ID.me

# Why the Industry is Moving Away from KBV

Knowledge-based verification (KBV), also sometimes referred to as knowledge-based authentication (KBA), is a method of verifying someone is who they say they are by asking them a series of questions before providing access to applications or websites that contain sensitive data or enable financial transactions.

# The Typical KBV Process:

1. The user provides personally-identifiable information, like their name, birthday, and home address.
2. The verifier automatically generates a series of multiple-choice questions – often sourced from publicly-available or legally-purchasable databases.
3. The user responds to the questions and their answers are checked against the information in the databases.

## Knowledge-Based (Identity) Verification versus Knowledge-Based Authentication

The terms KBA and KBV are often used interchangeably, which creates confusion. Knowledge-based tests can be used for two purposes: 1) identity verification of a new user and 2) authentication of an existing user.

## Identity Verification Use Case

Before granting access to an account that safeguards personal data or initiates financial transactions, organizations need the user to confirm their identity. In these scenarios, "dynamic" knowledge-based verification (KBV) is typically used.

A series of multiple-choice questions are generated dynamically from public and financial records then presented to the user. Common examples include:

- **Which of the following retail credit cards do you have?**
- **What state did you reside in during 2015?**
- **What is your approximate monthly mortgage payment?**

## Authentication Use Case

Before allowing an existing user to return to their account, organizations need to confirm that the user is the same person who previously created the account. In these scenarios, "static" knowledge-based authentication (KBA) is typically used.

The user selects 3-5 questions about themselves from a predefined list of options. Examples might include:

- **What is your favorite food?**
- **What middle school did you attend?**
- **What is your favorite book?**

The user supplies their answers to these secret questions, which are stored on file. Each time the user attempts to login, they must answer one or more of these questions to authenticate.

# The Problems with KBV

## "Secret" information is not actually secret

The clues that you would need to answer most of the dynamic KBV questions are easily researchable online, in many cases. You can find educational background, home values, mortgage payments, car registrations, birthdates, and social security numbers on:

- **Social media**
- **In public records**
- **On the dark web (thanks to earlier data breaches)**

Bots can easily collect this information and rapidly respond to the questions. In fact, some financial institutions are now putting controls in place that identify when questions are being answered too quickly as the speed might be indicative of a bot.

Systems that are solely reliant on KBV can't adequately defend against access from unwanted sources. The average person can often guess the right answer to the multiple-choice questions with a search engine and some common sense. When so much personal data has already been exposed by previous data breaches and uploaded to the dark web, the hacker's job only gets easier.

## IRS "Get Transcript" Incident

The weak security of KBV has caused many data breaches, the most infamous being the IRS data breach of 2015 that compromised the personal information of around **700,000** taxpayers. The breach was made possible by the IRS's Get Transcript feature, which enabled taxpayers to view most of the line items from their tax returns over the past several years.

There were two steps users had to take to access their tax transcripts. The first was to enter personal information such as social security number and address. Much of that personal information was already available on the dark web due to previous data breaches.

That meant the only other safeguard was the second step, KBV. The malicious actors were able to answer the personal questions, likely by taking advantage of patterns in the multiple-choice questions. For example, hackers can:

- **Write algorithms to select "none of the above," a frequent answer to KBV questions.**
- **Make educated guesses about the banks that own a mortgage by cross-referencing a person's home address with the locations of the banks presented as answers to the question, and choosing the bank that operates near the individual's home.**
- **Calculate mortgage payments using public-facing tools like Zillow that list how much homes have been purchased for in the past.**
- **Determine the type of car someone drives by looking at photos on social media, or even looking at the Google Street View of someone's house.**

According to the IRS, the hackers were unable to infiltrate about 500,000 accounts. However, a less than 50% success rate at blocking fraudulent attacks is not comforting, especially for the 700,000 people KBV failed to protect. As a result, the IRS discontinued the use of KBV shortly after the attacks.

## Low Pass Rates

Often, KBV pass rates don't break the 70th percentile. As questions increase in difficulty, and therefore security, the pass rate from legitimate users drops. In other words, KBV does not allow for a secure path that all legitimate individuals can pass.

When questions are generated from databases, they could ask for someone's past five addresses or details about their credit history that users might not remember (if they ever memorized the information in the first place). Difficulty remembering seemingly-random facts is one of the primary reasons legitimate users fail these checks.

## Google Study on KBA Effectiveness

Users even forget relatively simple answers when they choose the questions themselves for authentication purposes. A 2015 study by Google found that users forgot their favorite food **47%** of the time. Meanwhile, white hat hackers were able to guess the right answer 20% of the time by choosing "pizza" – the most common favorite food of Americans.

If legitimate users can't even remember their favorite food, remembering granular details from their past shouldn't be expected.

# What's the Solution?

## Identity Verification Alternatives

KBV creates a situation where it is too easy for legitimate individuals to fail and arguably even easier for remote-based fraudsters to successfully attack at scale. In its most recent Digital Identity Guidelines (800-63-3), the National Institute of Standards and Technology (NIST) no longer considers KBV a strong piece of evidence for identity verification.

Instead, NIST suggests that data requiring high-assurance protection should collect at least two pieces of strong evidence, like government IDs, and one piece of fair evidence, like checking phone ownership with telecom providers. Both of these pieces of evidence can be verified with a higher degree of certainty than KBV:

- **ID verification relies on machine vision to quickly confirm authenticity of the document and match the name on the document with the name submitted.**
- **Mobile phone verification relies on a combination of factors like checking device tenure and looking for signs of fraud, such as SIM swaps.**

The only way hackers could defeat these methods is to get physical possession of the ID or device, which can't be done in a remote and scalable manner.

In addition to being more secure, these methods are also more inclusive because they don't rely on someone's capacity to remember obscure details, and instead rely on IDs or mobile devices that most every American has.

## Authentication Alternatives

Many organizations are also moving away from KBA as an authentication mechanism. NIST also recommends replacing KBA security questions with multi-factor authentication (MFA). Both passwords and answers to security questions are considered "something you know" – they rely on information that should be in your head and nowhere else. However, as we've seen, much of that information is likely exposed on the dark web or can be easily guessed.

MFA, on the other hand, combines a password (something you know) with "something you have," like a phone that can receive randomized codes, or "something you are," like a fingerprint. That way, even if your password is exposed to hackers on the dark web, they won't be able to access your account because they don't have the physical device or biometrics.