TRUSTLOGIX

The Top 7 *Cloud Data Security* Issues That You Don't Know About

Chances are, if you are reading this, you already have many data projects in the cloud, and more are on the way.

Storage is cheap, and all the major cloud data platforms are enabling more types of data projects including data applications, data sharing, machine learning, and now generative AI, in addition to old-school analytics. Truth is, you're never going to be "done" with your cloud data journey.

Also, chances are you are probably already taking measures to keep your sensitive data safe and secure.

At least, you think your data is safe and secure. After all, you have access control policies in place. You are probably using the data platforms' native access control functionality, or perhaps using the legacy tools that worked for you while still on-premises. You have a process in place for reviewing and approving access requests from your users that, other than them complaining how slow it is, seems to be working fine. Your monitoring tools aren't alerting that anything suspicious has happened.

Bottom line, you haven't been breached and you haven't failed an audit, so everything's great, *right*?

Right. So thought many victims of breaches and failed audits before they actually happened. This is because they overlooked lurking issues that weren't obvious without specifically looking for them, and these issues are usually the ways that exploits and leaks happen. Moreover, these issues are easy to occur in the cloud, precisely because storage is so cheap, and it's so easy to stand up a multitude of data projects, making the sheer volume and distribution of data go beyond anything you have had to manage before.

Let's look at the top seven of these issues. Ask yourself: ***Are you sure you don't have these lurking in your organization? Do you have tools in place to detect these?***

1. Dark Data

In the cloud, it's very easy to create copies of data for specific one-off needs and then leave them lying around. Your application team needs some data for testing? Give them a copy. Your data scientists are training a machine learning model? Create a copy. Sharing data with a business partner? Copy it and grant them access to that copy. Your data engineers want to test out some new data integration or data quality jobs? More copies.

Storage is cheap, and so many copies can lie around and barely move the needle on your monthly cloud bill. And they aren't being accessed anymore, so they don't show up in your usage monitoring reports. One study discovered 30% of scanned customer cloud data stores on Amazon Web Services, Azure, and Google Cloud are dark/ghost data.

But the existence of this data, with the numerous accounts that could still access it, **represents an unnecessary potential attack surface**, and should just be deleted.

Think this isn't a big problem?

On average, **TrustLogix customers and prospects found tens of thousands of unused tables across their cloud databases**, and the sensitive data therein is, without exception, eye-opening.

2. Role Explosion

As more users and projects are added, and more datasets are onboarded to your cloud data platforms, chances are you create more and more roles to allow specific groups of users to access different subsets of data for different jobs. It's easy to create new roles, and most of the time, your administrators just "check the security box" and quickly define a role-based access policy for that new project. **And so, as given datasets are used across multiple projects, and given users get involved in multiple projects, you will end up with "role explosion," in which there exist many roles with overlapping privileges**.

Sometimes this results in a user not getting access that they should—and they will complain about this and you fix it, perhaps after painstakingly troubleshooting which role is to blame.

But what if this results in users being overprivileged? You will never get complaints from your users about too much access. This leads to increased risk of data misuse and compliance violations, or, even worse, a bigger "blast radius" if an overprivileged account got hijacked by a bad actor.

Think this isn't a big problem?

On average, **TrustLogix customers and prospects had many hundreds of overlapping roles, many of which did lead to overprivileged accounts**, and were able to rationalize and quickly reduce their role hierarchies to a more manageable state.

3. Ghost Accounts

When employees or other users leave or transfer, their account should be closed, or at least its access policies should change. Problem is, even in organizations where account revocation is part of the user offboarding process, it is easy to overlook one or more databases, given how easy it is to spin up new database instances and schemas in the cloud. In fact, a Varonis study estimated 46% of companies had more than 1,000 ghost user accounts.[1]

Think this isn't a big problem?

On average, **TrustLogix customers and prospects had many hundreds of ghost accounts, each one representing an unnecessary potential attack vector, and often the former employee or other user will still remember their account credentials**. It takes just one of them to have bad intentions for a leak or breach to occur.

1 https://www.darkreading.com/identity-and-access-management/one-third-of-internal-user-accounts-are-ghost-users-

4. Overprivileged Accounts

While role explosion is one way for overprivileged accounts to occur, far more common is simply assigning the user the wrong privileges in the first place. Somebody has admin privileges when they shouldn't, or was mistakenly assigned access to a schema that they shouldn't. Perhaps your database administrator (DBA) granted admin rights to your data science leads or your application testing leads just so they could stand up their database instance more quickly and get off the DBA's back, and this DBA had every intention to revoke admin rights when they were done, and just forgot.

Perhaps you are monitoring for unusual activity, using data loss prevention or other monitoring tools, but the truth is, these tools are just finding exfiltration as it is happening. They are not preventative. You need a way of discovering overprivileged accounts before they get hijacked, or better yet, preventing them from being assigned in the first place.

Think this isn't a big problem?

On average, **TrustLogix customers and prospects found over 30 accounts with admin privileges that they shouldn't have had**, each one representing a significant potential attack vector.

5. Data Sprawl

Because it's so easy to stand up new database instances and new data pipelines in the cloud, and storage is so cheap, chances are that you have critical datasets being shared across many databases across your organization. Modern cloud-native ETL tools like Fivetran and Matillion and streaming technologies like Kafka make it easier than ever to share data across database instances in the cloud. And with most of the cloud data platform vendors offering data-sharing features, this "data sprawl" can now extend to your customers and partners as well.

As if maintaining good access controls on individual databases wasn't hard enough, it's even harder across numerous database instances. And it's impossible for those database instances you don't know about.

Think this isn't a big problem?

On average, **TrustLogix customers and prospects did find multiple data pipelines in place where data was being transferred out to a target they weren't actively monitoring**, each one representing a potential vector for exfiltration.

6. Shadow IT

Your users want their data and they want it now! With the speed of business only increasing, and new data projects being easier than ever to stand up in the cloud, this is understandable. Your users complain about your access request approval processes being slow, but you are doing your best, and they get their data eventually, so not a problem, right?

The problem is, with business pressures being as they are, not all of your users will just wait patiently. With role explosions and overprivileged accounts and these other issues likely occurring in your organization, it isn't hard for an impatient user to find another way to get the data they need by using one of these accounts or by using a client tool they shouldn't be. Studies confirm this continues to be an issue, one noted 32% of workers use unapproved communication and collaboration tools.[2]

Think this isn't a big problem?

On average, **TrustLogix customers and prospects found several unauthorized clients accessing sensitive datasets, as well as single accounts being used by many different clients** (in other words, those accounts were being shared with many users who then logged in from clients the original account owner should not have).

2 https://www.beezy.net/2022-workplace-report

7. Believing You Don't Have These Problems!

This seventh issue is unlike the previous six. It is not a specific type of security risk like these others. Instead, it is the issue of human nature and organizational behavior.

This can happen in the best-run organizations: Business goals and deadlines are aggressive, and your business users are clamoring for data and they need it yesterday. Your data teams are stretched thin as it is, and your security teams are tired of being seen as the bottleneck for the business achieving its goals. You have tried to do the right thing, checked all the proverbial boxes by quickly defining the right access control policies, so you can move on and keep up with the business's next set of demands. You aren't seeing any alerts; nothing bad seems to be happening. So, why look for problems when your team is already so busy? Here's another example: Your data engineers and admins wrote some great scripts for managing access controls, and they have complete confidence they got it right, so data security is "solved." Why second-guess them when there is so much other work staring them in the face?

You can call this pride, blissful ignorance, or just laziness. We choose to call it human nature, and no organization is immune. But if you put a cloud-native data security posture management platform in place that automatically detects these issues, then this last issue is much less of a problem. You will see them in your alerts and your reports, and moreover we let you immediately act on them, so that it doesn't seem like yet more work for your data and security teams. Indeed, TrustLogix will lead to less work because you are putting a better foundation in place to manage access requests faster, manage roles better, eliminate dark data and ghost accounts faster, and provide better visibility and control across all your database instances from one place.

Still think your organization doesn't have these issues?

See for yourself with our **Free Data Security Audit Scan Report**. Within an hour, you will be able to see how many of these issues you have, and how serious they are.

TrustLogix
800 W. El Camino Real, Suite 180
Mountain View, CA 94040

hello@trustlogix.io
www.trustlogix.io
Copyright © 2023 TrustLogix Inc.