International Journal of Leading Research Publication (IJLRP)

E-ISSN: 2582-8010 ● Website: www.ijlrp.com ● Email: editor@ijlrp.com

Identity Security From Human, Machine, And AI Identity Perspective

Anand Athavale

Independent Researcher, Decades of Industry experience in Data Management
andyathavale@gmail.com

Abstract:
This article explores a different approach towards non-human or machine identities and looks at AI as part of it. Security teams looking at non-human identities as risk have more than one option than brute force monitoring and behavior analysis. There are different approaches to identity security which require evaluation for machine identities. Identity security for non-human identities has some unique challenges and solving those requires taking a empathy-based approach towards security teams than piling on more responsibilities. Gen AI identities also need to be evaluated on similar ground of non-human identities while considering the nuances.

Keywords: Identity Security, Tier-0, Human vs. Non-Human Identities, Gen AI controls.

Introduction

Non-human entities, formerly known as service accounts, became necessary for automating base system functions without manual intervention. As digital systems matured, authentication became a prerequisite for executing most actions, reducing the number of default or unauthenticated operations. Authorization mechanisms became necessary to control *what* authenticated users or systems could do. However, many core operating system processes and services needed to run continuously and reliably, without a human user authenticating and authorizing each step. To handle this requirement, service accounts were created. Thus, non-human service account identities were designed and created to execute automated tasks securely, often with minimal or scoped authentication and authorization requirements [1]. This approach balanced security needs for human users with the operational demands of automated systems.

Types of modern-day service accounts, or, Non-Human Identities

Service accounts were initially only limited to operating systems. However, service accounts as a concept was picked up by non-core applications to solve for the same automation requirements to carry out repetitive and underlying tasks. These application tasks were of both types, operational, which ran continuously, and functional, which were collections of tasks resulting from an action a human user initiated using the application. Applications also needed non-human accounts to carry out these tasks. Sometimes applications decided to simply re-use the existing service accounts provided by the operating systems. However, as the security mindset grew mature, the applications started pivoting to custom service accounts which had just enough privilege to do what was required by the application. This principle was called the Principle of Least Privilege [2].

Once cloud and similar environments were created, service account equivalents such as service principles were created. Service principles evolved from having static client secrets to support certificate-based authentication based on verifying the certificate against the public key generated and stored during the app registration process. Finally, cloud has also introduced managed identities which do not have any secrets or certificates.

Besides these, API tokens can be considered a form of non-human identity, but API tokens are primarily tied to an entity like a user or a client. A client could be a machine identity or a service principle. Common examples of machine identities are AWS IAM role for EC2 which is identity assigned to a virtual machine, Azure Managed Identity which is auto-managed service principle for VMs, apps and functions and GCP service account which is identity for Compute Engine (VMs), Google Kubernetes Engine or Cloud Run.
There are repeatable process automation platforms like UiPath, Automation Anywhere, Blue Prism but those still use traditional human like authentication mechanisms. Besides these, there are a few variations of practice, where even human identities are sub divided into daily use and privileged accounts. Daily use identities are used only for email and general tasks. IT use credentials are considered privileged and often use privileged access management. Some organizations have "break-glass accounts" which are only for emergency use only and are highly audited. Some practices go with temporary elevation of existing accounts often referred as Just-In-Time access. Privileged accounts are sometimes further subdivided based on roles and functions.

Known or Identified risks for Non-Human Identities

The currently known or identified risks for non-human identities can be grouped into two major categories.

1. Credential storing

This group mainly consists of credential sprawl, where service account secrets, API keys or certificates are stored across code, scripts, pipelines, and config files [3]. Hardcoded credentials embedded in source code, container images, or accidentally spewed in logs also come under this category.

2. Lack of Hygiene

This is a broader group which consists of over-privileged access, improper management and lack of governance [4]. Over-privileged access can refer to broad admin-level access instead of minimum-required-privilege, misconfigured trust policies with wildcards and improper scoping where identities are not bound to namespaces or environments. Improper management can be lack of rotation, expiry, and orphaned identities without services. Shared and reused identities used by multiple services or pipelines also falls under improper management. Lack of governance mainly points to inadequate or missing visibility into usage and lack of auditing, which can be also called lacking of accountability.

A different approach than "monitor all NHI behavior"

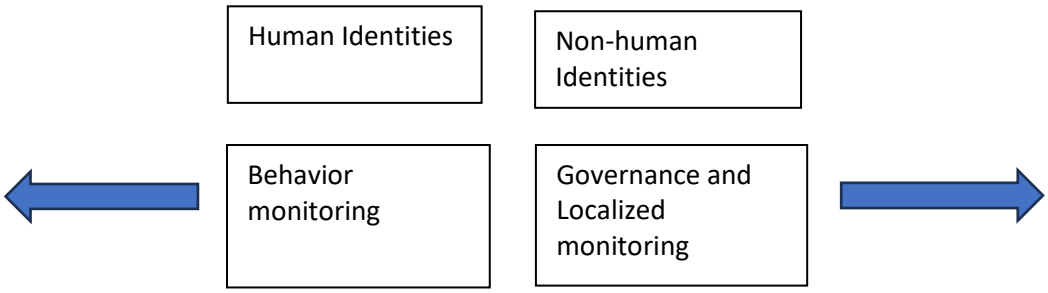
There is one school of thought emerging which suggests monitoring behavior of Non-Human Identities to mitigate risks brought by explosion of non-human or machine identities. This is originating from a fact that there is an explosion of non-human identities where non-human or machine identities are many times than human identities. Here is an analogy. Imagine a scenario where you have many robots for performing several tasks on your behalf. There is a robot for cleaning, a robot for cooking, one for driving and similar tasks. You have commanded and authorized these robots to carry out those tasks. But now you are worried about someone else misusing these robots. What do you do? Start monitoring all their actions? If you do, weren't you better before all these robots then? To give analogy of proposed solution for this scenario, the solution being proposed is to have another monitoring robot which would observe unexpected or unintended behavior and report back.

The risks related to non-human identities can be explained as follows. If you had one hundred human identities before, there was chance that one of those could be compromised and that would lead to a cyber-attack or malicious behavior. Now, you have five thousand and one hundred identities counting the added 1:50 machine identities. So, your risk has increased multi-fold. But, if we tried to look behind the attacks with machine-based identities, we are unable to find any breach which was directly through a machine identity. What we find instead that there was a human-identity compromise or failure of security practice first, which then led to machine identity compromise. An example of a failure of security practice is copying and pasting secrets into public facing document [5]. Even if we consider the credential dumps of machine identities sold on the

dark web, how were those dumps created in the first place? Those were obtained through a human identity compromise to get inside parameter walls of many organizations. There are occurrences of stolen data due to SQL injection which may not have involved human identity compromise, but those methods still require some level of code knowledge to then manipulate the SQL code to get more data out than intended. In very rare scenarios, the exposure of SQL code results in public, which means again the attacker has to compromise a human credential to get in first.

Given this to be the current situation, may be before we start monitoring behaviors of machine identities themselves, we could start where we track the interactions of human-identities to non-human identities while treating non-human identities as security sensitive data items. This way, the anomalies seen during creation, access, rotation, and expiry of non-human identities could first give some early indications. Hygiene evaluation for over exposure, reuse, sharing, non-rotation can also be tracked through these mechanisms like tracking access permissions to sensitive data and looking for duplicate files.

The difference here would be that the responsible team would be the non-human entity owner application team instead of compliance and governance team. To clarify, this will still need to be automated.



The proposed solution here is to have individual application teams to take ownership and visibility of the hygiene evaluations. The security team would get notified only when the hygiene evaluation activity from individual application teams is not reported to security teams. So, in this case, the inaction will be flagged to the security teams instead of abnormal activity. The reason behind proposing this alternate method is based out of empathy for security teams. Security teams simply would not have bandwidth to decipher normal vs. unusual behavior of machine entities. Security teams may not even recognize these machine entities, grasp the nature and impact of the issue, or, know the steps for remediation. For non-human identities, security teams should assume the role of auditors instead of watchers. This is the only way security teams can scale. Burdening them with monitoring behavior anomalies and misconfiguration for non-human identities will be like expecting someone to watch one more screen, who already has twenty screens to look at.

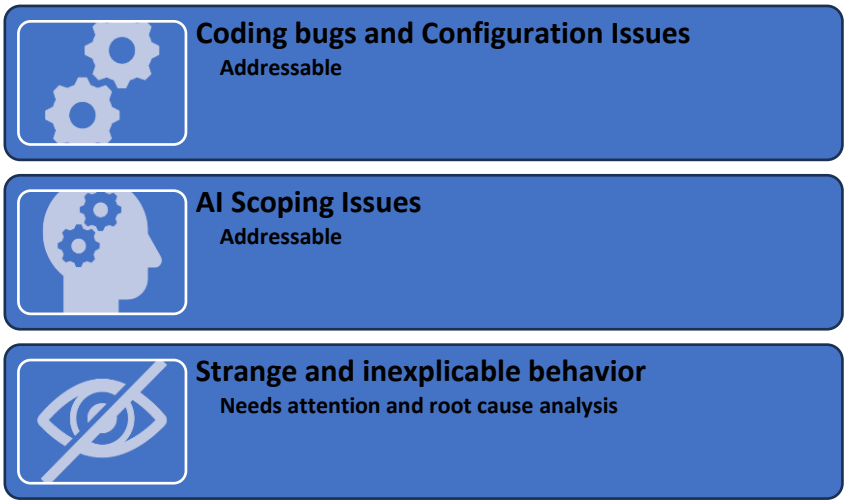
Challenges to Tier-0 segregation approach due to non-human identities

The angle of Tier-0 separation violation due to non-human identities does not appear to be discussed and acknowledge widely. Tier-0 separation revolves around human identities with critical privileges and domain accounts [6]. It did not consider non-human identities broadly. The challenge is also related to fewer number of identity authorities and identity management systems. Before, there were only handful of those like Active Directory, LDAP etc. But now there are more systems and authorities issuing identities. SPIFFE, Secure Production Identity Framework For Everyone, defines constructs needed for workload identities, with SPIRE, SPIFF Runtime Environment instantiates the workload identity mechanism [7]. Thus, there are new control mechanisms like SPIRE server which need to be brought under Tier-0 governance. This is just one example. There are many more authorities, applications and layers which deal with identities and sit at non-Tier-0 levels. The Federated OpenID Connect protocol lets a token generated by regular identity provider to be trusted by different systems or cloud platforms. If the trust is misconfigured, it can lead to misuse. These

challenges require non-human identities to be trusted as security sensitive data and be governed accordingly. Before attempting the surveillance of non-human identities which can quickly get out of hand, there are more fundamental steps required which are achievable. Going down the route of monitoring the non-human identity behavior will only end up creating more alert fatigue for security teams.

GenAI Identities considerations

With Gen AI, there is another type of identity which needs to be considered. First, the human interacting with Gen AI applications like chats. Second, API-based access to Gen AI applications, which can be a human written program, or, it can be another Gen AI application interacting with different Gen AI application using one of the discussed methods in machine identities. Here, the security of first type of identity is not much different than any other application which uses regular methods of authentication and authorization. In fact, for the most part, it would be integrated with identity management systems used by other applications and will carry the same security posture.



However, GenAI agent interactions need to be looked at closely from security point of view. For unintended results returned from a query, or, unexpected actions are performed by the agentic AI, a process to segregate those issues should be put in place. To clarify, this is not the solution for how to secure these identities. Instead, this is preparation in advance, to avoid confusion and chaos when the actual unintended or unexpected behavior is observed. The three categories for grouping these issues could be as follows. Coding bugs and configuration error caused unintended results or behaviors, insufficient scoping caused out of bounds unexpected behaviors and “this is really strange and can’t be explained” issues. The first two kinds of issues will be addressable. The third kind of issues should be rare. It means an independent monitoring system needs to be in place but it needs to extract and classify the results behavior besides just monitoring the AI identities and reporting suspicious behavior. Only then, it would be possible to carry out organized investigation and root cause analysis when such issues are seen.

Conclusion

Explosion of non-human identities is real. However, before applying the modern techniques of behavior monitoring for human identities to non-human identities, there are several considerations. First one is to consider alternatives to reduce shifting the entire burden to security teams. Second is to consider and implement governance principles like Tier-0 separation to non-human identities. Additionally, GenAI cannot be ignored for this aspect but it needs to be handled differently than other non-human identities. Variations with respect to behavior by GenAI identities need to be considered and processes need to be in place before monitoring non-human identities related to GenAI.

REFERENCES:

1. Tak Skyverer, The Service Accounts Guide Part 1: Origin, Types, Pitfalls and *(November 2024)*, https://astrix.security/learn/blog/the-service-accounts-guide-part-1-origin-types-pitfalls-and-fixes/ , (May, 2025)
2. [SentinelOne], What is the Principle of Least Privilege (PoLP)? *(Jan 2025)*, https://www.sentinelone.com/cybersecurity-101/identity-security/what-is-the-principle-of-least-privilege-polp/ , (May 2025)
3. Chris Tozzi, What is Secrets Sprawl & How to Avoid It with Secrets Management, CYBERARK®, *(January 2020), https://developer.cyberark.com/blog/what-is-secrets-sprawl-how-to-avoid-it-with-secrets-management/* , (June, 2025)
4. [Legit Security], *What Are Non-Human Identities? Challenges and Best Practices [ASPM Knowledge Base]*, (June 2025) https://www.legitsecurity.com/aspm-knowledge-base/what-are-non-human-identities , (June, 2025)
5. Dennis Fisher, The hidden threat in your stack: Why non-human identity management is the next cybersecurity frontier, (May 2025), https://claritysecurity.com/clarity-blog/governance-who-should-have-access , (June, 2025)
6. [Clarity Security], Governance for Tier 0/1 Systems - Who should have access?, (Jan 2025), https://www.kdnuggets.com/2018/08/named-entity-recognition-practitioners-guide-nlp-4.html, (June, 2025)
7. [spiffe.io], An overview of the SPIFFE specification, [spiffe.io – CLOUD NATIVE COMPUTING FUNDATION Projects] (November 2019) https://www.cortical.io/static/downloads/email-classification-white-paper-2019.pdf, (December, 2019)