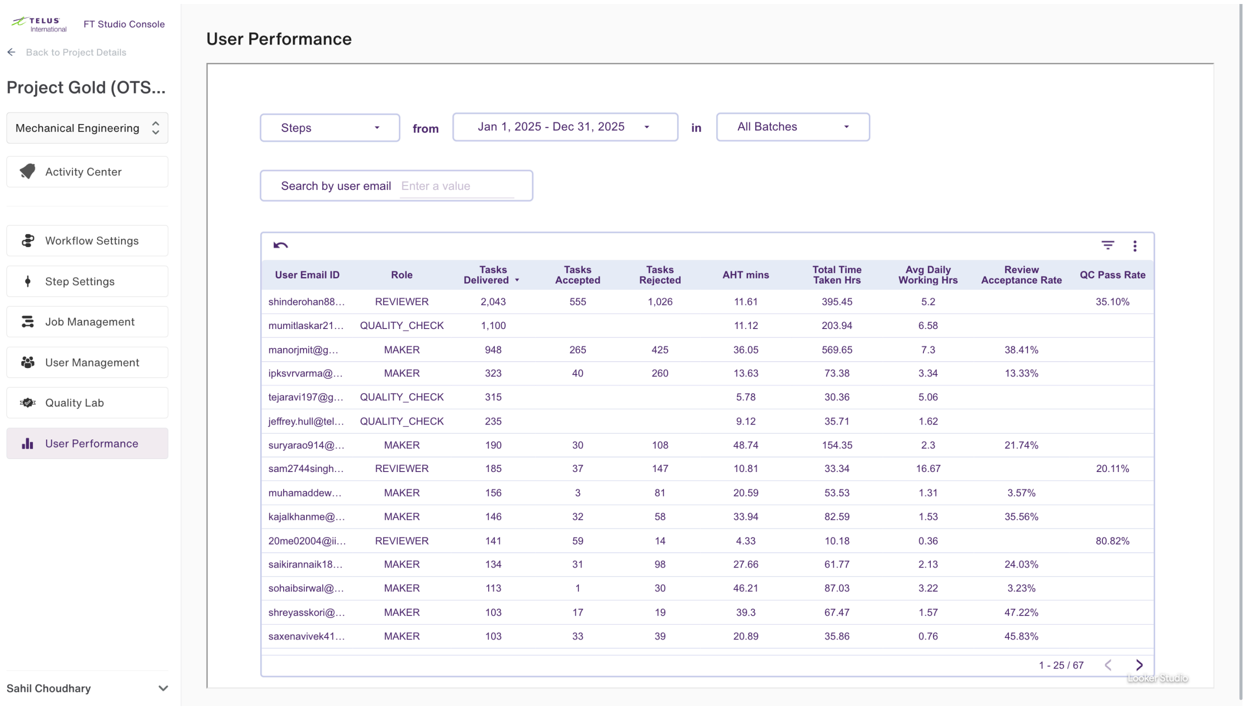
  Instead of scanning raw metrics across dozens or hundreds of contributors, PMs can quickly isolate groups such as *low performers*, *inactive users*, or *new contributors* and take action on them.

- Cohort creation is intentionally **action-driven**:

  PM decides *what they want to do* (disable, nudge, retrain, limit tasks, adjust QC sampling), and cohorts help identify *which contributors* should receive that action.

> Enable PMs to quickly group contributors into meaningful buckets so they can move from raw data → insights → corrective action in a structured, repeatable way.

Current State in the Product (What we delivered in V1)



User Performance

Steps from Jan 1, 2025 - Dec 31, 2025 in All Batches

Search by user email Enter a value

User Email ID	Role	Tasks Delivered	Tasks Accepted	Tasks Rejected	AHT mins	Total Time Taken Hrs	Avg Daily Working Hrs	Review Acceptance Rate	QC Pass Rate
shinderohan88…	REVIEWER	2,043	555	1,026	11.61	395.45	5.2		35.10%
mumitlaskar21…	QUALITY_CHECK	1,100			11.12	203.94	6.58		
manorjmit@g…	MAKER	948	265	425	36.05	569.65	7.3	38.41%	
ipksvrvarma@…	MAKER	323	40	260	13.63	73.38	3.34	13.33%	
tejaravi197@g…	QUALITY_CHECK	315			5.78	30.36	5.06		
jeffrey.hull@tel…	QUALITY_CHECK	235			9.12	35.71	1.62		
suryarao914@…	MAKER	190	30	108	48.74	154.35	2.3	21.74%	
sam2744singh…	REVIEWER	185	37	147	10.81	33.34	16.67		20.11%
muhamaddew…	MAKER	156	3	81	20.59	53.53	1.31	3.57%	
kajalkhanme@…	MAKER	146	32	58	33.94	82.59	1.53	35.56%	
20me02004@ii…	REVIEWER	141	59	14	4.33	10.18	0.36		80.82%
saikirannaik18…	MAKER	134	31	98	27.66	61.77	2.13	24.03%	
sohaibsirwal@…	MAKER	113	1	30	46.21	87.03	3.22	3.23%	
shreyasskori@…	MAKER	103	17	19	39.3	67.47	1.57	47.22%	
saxenavivek41…	MAKER	103	33	39	20.89	35.86	0.76	45.83%	

1 - 25 / 67

Today, FTS shows contributor performance via:

- A Looker Studio dashboard
- Embedded into FTS via iframe
- Read-only, dashboard-first experience

What this gives us:

- Correct metric definitions
- Alignment with BI
- Fast initial delivery

What it **does not** give us:

- Metric-level filtering inside the product
- Percentile calculation
- Step-complete coverage (Rework step is not included in the current version)
- A foundation for cohort creation or actions

This implementation was intentionally scoped as V1 (visibility of data within Product)

This table has the following finalised metrics after discussion and agreement with the BI team.

Metric	Definition
Tasks Delivered	Total number of tasks completed by the contributor for the selected workflow step.
Tasks Accepted	Number of tasks submitted by the contributor that passed all downstream steps in the workflow.
Tasks Rejected	Number of tasks submitted by the contributor that were rejected at any downstream step after the contributor’s step.
Average Handling Time (mins.)	Average time spent by the contributor per task, calculated using task assignment data.

Metric	Definition
Total Time Taken (hrs.)	Total time spent by the contributor on tasks during the selected time period.
Avg Daily Working Hours (hrs.)	Average number of hours the contributor was active per day within the selected date range.
Review Acceptance Rate	Percentage of **Maker (and Triage**, if applicable) tasks that passed all downstream steps, including Reviewer and QC.
QC Pass Rate	Percentage of reviewed tasks that successfully passed the Quality Check (**applicable to the Reviewer step**).

- We did not include the Rework step in this table currently, since the BI team doesn’t have it in their dashboards, but since its within the product so we need to include all the steps.

Why the Current State Is Not Sufficient

The embedded dashboard breaks down when PMs try to *act* on this data:

- Absolute thresholds don’t scale across the workflow as “Good” and “bad” performance is relative, not fixed → Percentile calculation is very much required
- Lists cannot be reused or connected to downstream actions

To move from **visibility → control**, FTS must internalize this data (instead of just having a Looker Studio e,bedding)

▼ Architectural Consideration: How Do We Populate Metrics for Cohort Creation?

To support cohort creation, FTS needs a reliable, scalable, and consistent data source for contributor metrics. There are two architectural approaches:

Fetch Metrics from BigQuery (BI as the Source of Truth)

Use the existing BigQuery tables populated via ETL from FTS Postgres and consume the **same metric definitions** the BI team already uses. Cohort creation in FTS would query BigQuery directly (or through a service layer), ensuring alignment across dashboards and analytics.

- **Single Source of Truth:** Metrics stay 100% consistent with BI dashboards — no mismatch or confusion.
- **SaaS-Ready:** Centralized query layer scales naturally when FTS becomes a multi-tenant SaaS offering.
- **Low duplication:** No need to re-compute or maintain metric logic inside FTS.
- **Future-proof:** Cohort-based automation (next scope) can run directly on BigQuery definitions.
- Requires building a **BigQuery access layer** inside FTS.

> Engineering is already aligned with this approach for implementation in Q1’26: https://telusdigital-aicv.slack.com/archives/C09M74T3JSY/p1762517718093199

PMUS01 - Contributor Metrics (Foundation)

User Story

As a Project Manager, I want to view and filter contributor performance data directly within FTS, so that I can identify performance patterns and take informed actions without leaving the product.

Scope & Intent

PMUS01 is **not about cohorts yet**.

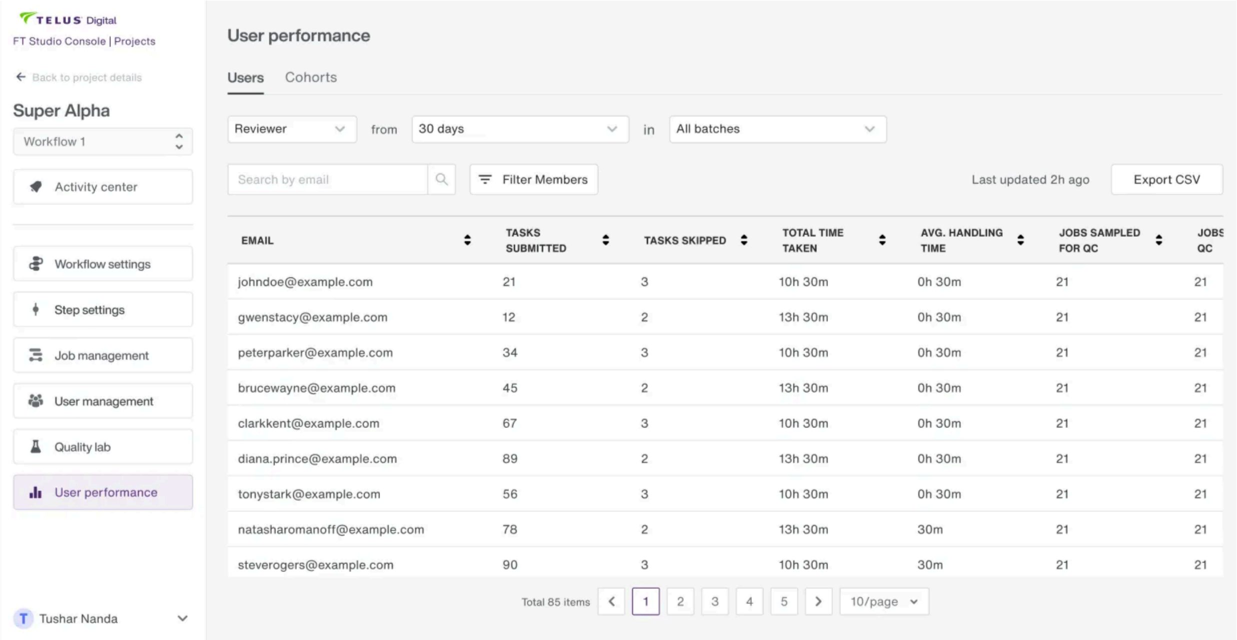
It is about making the User Performance (Contributor Metrics) table decision-ready.

This is the foundation everything else builds on.

Key Enhancements Required

Designs for Ref: https://www.figma.com/design/arjqhc1y5P0yuHq7CPbvQj/-ET--Q4-2025?node-id=1438-3774&p=f&t=k32UV38PLgP1v5i6-0

1. Internalized Metrics Table



User performance

Reviewer from 30 days in All batches

Last updated 2h ago

EMAIL	TASKS SUBMITTED	TASKS SKIPPED	TOTAL TIME TAKEN	AVG. HANDLING TIME	JOBS SAMPLED FOR QC	JOBS QC
johndoe@example.com	21	3	10h 30m	0h 30m	21	21
gwenstacy@example.com	12	2	13h 30m	0h 30m	21	21
peterparker@example.com	34	3	10h 30m	0h 30m	21	21
brucewayne@example.com	45	2	13h 30m	0h 30m	21	21
clarkkent@example.com	67	3	10h 30m	0h 30m	21	21
diana.prince@example.com	89	2	13h 30m	0h 30m	21	21
tonystark@example.com	56	3	10h 30m	0h 30m	21	21
natasharomanoff@example.com	78	2	13h 30m	30m	21	21
steverogers@example.com	90	3	10h 30m	30m	21	21

Total 85 items < 1 2 3 4 5 > 10/page

- Replace Looker iframe with a native FTS table
- Powered by BigQuery
- Same metric definitions as BI

UX behaviour: The ‘step’ filter is a single select (the PM can select only one step at a time), and whatever the first step is of the worlflow (Maker in most cases) is pre-selected here.

2. Percentile Computation (Critical)

For each metric that is available to the PM, system will keep a log of its percentile values:

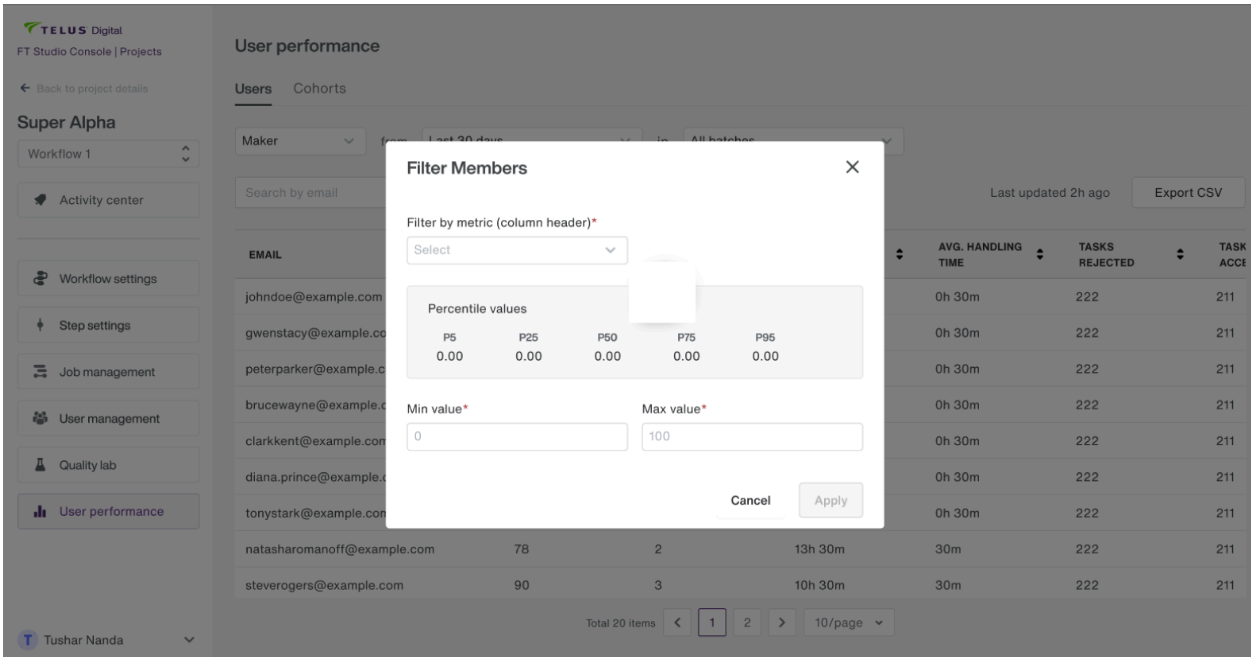
- Following percentile values will be captured by the system for every refresh:
  - Percentiles: `p5`, `p25`, `p50`, `p75`, `p95`
- Computed at **workflow + step level,** for the selected date range: Based on already-applied filters within the worklflow (date range, step)

- Consider this percentile level to be calculated as a further filter after the already selected filters, so let say the PM has selected a particular step and a date range, then the percentile calculation need to happen at this filtered data (we don’t include Batch level filtering for percentile calculation)

Why this matters:

- Enables relative performance evaluation
- Helps in identifying the true outliers (contributors at the extreme ends)
- Forms the backbone of cohort logic

3. Metric-Level Filtering



Filter Members

Filter by metric (column header)*

Select

Percentile values

P5	P25	P50	P75	P95
0.00	0.00	0.00	0.00	0.00

Min value*: 0

Max value*: 100

EMAIL				AVG. HANDLING TIME	TASKS REJECTED	TASK ACCE
johndoe@example.com				0h 30m	222	211
gwenstacy@example.co				0h 30m	222	211
peterparker@example.c				0h 30m	222	211
brucewayne@example.c				0h 30m	222	211
clarkkent@example.com				0h 30m	222	211
diana.prince@example.c				0h 30m	222	211
tonystark@example.com				0h 30m	222	211
natasharomanoff@example.com	78	2	13h 30m	30m	222	211
steverogers@example.com	90	3	10h 30m	30m	222	211

PMs can:

- Select a metric column
- View its percentile distribution
- Filter contributors using Min. and Max value range (they can view the percentile values for the metric in the modal, as shown in the screenshot above)

    - This percentile calculation happens at the step level of the workflow for the selected date range, its not differentiated by the batches.

This turns the table into an exploration and discovery tool.

4. Rework Step Inclusion

- For the metrics to be calculated, the Rework step should be treated similar to Reviewer
- Same quality metric calculation is applied [QC Pass Rate]
- This ensures full workflow visibility

5. “Tasks Skipped” Metric to be added

- Definition: Number of tasks skipped by the contributor
- This metric was not included in V1 because the BI team doesn’t include the same in their dashboards, but we received requests by PMs post-V1
- Number of tasks skipped by a contributor can be fairly treated as a quality & intent signal - high tasks skipped implies that the user skips tough tasks and only picks up easier tasks.

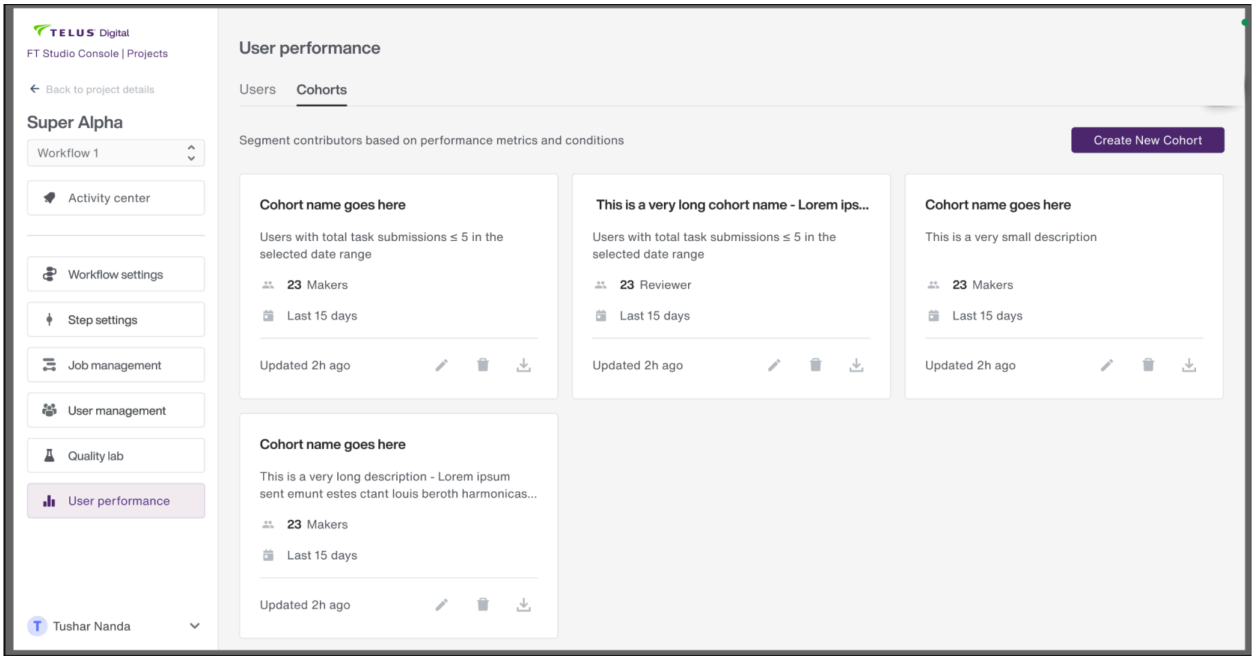
Outcome of PMUS01

After PMUS01:

- The table answers *who* and *why*
- Percentiles give context
- Filtering gives precision
- The system becomes cohort-ready

PMUS02A: Contributor Cohorts

As a **Project Manager**, I want to **create and view contributor cohorts (filtered lists of users)** based on specific performance metrics and thresholds, so that I can quickly identify contributors who share similar performance characteristicssuch as new users, low performers, or top performers- and take informed actions on them (e.g., limit task assignment, retrain, or promote).



Segment contributors based on performance metrics and conditions

Cohort name goes here
Users with total task submissions ≤ 5 in the selected date range
23 Makers
Last 15 days
Updated 2h ago

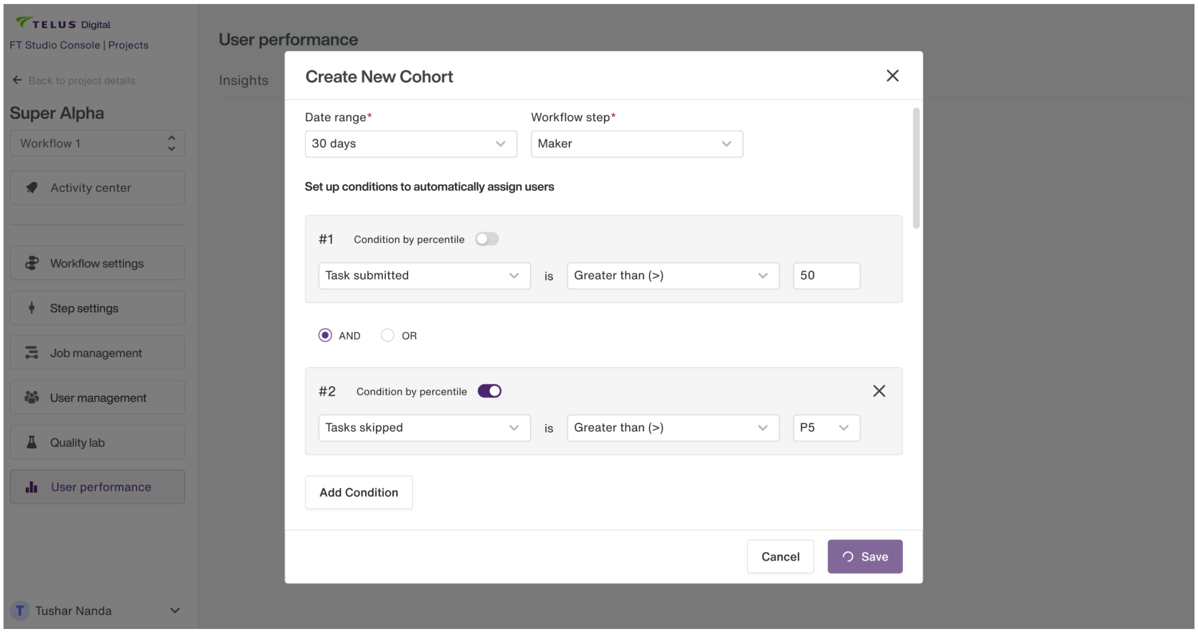
This is a very long cohort name - Lorem ips...
Users with total task submissions ≤ 5 in the selected date range
23 Reviewer
Last 15 days
Updated 2h ago

Cohort name goes here
This is a very small description
23 Makers
Last 15 days
Updated 2h ago

Cohort name goes here
This is a very long description - Lorem ipsum sent emunt estes ctant louis beroth harmonicas...
23 Makers
Last 15 days
Updated 2h ago

User Flow:

- **Cohort Creation (Custom Cohorts)**
  - PM can create new cohorts using a guided modal (button at the top of the page).



Create New Cohort

Date range*: 30 days

Workflow step*: Maker

Set up conditions to automatically assign users

#1 Condition by percentile ○
Task submitted is Greater than (>) 50

● AND ○ OR

#2 Condition by percentile ●
Tasks skipped is Greater than (>) P5

- **Steps in cohort creation:**
  1. **Select Date Range** [same filter, as we have on the user performance table]

     Defines the time window over which contributor metrics are aggregated (e.g., *Last 7 Days*, *Last 30 Days*, *Custom*).

     - **Fixed Ranges (e.g., Last 7 Days, Last 30 Days):**

       The system automatically determines the **FROM** and **TO** dates based on the current date.

       - Example: For *Last 7 Days*, the system uses today as the **TO** date and computes the **FROM** date as 7 days prior. Metrics are aggregated across this fixed rolling window.
     - **Custom Range:**

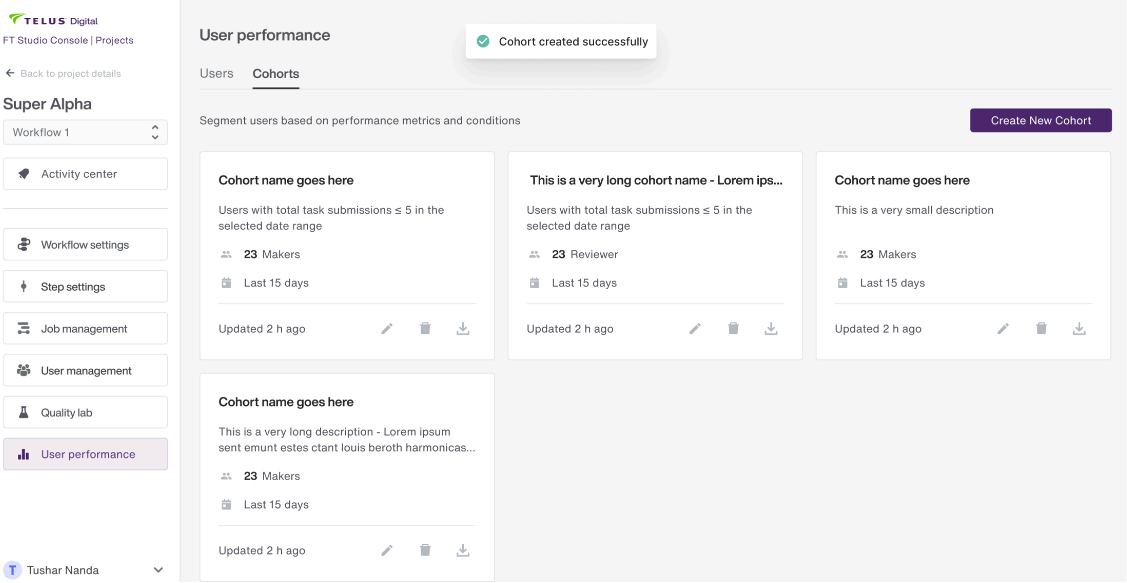
       Users manually select the **FROM** date and then choose either:

       - A specific **TO** date, or
       - **NOW**, which keeps the **TO** date dynamic.
       - When **NOW** is selected, the system continuously aggregates data from the selected **FROM** date up to the current day

(whatever data that has been updated for today), updating daily as “NOW” advances.

2. **Select Step-** choose workflow step such as Maker, Reviewer, or QC.
3. **Select Metric(s)-** choose one or more metrics (e.g., tasks_submitted, qc_pass_pct, aht_excl_idle).
4. **Define Conditions-** add threshold rules with AND/OR logic between metrics.
    - Threshold type can be **absolute** (e.g., QC Pass % < 80) or **percentile** (e.g., QC Pass % < p25).
5. **Add Name & Description –** clearly describe what the cohort represents and its intended use.
6. **Save & Generate –** system evaluates all contributors and lists those matching the conditions.

- **Cohort View**
    - Once created, the system displays a **Cohort Card** with:



User performance

Cohort created successfully

Segment users based on performance metrics and conditions

Cohort name goes here

Users with total task submissions ≤ 5 in the selected date range

23 Makers

Last 15 days

Updated 2 h ago

This is a very long cohort name - Lorem ips...

Users with total task submissions ≤ 5 in the selected date range

23 Reviewer

Last 15 days

Updated 2 h ago

Cohort name goes here

This is a very small description

23 Makers

Last 15 days

Updated 2 h ago

Cohort name goes here

This is a very long description - Lorem ipsum sent emunt estes ctant louis beroth harmonicas...

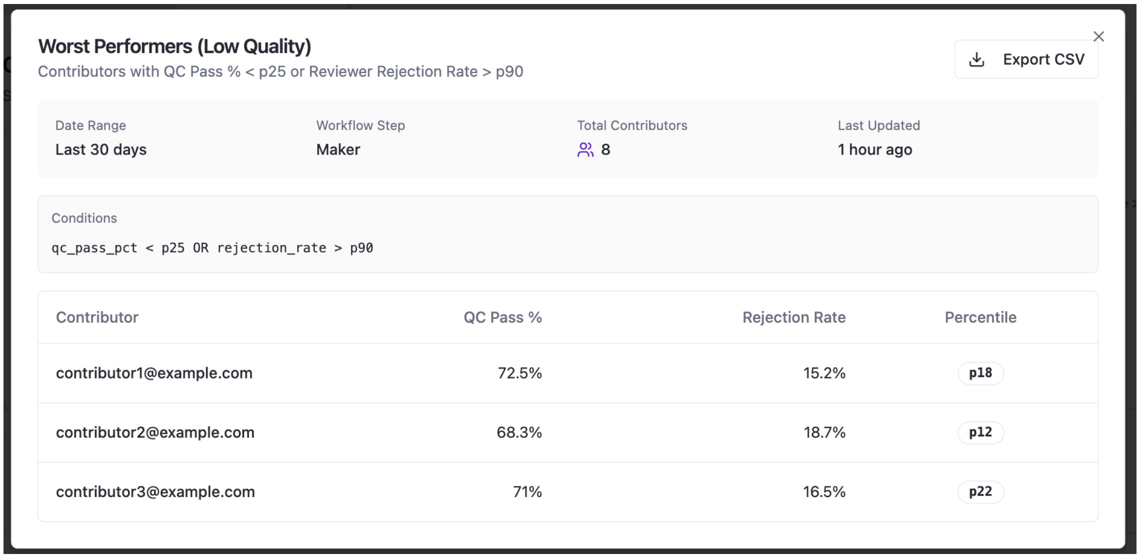
23 Makers

Last 15 days

Updated 2 h ago

- Name
- Description / Purpose

        - Contributor count
        - Last Updated timestamp
    - Clicking a card opens the detailed **Cohort List View**, showing:



Worst Performers (Low Quality)

Contributors with QC Pass % < p25 or Reviewer Rejection Rate > p90

Date Range	Workflow Step	Total Contributors	Last Updated
Last 30 days	Maker	8	1 hour ago

Conditions

```
qc_pass_pct < p25 OR rejection_rate > p90
```

Contributor	QC Pass %	Rejection Rate	Percentile
contributor1@example.com	72.5%	15.2%	p18
contributor2@example.com	68.3%	18.7%	p12
contributor3@example.com	71%	16.5%	p22

        - Contributor email/ID
        - Key metric columns (which were selected while configuring the cohort) and values
        - Percentile rank (if applicable)
        - Quick actions (e.g., download CSV, limit tasks, send training)
    - PMs can **download** the list of contributors as CSV.
- **Cohort Management:**

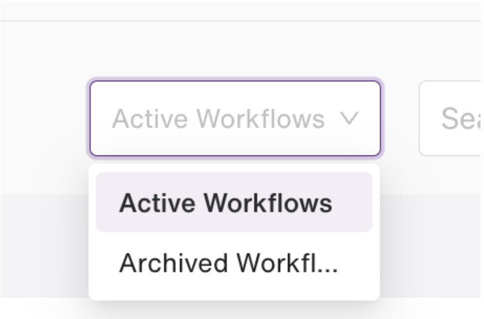
  Editing

    - Users can edit existing cohorts, including all the inputs taken from the user (date range, step, metrics etc.)

  Archiving

    - Cohorts can be **archived**, so we need to have a **Cohort Status** parameter (as part of cohort metadata) with two possible values:
        - LIVE

        - ARCHIVED
    - When a cohort is archived, its status updates to ARCHIVED.
        - If the cohort is currently linked to a feature, the behavior of the cohort in the feature after archiving depends on that feature's own logic (e.g., whether it continues using the archived cohort or stops using it).
    - Archived cohorts should be visually separated from live cohorts in the UI:
        - The Cohort Management page should include a simple switch or dropdown to toggle between **LIVE** and **ARCHIVED** cohorts (similar to the pattern used on the *Workflows* page - screenshot for reference).



        - Archived cohorts appear in a dedicated **Archived** section for clarity and discoverability.

- Cohorts can act as a reusable element across the product- for example, a PM could create a cohort of Makers who have submitted fewer than 5 tasks, and this filtered list can directly feed into features like 'Task Limit Configuration' [covered in PMUS03]

NOTE: the data in cohort cards will be updated every 24 hours (are we aggregating and calculating the base metrics every 24 hours OR are we updating the data every 1-2 hours?)

PMUS02B Pre-defined Cohorts

As a Project Manager, I want to quickly view a set of **pre-defined contributor cohorts** so that I can immediately identify the most critical groups of users (e.g., new contributors, low-quality performers, inactive contributors) without needing to configure filters myself. This helps me understand the feature, adopt it faster, and take timely actions on contributors who most directly impact project delivery and quality.

> The system should surface a small set of pre-defined cohorts that represent the **most common and high-impact contributor patterns** across projects.

Principles for Pre-defined Cohorts

- Pre-defined cohorts should map to common, recurring PM actions.
- Each cohort should answer a single PM question clearly.
- These Cohorts also should be editable, not rigid.
- Default cohorts use **simple, explainable thresholds** (absolute + percentile mix).

1. Worst Performers: Low Quality

Definition (default):

- `QC Pass % < p25` **OR** `Reviewer Acceptance Rate < p25`
- Minimum sample size (e.g. ≥ 5–10 tasks) can be defined by the `tasks_accepted` OR `tasks_rejected` metric

Why this matters:

This is the **highest-impact cohort,** contributors directly degrading data quality.

Typical actions that PM can take:

- Disable contributor
- Increase QC sampling
- Trigger retraining email

> This cohort alone justifies the entire feature.

2. New Contributors

Definition (default):

- Contributors with `tasks_delivered < 5` in the selected date range

Why this matters:

New contributors are the riskiest for quality. PMs frequently want to **treat them differently** until confidence is established.

Typical actions enabled:

- Limit number of tasks assigned
- Increase QC sampling
- Send onboarding / best-practice material

3. Inactive Contributors

Definition (default):

- `tasks_delivered = 0` in the last 7 days

Why this matters:

Inactive contributors silently affect throughput and planning.

Typical actions enabled:

- Send reminder / engagement email
- Remove from active roster
- Adjust capacity planning

PMUS03 - Automated Actions via Cohorts

As a Project Manager, I want to automate repetitive contributor-management actions

that are currently performed manually based on contributor performance metrics,

so that I can run/manage projects more proactively, improve quality and throughput, and significantly reduce time spent on operational upkeep.

Mental Model:

> **Cohorts define *who* is eligible for a given action.**
> **Each feature (and automation) defines *what*, *when*, and *how* the action executes.**

Buckets of Actions (Conceptual Framework)

Actions to be taken on Contributors fall under three broad categories:

1. User Information Actions - Updating contributor state and metadata

- User status (Active / Inactive)
- User tags

2. User Work Actions - Operational changes in task allocation (based on contributor perfomance and project requirements)

- Task limits (Task Allocation throttling based on quality, throughput, efficiency)
- Contributor QC Sampling

3. User Communication Actions - Automated email messages & nudges (communicating with the contributors)

- Training materials for low performers
- Emails at various phases (start/mid/end) of the project,
- Contextual alerts (”your quality is dipping, here are the next steps to improve”)

Feature-Wise User Stories

we'll be updating the following features with Cohort based config for this version:

1. Daily Task Limits
2. Contributor QC Sampling

▼ Daily Task Limits

As a Project Manager, I want contributor task limits to automatically adjust based on their performance cohort,
so that high risk (low-quality and new) contributors cannot overload the system with poor-quality work and high performers receive more task availability.

Goal - Create a quality-aligned task distribution system that protects daily output and reduces rework.

Why - Because task distribution based solely on volume (tasks completed) lets low performers dominate workload, increasing downstream QC and reviewer effort. So instead being able to limit the tasks based on Quality (along with throughput) is very impactful in increasing the overall quality of the project output.

User Flow

1. PM enables Task Limits feature
2. PM links one or more LIVE cohorts.
3. PM defines daily task limits for each cohort.
4. NOTE: **Multi cohort conflict resolution policy** - if a contributor belong to multiple cohorts (that have been linked in that feature) **take the lower limit**.
    a. example - contributor A belongs to 3 cohorts that have been linked to this feature:
        i. Cohort Alpha - Limit set equal to 5
        ii. Cohort Beta - Limit set equal to 10

iii. Cohort Gamma - Limit set equal to 15

In this case, the lowest limit is 5, so that is considered for contributor A and their daily tasks are limited to 5.

5. System enforces daily task limits based on these set configurations, for the contributors that are part of the linked cohorts.
6. PM can remove (delete) a cohort linking

Note: When a cohort’s data updates and contributor list changes in that cohort, the updated contributor list should take effect **immediately**. As soon as a linked cohort recalculates, the feature automatically re-applies the corresponding task limit configurations to the newly updated set of contributors.

Configurations

- Add the Daily limit per cohort while linking the cohort
- Multi-cohort conflict policy (stated out on the front end as tool tip, so the PM knows about it)

NOTE: Handling of archived cohorts - if a linked cohort gets archived then we consider it as a soft archival → a red warning message appears on the feature that a linked cohort has been archived and the PM can either delete that cohort linking or ignore it

Cases to keep in mind:

- Contributor in multiple cohorts → apply the higher limit(PM choice)
- Contributor in no cohort → default rules apply (no limit is applied to their task allocation)
- Cohort archived → system surfaces warning

- Since this is a step level config, we only allow the selection of the cohorts that have been created for that step

NOTE: we can set the **Daily Task Limit** for the relevant cohort to be **equal to 0**. This makes us in a way, operationally kind of Disable a contributor (as they won’t be receiving any tasks now), **without requiring us to create a separate automation in the Users section for disabling the user at the project level**

Why this makes sense: The decision to disable a user typically depends on additional signals outside the workflow - such as exam performance, malpractice detection, or project-level governance rules. Automating this action purely based on workflow-level performance can create inconsistencies, since disabling a contributor is a project-wide, high-impact decision with implications beyond a single workflow.

So, instead, if the PM is able to limit the tasks of such users to be equal to zero, then thats a better way to achieve this.

▼ Contributor-Level QC Sampling

As a Project Manager, I want the QC sampling percentage for each contributor to update automatically based on their performance cohorts,
so that high-risk (bad quality/low performing) contributors receive stricter QC and high performers are sampled less, without requiring constant manual adjustments.

Goal - Ensure QC sampling dynamically reflects contributor performance and reduces manual upkeep.

Why - Because contributor performance changes as they keep completing the tasks, and manual sampling updates are slow, inconsistent and require a lot of time to get the data and update it accordingly.

User Flow

1. PM opens **QC Sampling Config**.

2. PM sets default sampling percentage.
3. PM clicks **Link Cohort → selects a LIVE cohort**. (if a Cohort status = Archived, then it is not shown in the selection dropdown)
4. PM enters sampling % for that cohort.
5. PM saves the config
6. PM can link more cohorts.
7. PM can view the list of the contributors, as they see currently, which includes:
    a. the contributor email,
    b. No. of Jobs Completed,
    c. No. of Tasks Sampled
    d. Sampling % - this is the sampling % that is being applied to that contributor - this is non-editable (view only)

note- we remove the Export and Bulk Update functionality that is present in the feature currently (since its no longer required, the update is taken care by the cohorts)

NOTE: **Multi cohort conflict resolution policy** - if a contributor belong to multiple cohorts (that have been linked in that feature) **take the higher sampling %**.

- example - contributor A belongs to 3 cohorts that have been linked to this feature:
    1. Cohort Alpha - QC sampling set equal to 20%
    2. Cohort Beta - QC sampling set equal to 40%
    3. Cohort Gamma - QC sampling set equal to 60%

    In this case, the highest sampling is 60%, so that is considered for contributor A and their sampling % is set to 60%.

Note: When a cohort’s data updates and contributor list changes in that cohort, the updated contributor list should take effect **immediately**. As soon as a linked cohort recalculates, the feature automatically re-applies the corresponding QC sampling configurations to the newly updated set of contributors.

Configurations

- Default sampling %
- Cohort-wise sampling %

NOTE: Handling of archived cohorts - if a linked cohort gets archived then we consider it as a soft archival → a red warning message appears on the feature that a linked cohort has been archived and the PM can either delete that cohort linking or ignore it

Cases to keep in mind:

- Contributor in multiple cohorts → **apply the highest sampling %**
- Contributor in no cohorts → apply the set **default sampling**
- Cohort is archived → system surfaces warning
- since QC sampling only considers the last step which the job went through, so we need to allow the selection and configuration of the cohorts of that step only. For example - in a Maker-Reviewer workflow, the QC sampling can be done on ‘Reviewer’ step and on ‘Rework’ step, so while configuring the cohorts in this feature, we should allow the selection of cohorts that have been created for these two steps only.

▼ Key Metrics

Write down the metrics that this feature will improve

▼ Appendix

Contributor related actions

1. Throughput/Activity
    a. Adding a Tag to a User
    b. Inactivity nudges
2. Quality
    a. Task limits
    b. QC sampling rules
    c. Disable contributor from workflow
    d. Adding a Tag to a User
    e. Quality training emails

Items to change in current designs:

1. date range selection (having a ‘from’ and ‘to’ concept for date selection) - both in the table filters and in cohorts.
2. add ARCHIVE thing

question - if in BigQuery, the data updates every 1-2 hours, then when we select the date range to populate data, then do we get the data till the current time? for example if we select a date 1st December and another date lets say ‘Today’, then the data will be aggregated

Below are the cohorts which the PMs can create, which are not pre-defined by the system

▼ What other kind of cohorts can the PMs create: Detailed Examples

step level cohorts: why? since cohort is just a filtered group of contributors based on the action we want to take on them (actions to take, metrics to consider for filtering- will vary per step)

case: a user who is in multiple steps and has diverging actions accordingly (can be solved via union/intersection, based on the kind of actions - defined at the action level)

1) (Disable Users) Worst Performers: Poor Quality [High-Rejection + High QC Failure]

includes Makers, Reviewers.

doesn’t include QC’ers

- Condition: `tasks_submitted_30d >= 10` AND [`reviewer_rejection_rate_30d > p95(workflow)` OR `qc_pass_pct_30d < p25(workflow)`]
- Window: 30 days (and require min tasks)
- Why: Contributors causing most rework/cost.
- Action: Mark as Pending Disable → PM views the list of the users and then PM triggers disable action
- Safeguard: PM approval required; history + undo.

---

2) (Send Email) Low Activity [Low Tasks Submitted + Low Active Days Count]

includes Makers, Reviewers, QC’ers

- Condition: `tasks_submitted_7d < 3`
- Why: Capacity risk; inactive contributors reduce throughput.
- Action: Nudge email

---

3) (Send Email) Slow Executors [High AHT]

includes Makers, Reviewers, QC'ers

- Condition: `average_handling_time > p90(workflow)` **AND** `qc_pass_pct_7d < p50` (to avoid penalizing careful, high-quality workers??).
- Window: 30 days (and require min tasks)
- Why: Tasks taking too long = throughput drag.
- Action: Training Email

---

4) Limit Tasks in the queue of a Contributor

- Condition: `tasks_submitted_7d >= threshold` OR `reviewer_rejection_rate_7d > p90` OR `qc_pass_pct_7d < p10`
- Action:

5A) Dynamic QC sampling (one cohort to take action?)

- Condition:
    - Increase sampling-
    - Decrease sampling-
- Action:

5B-1)At-Risk: Low Quality [Increase QC sampling]

- Condition: `qc_pass_pct_30d < p10` OR `reviewer_rejection_rate_30d > p90` with `tasks_submitted_30d >= 5` .
- Window: 30 days (and require min tasks)
- Why: Likely needs retraining and more QC sampling, before we think of disabling them.
- Action: Email with training material; **increase QC sampling for these users(??)**

5B-2) High Performers [Decrease QC Sampling]

- Condition: `qc_pass_pct_30d > p95 AND reviewer_rejection_rate_30d < p10 AND tasks_submitted_30d >= threshold`

- Why: Contributors for whom we can reduce the QC sampling, *and/or promote them to higher-complexity jobs and promote to reviewers?*
- Action: **decrease QC sampling for these users(??)**

▼ Disable Contributor Feature

Disable Contributor - NOT REQUIRED

Disabling a contributor (i.e., updating their status to **Inactive**) should not be automated through cohort-based workflow data, for now. The decision to disable a user typically depends on additional signals outside the workflow: such as exam performance, malpractice detection, or project-level governance rules.

Automating this action purely based on workflow-level performance can create inconsistencies, since disabling a contributor is a **project-wide, high-impact decision** with implications beyond a single workflow.

Also, the operational outcome of disabling a user: **preventing them from receiving tasks** can already be safely achieved through the **Daily Task Limit** feature by setting the task limit to **0** for the relevant cohort. This makes a dedicated "Disable Contributor" automation unnecessary in the current version.

▼ Email / Notification Feature *[need to think more]*

Emails to be sent to the contributor - make use of the service developed by the platform team [define the items that our PMs would require from such a Email service - emails to be sent via hubspot and then tracking the analytics on the actions taken on those emails by the users etc.]

What This Feature Becomes

A rule-based communication system, not a one-off email sender.

Expectations from This Feature

A. Template Management

PMs can:

- Create multiple email templates
- Define purpose per template (nudge, training, warning, promotion)
- Use dynamic variables:
    - contributor name
    - metric context (e.g. “QC pass rate last 7 days”)
    - project / workflow name

Templates are project/workflow-scoped, not global.

B. Cohort Linking (Advanced Mode)

For each template, PM can:

- Link **one cohort**
- Define **send frequency**
- Define **cooldown period**

Example:

> “Send this email once every 7 days to contributors who newly enter this cohort.”

C. Anti-Spam Logic

System guarantees:

- Same template is not sent repeatedly to the same contributor
- Cooldown window enforced per (template × contributor)
- Re-entry logic:
    - Email can re-trigger only if contributor exits and re-enters cohort after cooldown

This logic is not configurable per PM, it is platform-enforced.

D. Visibility & Logs

PM can see:

- Last sent timestamp per contributor
- Why the email was sent (metric snapshot)
- Open / click rates (if available from email service)

Example User Flow (Email)

1. PM opens **Emails** feature
2. Creates or selects a template
3. Switches to **Advanced Mode**
4. Selects a cohort
5. Sets:
    - cooldown (e.g. 7 days)
    - trigger condition (on entry / periodic)
6. Saves & activates

From here, the system runs autonomously.

▼ Tag Addition to User