Hashed Email Onboarding

01/14/2026 5:07 am EST

Hashed Email Onboarding is an Alpha service available in the United States only. ID5 welcomes feedback and suggestions.

How the HEM Onboarding Service Works

Emails must be normalized and hashed using SHA-256. Normalization and hashing must strictly follow the instructions in section titled '**Normalizing Hashed Inputs**' available here.

Customers submit a file containing **normalized, hashed emails (HEMs)** to ID5. ID5 processes the file and returns a **matching output file** that maps each HEM to:

- ID5 IDs
- Mobile Advertising IDs (MAIDs)
- Any applicable Partner UIDs

Input Data Requirements

Submission Scope

Clients must inform ID5 which:

- Partner cookies that should be included in the output
- The platforms you intend to activate in so the ID5 ID can be encrypted appropriately

Each submission must include **all HEMs intended for matching**, not only incremental updates. This is required for:

1. **Opt-out compliance**
   ID5 expects to receive HEMs **only for users who have not opted out**. Files must already exclude opted-out users before upload.

2. **UID refresh**
   UIDs associated with a user may change over time.
   Submitting the full dataset ensures clients receive refreshed UIDs for optimal activation.

For measurement use cases, clients should retain all UIDs associated with their hashed emails for the measurement period of interest. Be sure to purge any data associated with opted out users.

File Format

While ID5 can support CSV for input and output files, we recommend using**Parquet** wherever possible. CSV is comparatively inefficient and more costly to process, and should generally be used only when Parquet is not an option.

- Format: **Parquet / CSV with headers, comma separated (could be gzipped)**
- Schema: File must contain one column named `hem`

Upload Location & Structure

- Data should be uploaded to an **S3 location agreed upon with ID5**, organized in folders named `/DATE=yyyy-mm-dd`, where `yyyy-mm-dd` represents the upload date.
- empty `_SUCCESS` file must be written to the folder upon completion

Upload & Delivery Frequency

- Supported frequencies:
  - Daily
  - Weekly

Output Matching File Content

The output file contains one row per matched UID.

Field Name	Description
`hem`	HEM from input
`uid`	A matching UID
`uid_type`	Type of the matching UID, which can be: `id5UID`, `encrypted_id5UID`, `android`, `ios`, `partner`
`partner_id`	An integer indicating ID5 partner ID when `uid_type` = `partner,` otherwise `null`
`ortb_matchmethod` (optional)	Enum (0–5 or 500+), per OpenRTB match method (see spec)

Note:
When the matching UID is a partner cookie UID:

- `uid_type` is always set to `partner`
- `partner_id` identifies the partner that owns the cookie

Output Delivery

- Output files are delivered in **Parquet or CSV format**
- Delivered to a **customer-specified location**
- Folder structure: `/DATE=yyyy-mm-dd/target_partner=partner_id`. When the same set of input HEMs are processed for multiple activation platforms, the output data may differ for each platform. In this case, results are stored separately under the corresponding `target_partner=partner_id` folder.
- Unless otherwise requested, ID5 will include the ID5 ID in encrypted format. This means the activation platform will have to decrypt the ID5 ID in order to activate

How does ID5 match UIDs to hashed emails?

ID5 uses both probabilistic and deterministic methods to match UIDs to hashed emails. Depending on whether your priority is precision vs scale, ID5 can adapt the pipeline configurations to help you achieve your goals.

Direct Platform Activation

If you would like ID5 to push your onboarded HEMs into an activation platform, please get in touch to discuss requirements.