

**Find out why Bud's transaction categorization is market-leading**



# Introducing transaction categorization

As someone who buys digital technology, it's important to know about the technical ins and outs of a solution before you decide to partner up.

Financial transaction categorization is offered by many providers, but it's often just an add-on service, and rarely a core capability. Even providers focusing on categorization, like open banking providers and money management tools, may fall short on top-quality enrichment.

That's because the accuracy and detail of AI-powered processing is crucial when it comes to transaction categorization - with precise, granular results required for maximum output and value. In some cases, these features are the difference between frustrating errors and a smooth categorization experience (hello, satisfied customers!).

Not all categorization vendors are created equal, and there's a reason why Bud's transaction categorization is market-leading. In this piece, you'll get an overview of Bud versus industry standard transaction categorization technology:

## Overview

- What is Bud?
- Why is Bud's transaction categorization among the best?
  1. Granularity
  2. Accuracy
  3. Machine learning
- What does industry standard categorization look like?
- Accurate categorization underpins growth
- Summary: Why Bud's categorization is market-leading

If you're already familiar with Bud's market-leading transaction categorization, [talk to us](#) if you'd like to know more.



## What is Bud?

Bud is a financial data intelligence provider using AI & ML to transform messy, raw financial data into digestible, actionable insights. Our data intelligence platform includes transaction categorization and merchant identification, allowing our clients and their customers to better understand transactions.

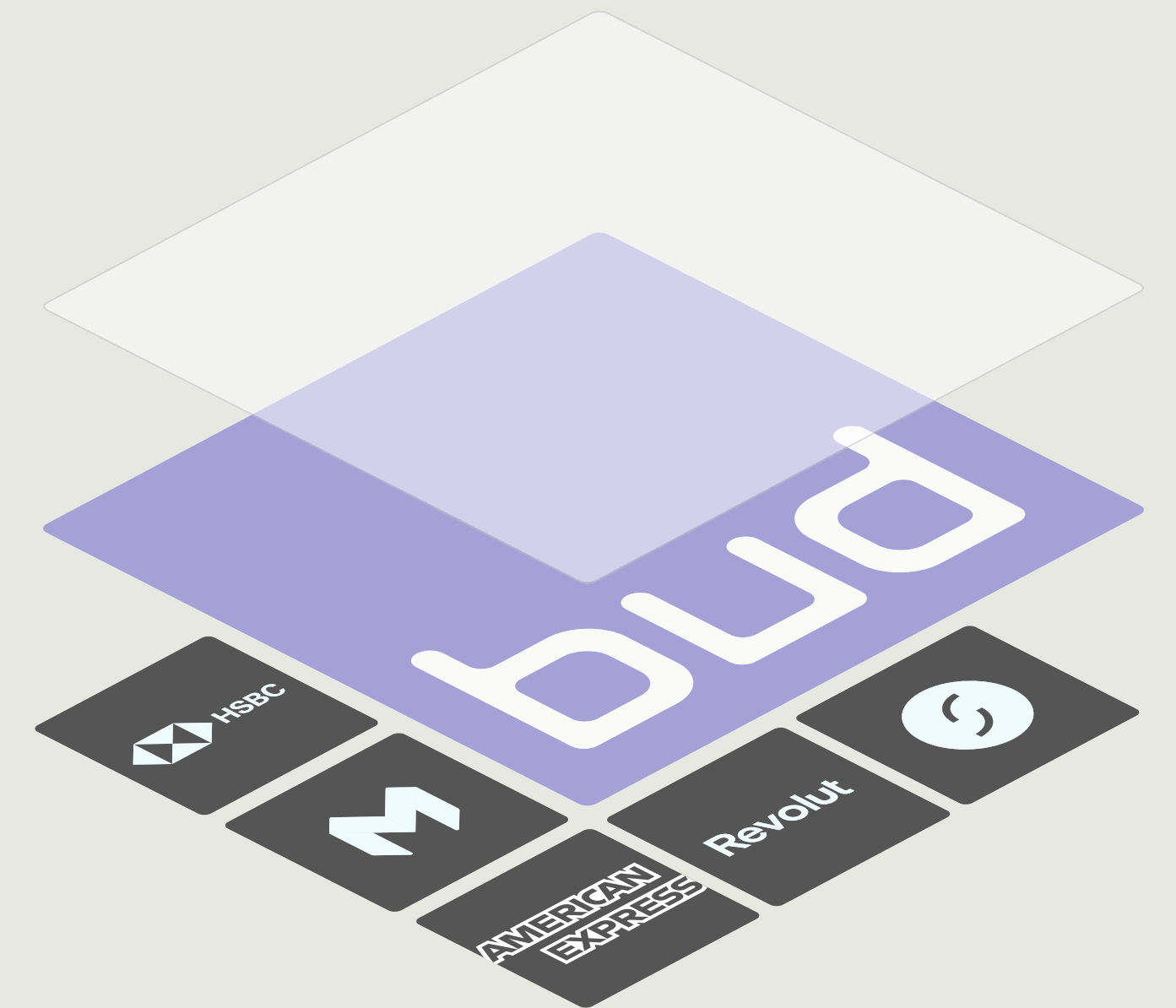
Our platform adds context to transactional data, enabling banks to better understand their customers either by using Bud's products or by feeding directly into their internal systems.

This empowers lenders to make faster and more confident credit decisions as part of their affordability calculations and, in turn, enhances the customer experience and improves their financial habits.

At Bud, our products are powered by industry-leading AI that provides unparalleled insight into transactional data. By leveraging cutting-edge natural language processing and generative models, we can categorize and enrich data for even more value - allowing us to unlock hidden insights and new dimensions that were previously out of reach.

## Why is Bud's transaction categorization market-leading?

Because we recognize AI is only as powerful as the data it's trained on, Bud's robust and powerful AI core ensures every transactional categorization is based on laser-accurate, real-time data sourced granularly.



## Granularity

Granularity refers to the level of detail in your data. At Bud, we have three levels of categorization. You can think of it as a category (L1), a sub-category (L2) plus a further set of sub-categories (L3) underneath those.

Level one (L1) could, for example, be the 'transport' category. The second level (L2) adds more context (such as: 'gas' and vehicle charging'), and the third (L3) is even more granular (such as: 'electric vehicle charging').

We only expose the L1 (category) and L2 (sub-category) to our clients for use in their apps and services. But nonetheless, the L3 category is important in identifying transactions and adding confidence to L1 and L2 categorization.

An example of how this works in practice is for lenders undertaking affordability assessments. Some competitors may find it difficult to identify transactions, such as a mortgage payment, with their categorization technology alone.

But Bud creates a 'smart finder' to identify the mortgage payment, combining categorization, extra statistical logic, plus two extra core enrichments:

### Bud's taxonomy has

- ✓ 22 categories at L1
- ✓ 112 categories at L2
- ✓ Over 200 categories at L3

1. Regularity (mortgage payments typically happen once per month)
2. Merchant identification

This helps lenders to make informed decisions.

**Description** "iz\*bank of america 01jun"  
**Amount** -1578.00  
**ID** "791fbf0e-bda6-11e9-8639-8c85905e8f95"  
**Currency** "GBP"

L1 category matched



Bank of America  
Mortgage

\$2317.00  
Monthly

## Accuracy

Bud's market-leading AI-powered transaction categorization is >96.9% accurate. Clients benefit from better data interpretation, and can build more effective money management and lending features as a result.

Bud's accuracy is industry-leading, in part, because the technologies for transaction categorization and merchant identification are decoupled. This provides flexibility. For example: a spend at 'Walmart' can map to different categories, such as 'groceries', 'gas and vehicle charging', 'mobile' or 'banking and insurance'.

But, some clients may need categorization and merchant identification models to work together. For example, recognizing a mortgage payment might involve some combination of identifying the lender as well as the amount and regularity of the payment, plus whatever we can glean from the raw transaction description. Bud's solution facilitates this approach.

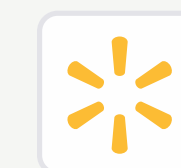
Groceries

Gas and vehicle charging

Banking and insurance

Mobile

L2 category matched



Walmart

Groceries

\$62.19

5 May

## Machine-learning

By leveraging deep learning, a subset of machine learning, Bud achieves effective natural language processing that allows us to accelerate categorization with precision and accuracy. Bud's language model not only considers which words appear in the data, but the order of the words and the transaction amount.

Let's say you spend \$2.19 at Walmart on gas.

Other competitors might categorize this transaction as 'supermarket' (Walmart) or 'transport' (gas). But the transaction amount is small, so Bud's data intelligence platform correctly categorizes this purchase as 'convenience store'.

The other advantages of Bud's machine-learning algorithm include:

- Ongoing iterative improvements to categorization at L3 level without service interruption (made possible because that categorization is vital but not exposed to clients)
- Behind-the-scenes experimentation to match evolving merchant technologies.

Unlike industry-standard rule-based approach, there's no need to rely on your service provider to edit thousands of rules as each new merchant enters the market. Bud maintains the system, keeping it performant and up-to-date.

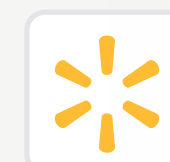
✓ 98% accuracy for L2 categorization

**Description** "iz\*walmart gas 24apr"

**Amount** -2.19

L2 category matched

Convenience store



Walmart

Convenience store

\$2.19

24 Apr



# What does industry standard transaction categorization look like?

## Poor categorization granularity

The first thing to note when considering what industry standard categorization looks like, is that not all transactions involve merchants. Salaries, benefits, rent, loan payments, transfers to peers and more – the ability to properly categorize transactions that lack traditional ‘merchants’ through text classification is key.

Some vendors struggle with this, and the result can be a money management tool with lots of important and high value transactions categorized as ‘general’ or miscellaneous’.

Elsewhere, those using transactional data to get a view of the financial position of their customer or product applicant risk making a poor decision, given they don’t have the full picture.

## Open banking providers typically use less granular taxonomies than Bud

At Bud, we’re confident that our AI-powered categorization granularity can power pretty much any use case and, of course, we easily and iteratively add more granularity when data becomes available.



LIHEAP  
General

\$700  
Monthly



LIHEAP  
Benefits

\$700  
Monthly



Another limitation to traditional categorization is the rule-based approach typically used. Their understanding of transactional data can be poor in the case where they only recognize particular or exact strings of words or patterns.

This is highly susceptible to errors when raw transaction descriptions are low quality, when words are abbreviated or when there are no spaces between words.

Some competitors use MCCs (merchant category codes) to attempt to categorize transactions. These four-digit codes classify businesses according to the types of goods or services they provide.

There are numerous problems with this approach – the most pertinent being that they're not always present. Even when businesses do have these codes assigned, they can be incorrect, extremely broad or not especially consumer-friendly. In addition, only one category is associated with a merchant which doesn't account for those businesses that do lots of things.

Likewise, combining transaction categorization and merchant identification into a single capability is a less flexible approach. Without the consolidation of regularity and noise identification, clients will find it difficult to identify or analyze spending patterns.

With transaction categorization and merchant identification merged into a single capability, financial institutions won't benefit from the safety net of Bud's double validation.

A lack of understanding around transaction data means a financial institution is forced to keep their personal financial management suggestions vague and customers may look elsewhere for better services.

## Unrecognized new merchants

In the US, [new business formation runs around 5 million per year](#). Without significant effort, these would likely be unrecognizable and categorized under 'general', which is frustrating to customers and unhelpful for credit assessment advisors.

Even among some competitors with machine-learning based algorithms, many lack regular updates and it can be painful for clients to upgrade as technology evolves.

Instead, Bud can recategorize transactions when customers provide feedback, providing flexibility in the system.

## Reliance on MCC codes

When first-party data is available, it's not uncommon for most providers to settle for the simple mapping of MCC (merchant category codes, a market standard for card networks and other payment providers). Initially, MCC codes were designed with a different purpose in mind, focusing mainly on the business aspect, instead of actual customer-facing offerings. Therefore, these codes are often configured per POS (point of sale), are rarely updated and tend to be inaccurate - leading to suboptimal quality of categorization, with incorrect codes being sent and data being underutilized.

Bud's approach is to look at all data provided with every transaction and build models that are able to spot patterns across the whole dataset. Combined with superior AI models and the ability to quickly train for new signals, Bud has the ability to outperform providers utilizing a MCC-centric approach to categorization.

## Accurate categorization underpins growth

If you're a financial services company that doesn't require a high level of detail within transaction data, then industry-standard transaction categorization might do the job for now. However, considering how things will change as you scale is an important factor.

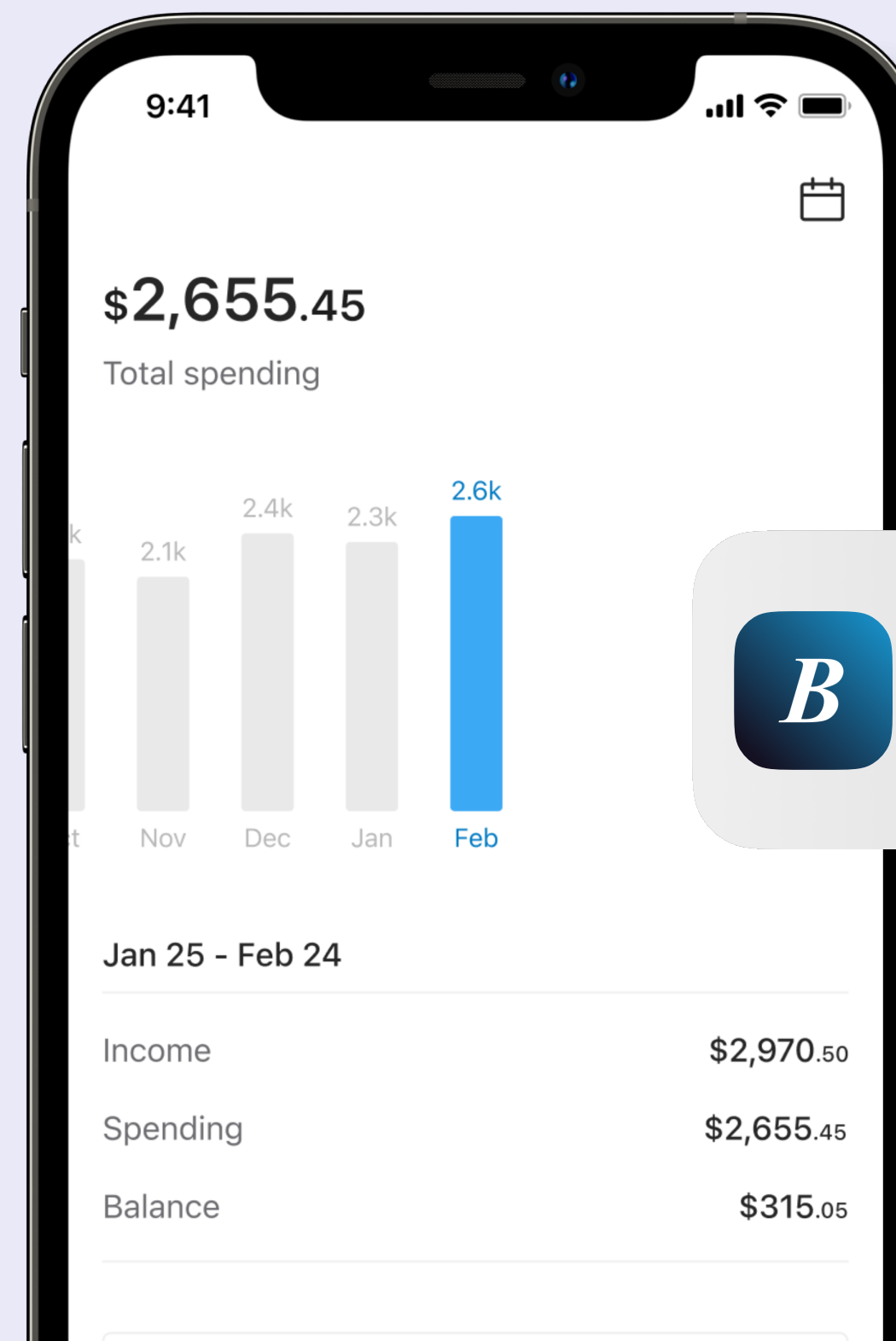
For those who want to greatly improve customer engagement and retention, or drive cost savings and financial inclusion for lending, ultra-accurate categorization is a must. Bud's gold standard transaction categorization is more likely to fulfill your needs.


## But why is accurate categorization so important?


With quality categorization, understanding the customer becomes much easier. Bud uses it to power several high-impact products, from supporting lending decisions


with Assess, supercharging digital channels with insights and money management tools available with Engage, to generative AI-assisted customer data platform - Drive.

However, many of our clients leverage the power of categorized data directly, which we can further enrich with location information, regularity detection, subscription and income finders and more.



 **You've exceeded your daily budget** now  
Take a look at recent trends to get more insight into your spending

 **SHOPPING**

 **BILLS**



## Summary: Why Bud's categorization is market-leading

### Bud's categorization

- Taxonomy of over 300 categories for industry-leading accuracy.
- Contextual recognition – every word has meaning and the order of words has weight.
- Transaction categorization and merchant identification are separate capabilities.
- L3 categorization is constantly iterated without interruption to the service (including the ability to re-categorize based on customer feedback).

### What this means for you

- Improved granularity equals minimal transactions left uncategorized – heightening accuracy.
- With complete contextualization, categories are almost always accurately defined and updated – regardless of phrasing.
- Bud's customizable categorization provides enhanced power and flexibility.
- Ensure categories remain relevant, fluid and accurate with continuous learning and re-categorization.

Bud.ai

**Get in touch with the team at Bud to understand how ultra-precise transaction categorization can transform your business**

**Get started**

