Account Opening Express 1.1 API Data Dictionary

Document History

Date	Changes
11 November 2025	**Update:** Update wording for ERM
27 August 2025	**Update:** Update Email Risk Score "Good" and "Bad" examples for accuracy
7 February 2025	**Update:** Update Email Velocity ranges for Null
11 December 2024	**Update:** Included disclaimer for usage of Identity Risk Model
10 October 2024	**Update:** Branding update to remove Ekata and reflect Mastercard, and Inclusion of ERS into documentation; order of IP and phone signals in response updated to match what customers see. Updated values for Phone Line Type.

Table of Contents

Data Dictionary Overview

The following data dictionary provides information to help you understand what each of the response attributes mean, their field values, and how to assess the risk based on the attribute output in relation to your customer's data.

Note that for any attribute that lists a numeric range for field values, specific thresholds may vary by customer.

Definition & Assessing Risk

Email Domain Creation Date

Definition

- This feature comes from our Ekata Identity Graph which is sourced from authoritative data providers.
- This attribute indicates the date that the email domain was registered. Gmail, for example, will be dated to 2004.

Assessing Risk

Field Values	Good Consumers	Fraudsters
null: input email is missing, or email domain is invalid Timestamped date, e.g., 2004-08-26	**Behavior:** Submits a real email with an established domain such as Gmail, Outlook, etc. **Ekata data:** often returns a domain creation date that is dated at least 5 years back from current date.	**Behavior:** Prefers creating a new email domain or using a higher risk domain to avoid having a history of use. **Ekata data:** will often return a domain creation date that is dated very recently, within the last year.

Email is Disposable

Definition

- This attribute returns a true/false response that signals whether the input email is from a known disposable domain or not.

Assessing Risk

Field Values	Good Consumers	Fraudsters
null: Input email is missing *true*: High Risk *false*: Low Risk	**Behavior:** Submits their real email address, which is a normal working static email. **Ekata data:** Identifies the email as disposable or not if available in Ekata's Network.	**Behavior:** Prefers temporary disposable emails. **Ekata data:** Emails may be identified as disposable.

Email First Seen Days

Definition

- This attribute comes from the Ekata Identity Network, which is sourced from Ekata's global customers.
- This attribute indicates the first time that the input email has been seen in Ekata's network.
- The email address does not need to be active for it to be in Ekata's network.

Assessing Risk

Field Values	Good Consumers	Fraudsters
null: Input email is missing 0: Never seen before, high risk 1-90: Very high risk 91-365: Neutral risk 366+: Low risk	**Behavior:** Submits their real email address, which is typically used online frequently. **Ekata data:** Likely first saw the email a long time ago. Some emails have never been seen where coverage is low. Emails seen only recently are less common, since consumers do not change emails often.	**Behavior:** Typically utilizes disposable or temporary emails and change email addresses frequently. If they are impersonating a victim, they rarely use the victim's email unless they have gained full login access. **Ekata data:** Likely have never seen the email before or have seen it only recently. Emails first seen a long time ago are uncommon.

Email Velocity

Definition

- This attribute indicates the max number of times the input email has been seen in Ekata's Identity Network over the last 90 days.
- If the email hasn't been seen in the network in the last 90 days, velocity will be 0.

Assessing Risk

Field Values	Good Consumers	Fraudsters
0: Neutral risk 1: Low risk 2-4: Neutral risk 5-9: Medium-high risk 10+: High risk	**Behavior:** Typically, only transact or create accounts once a month or less. Few have weekly or daily online interactions. **Ekata data:** Returns a low velocity score.	**Behavior:** Takes advantage of victim's personal information to shop or frequently open fraudulent accounts. **Ekata data:** Returns a high velocity score.

Email Volatility

Definition

- This attribute indicates the max number of times the input email address has been seen with other identity elements (phone, IP, address) in the last 90 days in Ekata's network.
- If the email hasn't been seen with another identity element in the last 90 days, volatility will be 0.

Assessing Risk

Field Values	Good Consumers	Fraudsters
0: Neutral Risk 1-2: Medium-High Risk 3-5: High Risk 6+: Very High Risk Score <0.6: Low Risk 0.6-0.9: Neutral Risk >.9: High Risk	**Behavior:** Use the same set of identity data consistently. May sometimes use secondary identity elements in place of their primary (e.g. landline instead of their phone, or a shipping address instead of billing) in online transactions or signups. **Ekata data:** Returns a low volatility score.	**Behavior:** Match victim cardholders' names and addresses with burner phones, throwaway emails, and other temporary identity elements that they change frequently. **Ekata data:** Returns a high volatility score.

Email Risk Score

Definition

- Email Risk Score assesses the risk level of an email address. The score is derived from a model that leverages insights from the Identity Network including email tumbling detection, email linkages to other PII inputs and disposable email domain list service.
- To return a score, an email address must be provided. For best results, it is recommended sending at least one of the following as well for geolocation purposes:
    - Phone (primary or secondary)
    - Address (primary or secondary)
    - IP Address

Assessing Risk

Field Values	Good Consumers	Fraudsters
null: A service timeout is experienced Range: 0.000 – 1.000 <0.3 Low Risk 0.31 - 0.5 Neutral Risk 0.51 - 0.8: Medium Risk 0.8-0.97: High risk >0.97: Very high risk **Note:** Risk thresholds will vary per customer and country as it depends on customer's specific risk tolerance and regional behavior. This is why providing phone, address, or IP address is important for geolocation purposes.	Examples: • Mailbox velocity is low • Email domain is not disposable • Email first seen 2 or more years ago	Examples: • Mailbox velocity is high • Email first seen today • Email domain is disposable

Email Mailbox Velocity

Definition

- An integer value for the velocity (frequency) a mailbox has been seen in the past 180 days.
- A mailbox is the un-tumbled name part of an email address. For example: johndoe@gmail.com, john.doe@gmail.com, and johndoe+123abc@gmail.com all resolve to the same mailbox.

Assessing Risk

Field Values	Good Consumers	Fraudsters
null: The mailbox has not been seen in 180 days, or a service timeout has been experienced. Range: 1 – Infinity *Null*: Neutral Risk 1 – 5: Low Risk 5 – 10: Neutral Risk 11 – 20: Medium Risk 21 – 100: High risk >100: Very high risk **Note:** Risk thresholds will vary per customer and country as it depends on customer's specific risk tolerance and regional behavior. This is why providing phone, address, or IP address is important for geolocation purposes.	**Behavior:** Typically, they will have a mailbox velocity (mailbox was seen) of less than 10 times in the past 180 days.	**Behavior:** Typically, it will have a mailbox velocity (mailbox was seen) of more than 20 times in the past 180 days.

IP Risk Score

Definition

- This attribute helps identify the overall risk of an IP. The scores are powered by the Ekata risk model.
- The attribute considers metadata sourced by Ekata's Identity Graph, e.g., IP is associated with a VPN, the relative risk that the VPN carries, and behavioral characteristics sourced by Ekata's Identity Network for the associated VPN and the IP address.

Assessing Risk

Field Values	Good Consumers	Fraudsters
null: Input IP is missing, or IP is invalid <0.6: Low risk 0.6-0.9: Neutral risk ≥ 0.936: High risk	**Behavior:** Uses their normal IP address on their laptop, mobile phone, desktop, or other device, from the various places they shop. **Ekata data:** Identifies the IP address as low risk based on IP attributes and behavioral patterns that match normal good transactions if available.	**Behavior:** Prefers proxy IPs, public Wi-Fi, and other methods to hide their identity, and behaves in different ways, e.g., submitting several account openings across different merchants in quick succession. **Ekata data:** Often identifies the IP address as risky based on IP attributes and behavioral patterns that are associated with fraudulent account openings.

IP Geolocation Country Code

Definition

- This attribute comes from the Ekata Identity Graph, which is sourced from authoritative data providers.
- This attribute indicates the country location of the IP address, returned as a two-character country abbreviation (e.g. US, CA, SG, DE, etc.).
- This attribute is derived from the latitude and longitude coordinates of the IP address to return the accompanying country.

Assessing risk

Field Values	Good Consumers	Fraudsters
null: when input IP address is missing, is invalid, is in a private range, or on provider timeout Otherwise, we will return a two-character country code.	**Behavior:** Uses their normal IP address associated with their country of residence **Ekata data:** Almost always returns a country code that matches the country of residence.	**Behavior:** Prefers proxy IPs or other ways to hide their real identity, or uses their real IP in conjunction with mismatched stolen identities **Ekata data:** Can sometimes return a country code that does not match country of primary address, or is originating from a country that have higher incidences of fraud (Russian Federation, Algeria, Morocco, etc.)

IP Velocity

Definition

- This attribute indicates the max number of times the IP address has been seen in Ekata's Identity Network over the last 90 days.
- If the email has not been seen in the network in the last 90 days, velocity will be 0.

Assessing Risk

Field Values	Good Consumers	Fraudsters
0: Neutral risk 1: Low risk 2-4: Neutral risk 5-9: Medium-high risk 10+: High risk	**Behavior:** Typically, only transact or create accounts once a month or less. Few have weekly or daily online interactions. **Ekata data:** Returns a low velocity score.	**Behavior:** Takes advantage of victim's personal information to shop or frequently open fraudulent accounts. **Ekata data:** Returns a high velocity score.

IP Volatility

Definition

- This attribute is derived from the Ekata Identity Graph which is sourced from authoritative data providers.
- This attribute indicates the max number of times the IP address is seen with other identity elements (e.g., phone numbers, addresses, emails, etc.) in the past 90 days.
- If the IP address hasn't been seen in the network in the last 90 days, volatility will be 0.

Assessing Risk

Field Values	Good Consumers	Fraudsters
0: Neutral Risk 1-2: Medium-High Risk 3-5: High Risk 6+: Very High Risk Score <0.6: Low Risk 0.6-0.9: Neutral Risk >.9: High Risk	**Behavior:** Use the same set of identity data consistently. May sometimes use secondary identity elements in place of their primary (e.g. landline instead of their phone, or a shipping address instead of billing) in online transactions or signups. **Ekata data:** Returns a low volatility score.	**Behavior:** Match victim cardholders' names and addresses with burner phones, throwaway emails, and other temporary identity elements that they change frequently. **Ekata data:** Returns a high volatility score.

IP Phone Distance

Definition

- This attribute is derived from the Ekata Identity Graph which is sourced from authoritative data providers.
- This attribute calculates the distance in miles between a given IP address's geolocation and phone location if both are provided.
- The attribute is derived by measuring the latitude and longitude of both the IP address and the phone location and comparing the two to provide a distance calculation using Haversine formula distance in miles.

Assessing risk

Field Values	Good Consumers	Fraudsters
null: input IP or phone is missing, or IP or phone is invalid 0: medium-low risk (IP geo-locations are usually only precise to the postal level, so a zero usually indicates a partial address was given, rather than the IP and phone being the exact same location) 1-9: Low risk 10-99: Neutral risk 100+: High risk	**Behavior:** Typically sign up to accounts from home, work, or on commute, and only rarely from a great distance from their phone's billing address except in cases the phone is billed to a different address than the person resides. **Ekata data:** Generally, sees a small distance between IP location and phone location.	**Behavior:** Prefers proxy IPs or public Wi-Fi to hide their identity, and burner or temporary phones for which the associated location can vary. **Ekata data:** Often sees a large distance between IP location and phone location.

Phone Line Type

Definition

- This attribute comes from the Ekata Identity Graph, which is sourced from authoritative data providers.
- This attribute returns the associated line type with the input phone number and how it is used. It returns one of the following values below:

Assessing risk

Field Values	Good Consumers	Fraudsters
null: when phone number is invalid or line type is unknown (occurs rarely for some international numbers) *landline*: medium-high risk *fixed-VoIP*: medium-high risk *mobile*: neutral risk voicemail: high risk *toll-free*: high risk *premium:* high risk non-fixed-VoIP: high risk other: high risk	**Behavior:** Uses their normal personal phone, which is almost always mobile, and only rarely landline, fixed VoIP or non-fixed VoIP. **Ekata data:** Almost all phones will be identified as 'mobile'.	**Behavior:** Prefers burner phones which are often non-fixed VoIP or may enter the victim's phone or a fake phone number if they do not expect it to be called or texted. **Ekata data:** Some phones will be identified as Non-fixed VoIP, landline, or other non-mobile line types.

Phone Carrier

Definition

- This attribute comes from our Ekata Identity Graph, which is sourced from authoritative data providers.
- This attribute indicates the carrier associated with the input phone number. The phone number must be valid for phone carrier to be returned.
- The coverage is at the MVNO (mobile virtual network operator) level for most countries meaning the value returned is typically the company who is provisioning the number to the end user rather than the major carrier who owns it.

Assessing Risk

Field Values	Good Consumers	Fraudsters
Ekata returns thousands of phone carriers across the world and will return new carriers as they arise. *null*: when phone number is invalid, or carrier is unknown Carrier among the top 3 most popular for the country: neutral risk Other carriers: medium-high risk Carriers associated with burner phones: very high risk	**Behavior:** Uses their normal personal phone, which is usually a popular mobile carrier. **Ekata data:** Usually returns the name of a popular phone carrier for the country.	**Behavior:** Prefers burner phones which are often non-fixed VoIP or may enter an impersonated victim's phone or a fake phone number if they do not expect it to be called or texted. **Ekata data:** Often returns the name of a popular phone carrier, but normally uncommon carriers, especially those associated with prepaid or VoIP services, occur much more frequently.

Phone First Seen Days

Definition

- This attribute comes from the Ekata Identity Network, which is sourced from Ekata's global customers.
- This attribute indicated the first time the input phone number has been seen in our network.

Assessing Risk

Field Values	Good Consumers	Fraudsters
null: Input phone is missing or invalid 0: Never seen before; High risk 1-90: High risk 91-365: Neutral risk 366+: Low risk	**Behavior:** Submits their real phone number, which is generally kept for years, and is frequently used online. **Ekata data:** Likely first saw the phone a long time ago. Some phone numbers have never been seen where coverage is low. Phone numbers seen only recently are uncommon, since consumers do not change them often.	**Behavior:** Prefers disposable or temporary phone numbers that they change frequently. If they are impersonating a victim, then they may use the victim's phone number. **Ekata data:** Likely we have never seen the phone number before or have seen it only recently. Phone numbers first seen a long time ago are less common.

Phone Velocity

Definition

- This attribute indicates the max number of times the input phone has been seen in Ekata's Identity Network over the last 90 days.
- If the phone hasn't been seen in the network in the last 90 days, velocity will be 0.

Assessing Risk

Field Values	Good Consumers	Fraudsters
0: Neutral risk 1: Low risk 2-4: Neutral risk 5-9: Medium-high risk 10+: High risk	**Behavior:** Typically, only transact or create accounts once a month or less. Few have weekly or daily online interactions. **Ekata data:** Returns a low velocity score.	**Behavior:** Takes advantage of victim's personal information to shop or frequently open fraudulent accounts. **Ekata data:** Returns a high velocity score.

Identity Network Score

Definition

- Identity Network Score is a machine learning prediction that provides insight into how risky a digital interaction is based on activity patterns of the identity elements that are being used.
- Activity patterns that the Network Score focuses on include velocity, popularity, volatility, and age/maturity of the element(s).
    - Velocity: how often element(s) are used
    - Popularity: At how many merchants element(s) are used
    - Volatility: how often element(s) are used with other elements
    - Age/maturity: when elements were first/last seen
- Network Score is derived from the Ekata Identity Network, which is made up of more than 400M global monthly queries that surface usage patterns of identity data provided by Ekata's network of customers.
- To return a score, at least one valid element is needed. For best results, it is recommended sending as much of the following information as possible:
    - Name
    - Phone
    - Address
    - Email Address
    - IP Address

Assessing Risk

Field Values	Good Consumers	Fraudsters
null: A service timeout is experienced or did not get enough Inputs Range: 0.000 - 1.000 <.33: Low risk .33-.8: Uncertain >.8: High risk	**Examples:** ● 3 IP addresses used with primary address in last 3 months ● Primary address and email used together at 5 businesses in last month ● Email seen in 0 transactions in last 24	**Examples:** ● 15 IP addresses used with same email in last 6 months ● Primary address used in 28 digital interactions in 1 month ● 20 primary addresses used with one secondary

Field Values	Good Consumers	Fraudsters
Note: Risk thresholds will vary per customer as it depends on customer's specific distribution.	hours • Phone and email first seen together 2 or more years ago	address in 3 months • Email seen in 15 more transactions in last 2 weeks vs. 3 months

Identity Risk Score

Definition

- The Identity Risk Score is a comprehensive risk score calculated in real time that combines authoritative data (match statuses, metadata, linkages) from the Ekata Identity Graph as well as usage patterns of elements in the Ekata Identity Network.
- To return a score, the required Inputs must be provided. For best results, it is recommended sending as much of the following information as possible:
    - Name
    - Phone
    - Address
    - Email Address
    - IP Address

Assessing Risk

Field Values	Good Consumers	Fraudsters
null: A service timeout is experienced Range: 0 - 500 <250: Low risk 250-350: Uncertain >350: High risk Note: Risk thresholds will vary per customer as it depends on customer's specific distribution	**Examples:** ● Email, phone, and address are all valid and match to name ● IP is valid and risk is low ● Short distance between address and IP or phone ● Email and address first seen together 2 or more years ago ● Email first seen 2 or more years ago	**Examples:** ● Email, phone, and address are either not found or could not be validated ● IP geolocation does not match physical address ● Large distance between address and IP or phone ● 15 IP addresses used with same email in last 6 months ● Primary address used in 28 digital interactions in 1 month

Email Risk Score

Definition

- Email Risk Score assesses the risk level of an email address. The score is derived from a model that leverages features from the Identity Network's Interactions and Graph, email tumbling detection, email linkages to other PII elements, and a new disposable email domain list service.
- To return a score, an email address must be provided. For best results, it is recommended sending at least one of the following as well for geolocation purposes:
    - Phone (primary or secondary)
    - Address (primary or secondary)
    - IP Address

Assessing Risk

Field Values	Good Consumers	Fraudsters
null: A service timeout is experienced Range: 0 - 1 <~0.8: Low risk ~0.8-0.97: Medium risk >~0.97: High risk **Note:** Risk thresholds will vary per customer and country as it depends on customer's specific risk tolerance and regional behavior. This is why providing phone, address, or IP address is important for geolocation purposes.	**Examples:** ● Mailbox velocity is low ● Email domain is not disposable ● Email first seen 2 or more years ago	**Examples:** ● Mailbox velocity is high ● Email first seen today ● Email domain is disposable

Upgrading Score Versions

Ekata provides a transition period to upgrade from the current version to the new release version of models. During the transition period, both model versions are made available for test and validation. At the close of the transition period, the newly released models will automatically become the production model.

Admins can choose which model version outputs are returned in the admin panel at app.ekata.com.

Admins can also choose when to migrate each API key independently.

Your specific score distribution may be impacted with a new version upgrade. To prepare for the transition, Ekata recommends:

- Retuning your rulesets or retraining your fraud model with the new Ekata scores on a historical dataset. Ekata can provide a score backfill file that contains the account_signup_ids passed in the original queries along with the new and current scores for a given time.
- Joining the score backfill file with your internal dataset using the account_signup_ids allows retuning and retraining fraud rulesets with the newer score against known fraud outcome labels.
- Reviewing the score thresholds in your current fraud implementation with the newer score model version to determine if changes are needed.

Disclaimers

This model is designed to be an informational tool only. This model is provided as a rough estimate of authentication-based risk decisioning performance. The analysis performed by this model is a series of general estimates which are based upon the underlying information and assumptions now available. That information may change over time, and the analysis would need to be updated to reflect those changes for the analysis to be useful. The assumptions regarding authorization rates are hypothetical and there can be no guarantee that they will be achieved. Actual results may vary substantially from the figures shown. Mastercard accepts no responsibility for any losses arising from any use of or reliance upon any calculations or conclusions reached using this Model.

MASTERCARD MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING (A) THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE, AND NON-INFRINGEMENT; (B) RELATING TO THE PERFORMANCE OF SMART AUTHENTICATION OR USE OF RISK INFORMATION; (C) THAT USE OF SMART AUTHENTICATION OR RISK INFORMATION SHALL BE UNINTERRUPTED OR ERROR-FREE; OR (D) CONCERNING THE ACCURACY, QUALITY, RELIABILITY, SUITABILITY, OR EFFECTIVENESS OF THE RISK INFORMATION OR ANY OTHER DATA, RESULTS, CONTENT, OR OTHER INFORMATION OBTAINED OR GENERATED BY COMPANY THROUGH ITS USE OF SMART AUTHENTICATION OR ANY RISK INFORMATION. SMART AUTHENTICATION, RISK INFORMATION, AND OTHER MASTERCARD IP IS PROVIDED "AS IS, " WITH ALL FAULTS, KNOWN AND UNKNOWN. THE COMPANY ASSUMES THE ENTIRE RISK ARISING OUT OF ITS USE OF SMART AUTHENTICATION AND ITS USE OF THE RISK INFORMATION UNDER ALL APPLICABLE LAWS, INCLUDING THOSE RELATING TO PRIVACY AND DATA PROTECTION, BANKING, CREDIT, AND ANTI-DISCRIMINATION.