



OFFICE OF THE
STATE AUDITOR

March 26, 2021

The Honorable Sean D. Reyes
Attorney General, State of Utah
Utah State Capitol Building
350 North State Street, Suite 230
SLC, Utah 84114

Re: Limited Review of Banjo

Attorney General Reyes:

Pursuant to your request, the Office of the State Auditor (OSA) has reviewed the Office of the Attorney General's (AGO's) former contract for the Banjo¹ public safety application (Live Time) in response to a Request for Proposal (RFP) for a Public Safety and Emergency Event Monitoring Solution. The circumstances surrounding Live Time posed several complex issues, including concerns regarding privacy and potential algorithmic bias. We recognize the compelling law enforcement interests underlying the associated procurement.

Key Takeaways

- The actual capabilities of Live Time appeared inconsistent with Banjo's claims in its response to the RFP². The AGO should have verified these claims before issuing a significant contract and recommending public safety entities to cooperate and open their systems to Banjo.
- Other competing vendors might have been able to meet the "lower" standard of actual Live Time capabilities, but were not given consideration because the RFP responses were judged based on "claims" rather than actual capability. The touted example of the system assisting in "solving" a simulated child abduction was not validated by the AGO and was simply accepted based on Banjo's representation. In other words, it would appear that the result could have been that of a skilled operator as Live Time lacked the advertised AI technology.

¹ Recently renamed safeXai.

² The RFP was evaluated by a committee comprised of representatives from the AGO, Utah Department of Technology Services (DTS), Utah Department of Transportation (UDOT), and Utah Department of Public Safety (DPS).

- Because of the reduced capability of the Live Time system, it appears much less likely personally-identifiable information (PII) was accessed, transferred, and used than was previously feared.
- The architecture of Live Time’s access to certain public safety systems should not have been permitted based on existing industry best practices.

Methodology

The OSA performed the following:

- Convened the Commission on Protecting Privacy and Preventing Discrimination (Commission). The Commission was intended to advise me regarding the merits of and focus of an in-depth review of Live Time.
- Facilitated Banjo briefing for the Commission. The Commission inquired into Live Time’s capabilities and methodology.
- Performed IT audit limited procedures to understand Banjo’s security processes and controls for personal data.
- Attempted to review Banjo’s financing history but limited public information is available and Banjo did not provide any documentation that demonstrated that its founder was no longer a beneficial owner of any part of the company's stock.

Commission Review of Live Time’s Capabilities

Personally Identifiable Information

Utahns trust the State and its political subdivisions to protect their privacy, particularly as it relates to the collection and storage of PII. Although there is not a single, universally accepted definition of PII, the National Institute of Standards and Technology (NIST)³ provides a thorough definition:

“...any information about an individual that is maintained by an agency, including information that can be used to distinguish or trace an individual’s identity, such as name, social security number, date and place of birth, mother’s maiden name, or biometric records; and any other information that is linked or linkable to an individual, such as medical, educational, financial, and employment information...”

Sensitive PII generally includes legal statistics such as an individual’s full name, driver’s license number, and social security number (SSN). Indirect or non-sensitive PII is data that is often readily available from public sources, such as online directories or social media. Traditionally, law enforcement does not consider physical address or phone number as sensitive PII, but privacy experts generally consider both otherwise.

³ NIST, Information Technology Laboratory Computer Security Resource Center ITL Bulletin for April 2010, Guide to Protecting Personally Identifiable Information, page 1.

<https://csrc.nist.gov/CSRC/media/Publications/Shared/documents/itl-bulletin/itlbul2010-04.pdf>

Live Time’s Capabilities Inconsistent with Banjo RFP Response

On June 12, 2020, Banjo presented the Live Time capabilities. The Commission compared those representations to Banjo’s RFP response^{4 5}. Banjo previously represented the ability for Live Time to perform live “event detection” based on integration of data from outside sources such as social media or private security data. Banjo touted the ability for Live Time to identify child abduction cases, active shooter cases, traffic accidents, event detection, and real-time events. These are similar to representations included in its RFP response to the AGO, which led to a contract⁶. Live Time, as presented to the Commission, appears to be a dashboard of data aggregated from Utah governmental sources, such as 911 dispatch centers, police agencies, and UDOT traffic cameras.

Banjo expressly represented to the Commission that Banjo does not use techniques that meet the industry definition of Artificial Intelligence. Banjo indicated they had an agreement to gather data from Twitter, but there was no evidence of any Twitter data incorporated into Live Time.

Further details of these discrepancies come from Banjo’s (redacted) RFP response⁷ (Response). Many of the claims made in the Response that would have been of most interest to the AGO did not appear to exist in the currently available product.

Inadequate Vetting of Key Personnel

In addition, a significant concern to the Commission was the lack of adequate vetting of key personnel, including Banjo’s founder. The lack of a rigorous background check process is heightened when one has the potential to access sensitive PII as well as the capability to steer law enforcement investigatory resources.

IT Audit Technical Review

At the time of our review, Live Time no longer contained Utah data. As such, our review was limited to processes, procedures, and systems as they existed at the time of this review.

We spoke with the following 911 dispatch centers/public safety answering points (PSAPs):

- Weber 911
- Summit County
- Wasatch County
- Central Utah Dispatch

⁴ SK19019-1 – Banjo Response | REDACTED VERSION, courtesy of State Purchasing

⁵ For example see video of presentation and demonstration to Lehi City Council by Banjo then-employee Brian Smith at https://lehi.granicus.com/player/clip/265?view_id=2&meta_id=53411

⁶ Contract #AR3205, “Public Safety and Emergency Event Monitoring Solution”, awarded June 4, 2019 as well as the prior Sole Source contract #196162, effective November 29, 2018

⁷ SK19019-1 – Page 1 of 2, Banjo 4.1 Response | REDACTED VERSION, courtesy of State Purchasing

- Salt Lake Valley Emergency Communications Center (VECC)

Each PSAP cooperated with Banjo during an initial configuration process. It appears these entities were recommended to participate in the Live Time program. Each entity was required to have a data sharing agreement (DSA) with Banjo before sharing data. Banjo then worked with each entity to gain access to the requested data. Since Banjo is not currently accessing Utah government data, we could not observe this process.

According to PSAP personnel, none of the information law enforcement agencies shared with Banjo was considered sensitive PII.

Based on our limited procedures, other than the concerns identified below, we did not identify other deficiencies in Banjo's processes or controls related to the transfer and storage of government data.

Data Access Concerns

We corroborated the data query process with the listed PSAPs. They reported they could limit the data Banjo queried. The extent of these limitations was unclear and considered a security risk as a result of misconfigured or insufficient security configurations.

From our discussions, Banjo indicated that some PSAPs provided more information than required. When configuring connections, Banjo employees worked with each PSAP to identify specific fields and used their own (Banjo) query as a method to filter out unnecessary information. It is likely that PSAPs provided Banjo access to database tables that contained excess information, possibly including sensitive PII, which should have been restricted.

Banjo would query a PSAP's database over an encrypted Virtual Private Network (VPN) link. 911 queries would happen in real time when new data was made available. Additional data was queried as needed to operate effectively. Even with data query configurations intended to limit Banjo access to data, this direct query by Banjo against the PSAP databases posed a security risk to both PII, as well as PSAP resources.

One significant risk is that Live Time was configured by Banjo personnel to make direct (SQL) database queries to PSAP databases. In theory, Banjo could alter those queries without knowledge of the PSAP. This could have allowed unauthorized access to other sensitive PII.

A second significant concern was the ability for Banjo to directly query these PSAP databases. Although entities we talked to said that they configured "tables" to limit data available to Banjo, it was clear that some agencies were making more information available than was necessary. Our database experts also indicated that it is easy to make a table configuration mistake, or future change, that again makes too much information available.

As an industry best practice, an entity should never allow a third party to make a direct database query against its database. This practice is even more essential in the case of a

database that may include any type of PII. Permitting a third party to have direct database query, as permitted under the agreements with Banjo, exposes the entity and its data to: 1) the risk of misconfigured security leading to inappropriate access to restricted, sensitive or personal data and 2) the risk of misappropriation or theft of data. For misconfigured databases, additional risks include 1) the risk of modification of data or 2) the risk of the injection of malicious code into the database, hijacking the entity's resources for nefarious purposes.

University of California, Berkeley's Information Security Office discusses database hardening best practices⁸ as, "...provid[ing] guidance for securing databases storing sensitive or protected data. Implementing these security controls will help to prevent data loss, leakage, or unauthorized access to your databases." They recommend all application code is reviewed for SQL injection vulnerabilities and that the database server firewall is opened only to specific application or web servers, and firewall rules do not allow direct client access.

A best practice would have been for the PSAP to develop an application programming interface (API), allowing restricted data access with limits controlled by the PSAP. Each PSAP should have an API developed and controlled by that PSAP and hosted within its security perimeter that exposes only permitted data to the third party, which can then "fetch" the permitted data by triggering the API. Three examples of types of APIs are described below.

1. In an online discussion of ways to prevent SQL injection attacks, author Paul Rubens⁹ reminds readers that, "SQL injection is a hacking technique that was discovered more than fifteen years ago and is still proving to be devastatingly effective today, remaining a top database security priority." His key suggestions include, "Don't use dynamic SQL – don't construct queries with user input." Instead, he suggests using what we refer to as an API, that is, "...prepared statements, parameterized queries or stored procedures instead whenever possible." He ends with this counsel, "Trust no one: Assume all user-submitted data is evil so use input validation via a function." Clearly the Banjo relationship required that Utah citizens and government trust Banjo.
2. Other experts recommend an even more robust "three tier system"¹⁰ where "The three-tier model effectively isolates the end user from the database by introducing a middle-tier server. This server accepts requests from clients, evaluates them and sends them on to the database server for processing. The database server sends the data back to the middle-tier server, which then sends the data to the client system."

⁸ Berkeley Information Security Office - Database Hardening Best Practices
<https://security.berkeley.edu/education-awareness/best-practices-how-tos/system-application-security/database-hardening-best>

⁹ How to Prevent SQL Injection Attacks, By Paul Rubens, Posted May 2, 2018

<https://www.esecurityplanet.com/threats/how-to-prevent-sql-injection-attacks.html>

¹⁰ Top 14 Data Security Best Practices, https://www.netwrix.com/data_security_best_practices.html

3. And finally, industry experts recommend¹¹ the Doctrine of Least Access, meaning that if you don't need to work with it, you shouldn't have access to it. This includes limiting access for all IT personnel. The Enterprise Systems Journal closes with this concerning summary, "According to one poll of almost 650 IT professionals conducted last year, 10 percent admit to regularly abusing their security privileges and inappropriately accessing corporate data."

Clearly, Live Time's configuration lacked certain key security features and Banjo's approach didn't follow best practices.

Data Transfer Concerns

In addition, Banjo established and controlled the security keys to the VPN between Banjo's systems and those of the PSAPs. Establishing a VPN with a third party could essentially tie together the third party's network to the public entity's network, exposing the public entity to any attacks or weaknesses in the third party's network. This is a significant breach of network security best practices as further described below.

To gain access to PSAP data, Banjo worked with each agency to establish a secure VPN, which Banjo established and controlled the keys to. This allowed Banjo to access police agency databases using custom SQL queries. This query happened in real-time on an agreed-upon "cadence." This opens significant security and privacy risks, particularly as Banjo noted that police agencies had "made available" to Banjo more data than Banjo required for operation. Data in transit over the VPN was encrypted.

The Banjo configuration, where Banjo was granted access by a VPN connection directly into the networks of some of these PSAPs is a clear violation of this guidance to "not allow direct client access." Again, a VPN directly ties two networks together. Even more concerning, Banjo issued and controlled the VPN keys. Although the agencies could remove the keys from their firewall to terminate Banjo access, a more secure model is for the agency to issue and control VPN keys, in the few cases where that is appropriate.

Data Storage

Banjo used Amazon Web Services (AWS) to store and secure its data. All data stored on the AWS servers appeared appropriately restricted to authorized employees. Banjo provided a screenshot of the AWS control panel showing the list of users authorized to access the data. Access required two-factor authentication for log-in. Banjo represented that their policy is that no customer data is allowed to be stored on personal workstations or laptops. However, we could not verify this because they do not have any Utah data.

¹¹ Enterprise Systems Journal , Access Control: 10 Best Practices, <https://esj.com/articles/2007/03/27/access-control-10-best-practices.aspx>

Banjo relies on AWS to secure data at rest. Ernst & Young performed an audit of the AWS infrastructure and generated an audit report referred to as a “SOC 2 Type 2” report for AWS, which determined that the design, implementation, and operation of data security controls are effective. Per this report, all data at rest is encrypted using AES-256 encryption.

Banjo claimed that data is stored on a temporary basis. Data is stored in folders on AWS, which are configured with auto-deletion parameters. AWS automatically deletes the data as scheduled. Banjo configures the folders to delete the data as desired. We reviewed the current deletion configuration settings, which showed data, configured to be deleted within 36 hours. Banjo claimed that most data is stored for no more than 24 hours, but that some data is stored for up to one week.

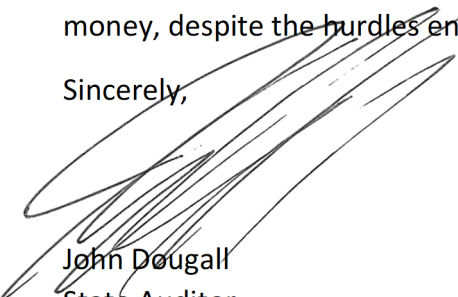
Conclusion

The Commission recently issued:

- Software Application Procurement Principles for Utah Government Entities
- Questions from the Commission on Protecting Privacy and Preventing Discrimination

These documents are intended to help State agencies and Utah government entities evaluate advanced technologies, particularly those that might place PII at risk or could embed unintended bias into government systems. These Principles will help government entities avoid the procurement pitfall experienced by the AGO. Following these Principles might have helped the AGO, DSP, and the UOU more appropriately assess both Live Time’s purported capabilities as well as identifying weaknesses in Banjo’s system design approach. I encourage you to use these Principles as your team assesses potential new advanced technology. Also, I would encourage you to encourage other government entities to embrace these Principles as they considering procuring sophisticated software applications leveraging advanced technologies. I recognize the challenges of finding innovative ways to better perform critical government functions. I also understand the complexities of translating one’s vision into operational tools. I commend government entities who continually strive to strengthen operations while saving money, despite the hurdles encountered pursuing those objectives.

Sincerely,



John Dougall
State Auditor

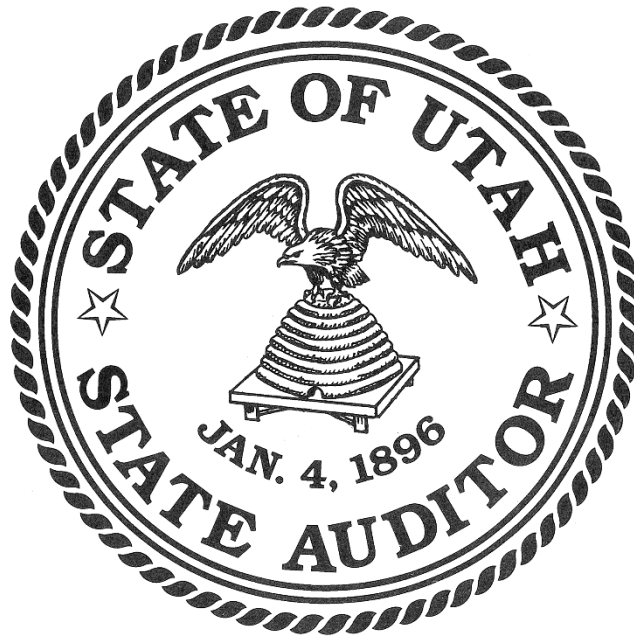
Attachments:

Software Application Procurement Principles for Utah Government Entities
Questions from the Commission on Protecting Privacy and Preventing Discrimination

Software Application Procurement Principles for Utah Government Entities

Commission on Protecting Privacy and Preventing Discrimination

Office of the State Auditor
Issued February 1, 2021



OFFICE OF THE
STATE AUDITOR

AUDIT LEADERSHIP:

John Dougall, State Auditor
The Commission on Protecting Privacy and
Preventing Discrimination



OFFICE OF THE
STATE AUDITOR

**Software Application Procurement Principles
for Utah Government Entities**

Commission on Protecting Privacy and Preventing Discrimination

**Office of the State Auditor
Issued February 1, 2021**

We recommend government entities apply the following principles as they procure commercial software applications or engage in development of custom software applications that use, gather or consolidate personally-identifiable information (PII) or other sensitive data. These principles are most suited to new or emerging technologies, such as artificial intelligence (AI) or machine learning (ML), that may not have a long history to draw upon for software application and vendor evaluation as well as for “startup” or young vendors that likewise may not have extensive history.

1. **Limit Sharing of Sensitive Data:** Government entities should fully understand their data. They should limit sharing of sensitive data (private data, PII, etc.) to the greatest extent possible to protect individual privacy and should not share more than is necessary to perform the required task. Data should be filtered and restricted within the government’s systems before being transferred into the vendor’s application. Wherever possible, a government entity should anonymize data, but government entities should recognize that sensitive data can be reconstructed from previously anonymized sources.
2. **Minimize Sensitive Data Collection and Accumulation:** A software application should collect no more sensitive data than it needs, and should retain that sensitive data no longer than it needs to accomplish the specified purpose.
3. **Validate Technology Claims - including Capability Review:** A vendor should clearly demonstrate the validity of their marketing claims. Example claims that warrant particular caution include
 - a. Asserted use of AI or ML,
 - b. Proposed use of disparate data sources, especially social media or integration of government and private sources, and
 - c. Real-time capabilities, especially real-time data intelligence or analysis.

4. **Rely on Objective, Repeatable Metrics:** Vendors make various claims about the ability of their software applications to deliver value within a given accuracy or efficiency measure. Do not rely on anecdotes as validation of these claims. Government entities should invest in software applications where the value can be measured on an ongoing basis. A reputable vendor should include success criteria in any Request For Proposal (RFP) response, and these should include metrics that are easy to measure and compare across time and vendors. The RFP should also request that work to automate the gathering and reporting of these metrics be included in the project definition.
5. **Assess Threat Models:** The vendor should be able to enumerate the people, processes, and technological interfaces that constitute an attack or risk surface for their proposed software solution. These threats should be prioritized, and high-priority threats should have recommended mitigations. The vendor should have a vulnerability reporting process. A documented history of conducting and remediating penetration tests is a significant benefit.
6. **Perform In-Depth Review of AI/ML Algorithms:** All claims of AI or ML should be clearly validated and explained, including:
 - a. AI algorithms used in the software application
 - b. model training and evaluation procedures used
 - c. model outputs utilized by product / feature
 - d. source and composition of the training, validation and test datasets
 - e. demonstration of having a license to use those datasets
 - f. pre- and post-processing methods applied to data
 - g. processes for model lifecycle management, including on-going evaluation

The output of an AI-based software application may still have issues of bias or lack of fairness, even if the inputs and system are reasonably judged not to include such failings. The output of the software application should be monitored to ensure protection of privacy and avoidance of improper or unexpected bias.

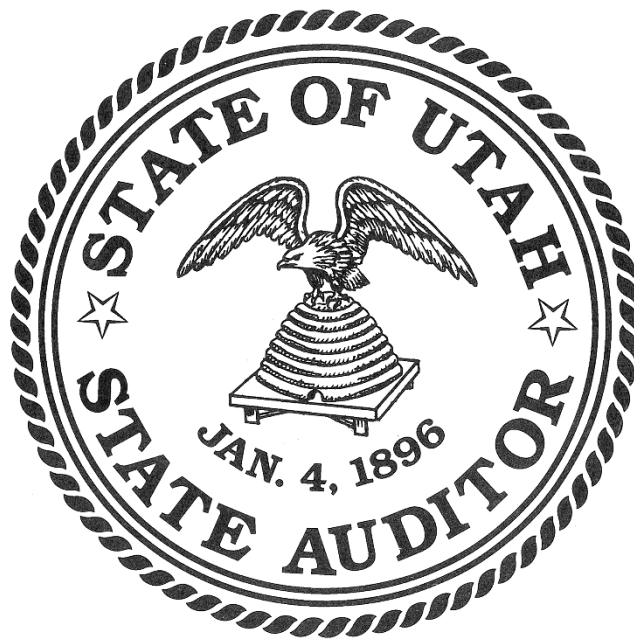
7. **Demonstrate Privacy Compliance: Privacy Specific Items and Protection:** The vendor should demonstrate compliance with privacy regulations, such as [CCPA](#) or any similar laws enacted by Utah.
 - a. The vendor should define what constitutes PII under the contract. A government entity's default definition may be overly narrow and may need adjustment for particular problems.
 - b. The vendor should describe their anonymization process and how it protects against the use of secondary data to de-anonymize data. A governmental entity should evaluate the effectiveness of that process to mitigate de-anonymization in the context of the software application.

- c. Specific certifications may be required for a specific application, but such certifications may still be insufficient to protect privacy.
8. **Review Steps Taken to Mitigate Discrimination:** Ensure that the vendor has considered the question of bias and discrimination within their software application and that the vendor has mechanisms, such as audit results, to demonstrate that their software application does not disproportionately affect various categories of individuals, particularly any federally protected class.
- a. For example, consider sources of data that may include implicit or historic bias (e.g., distribution of video cameras by region or neighborhood)
 - b. For example, consider how model choice and training may introduce bias.
 - c. Understand the interpretation of model output.
 - d. For AI-based or ML-based software applications, determine whether the source of training and model data has been evaluated by subject matter experts.
 - e. Entity should use best in class models for evaluation to prevent discrimination, particularly in the case of biometric analysis or facial recognition. As an example, the [U.S. NIST](#) provides evaluations of the accuracy of facial recognition based on demographic differences.
9. **Determine Ongoing Validation Procedures:** The government entity must have a plan to oversee the vendor and vendor's solution to ensure the protection of privacy and the prevention of discrimination, especially as new features/capabilities are included.
10. **Require Vendor to Obtain Consent of Individuals Contained Within Training Datasets:** Many biometric characteristics may be captured without an individual's knowledge or consent. Examples may include facial recognition or gait analysis. Ensure that a vendor has consent from the individuals whose biometric characteristics are used in training datasets.
11. **Vet Key Vendor Personnel:** Key vendor personnel may need background checks. The type of checks may vary depending upon the sensitivity of data that personnel have access to. These checks need to be validated to the procuring government entity.
12. **Evaluate Vendor Corporate Management and Vendor Solvency:** Evaluate the financing history of the vendor and their solvency to ensure they can carry out the contract. Placing code in escrow may be a compensating mechanism here.

Questions from the Commission on Protecting Privacy and Preventing Discrimination

Companion to Software Application
Procurement Principles
for Utah Government Entities

Issued February 1, 2021



OFFICE OF THE
STATE AUDITOR

AUDIT LEADERSHIP:

John Dougall, State Auditor
The Commission on Protecting Privacy and
Preventing Discrimination



OFFICE OF THE
STATE AUDITOR

**Questions from the Commission on
Protecting Privacy and Preventing Discrimination**

**Companion to Software Application Procurement Principles
for Utah Government Entities**

**Office of the State Auditor
Issued February 1, 2021**

HOW TO USE THIS DOCUMENT

This document is intended to help government entities within the State of Utah with their procurement of advanced technologies that have the potential to impair the privacy of, or lead to discrimination against, Utahns. The following list of questions and queries are companions to the referenced Principles document. Asking these questions should help government entities improve their evaluation process as they evaluate software vendor offerings and as they issue RFPs or procure commercial software products.

Consider including the questions or queries from the appropriate sections in your RFP and/or RFI and in appropriate proposal scorecards.

These are complex topics and may require subject matter expertise to effectively evaluate proposals. It is suggested that the public entity seek out relevant experts to pose these questions and evaluate the answers.

KEY STEPS

- Identify which sections apply to your particular application. For example, does the application or solution being considered have a potential impact on the privacy of Utahns?
- Could information used or gathered by the solution impact equity and inclusion of Utahns or lead to discrimination?
- Does the vendor use words or phrases that reference artificial intelligence, machine learning, computer vision, surveillance, recording or similar terms?

If so, review the sections below and identify those that apply to your solution.

DEFINITIONS

Artificial Intelligence (AI) - The theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages. (Oxford Languages via Google Dictionary)

Machine Learning (ML) - A collection of methods used to learn an algorithm for a task from historical data about that task.

Natural-Language Processing (NLP) - Broadly defined as the automatic manipulation and making sense of natural language, like text, by software.
(<https://machinelearningmastery.com/natural-language-processing/>).

Computer Vision - An interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos.

Public Entity or Government Entity - A state agency, office, board, or commission or a local government entity such as county, city or service district or any political subdivision of the State of Utah, including K-12 education and higher education entities.

PII - Personally Identifiable Information - Information that may lead to an individual being identified. Although there is not a single, universally accepted definition of PII, National Institute of Standards and Technology (NIST) provides a good working definition.¹ “Personally identifiable information (PII) is any information about an individual that is maintained by an agency, including information that can be used to distinguish or trace an individual’s identity, such as name, social security number, date and place of birth, mother’s maiden name, or biometric records; and any other information that is linked or linkable to an individual, such as medical, educational, financial, and employment information” (based on General Accountability Office and Office of Management and Budget definitions).

¹ NIST, Information Technology Laboratory - COMPUTER SECURITY RESOURCE CENTER
ITL BULLETIN FOR APRIL 2010, GUIDE TO PROTECTING PERSONALLY IDENTIFIABLE INFORMATION,
NIST - <https://csrc.nist.gov/CSRC/media/Publications/Shared/documents/itl-bulletin/itlbul2010-04.pdf>

1. LIMIT SHARING OF SENSITIVE DATA

- 1.1. What data elements are we considering sharing? Which elements are PII?
- 1.2. What data is the minimum set of data required to accomplish the task?
- 1.3. How personal and sensitive is that data?
- 1.4. Do we need to share that data?
- 1.5. Do any statutory or policy restrictions or obligations exist? Does the potential recipient understand and agree to these restrictions or obligations? If statutory or policy restrictions or obligations are unclear or non-existent, then the entity should default to caution.
- 1.6. What expectation of privacy do data subjects have of this data? (e.g. Do Utahns who receive a driver's license expect that data to be used for other purposes?)
- 1.7. What is the recipient's data retention policy associated with the data? How do we test/verify that they are in compliance?
- 1.8. How long will the data sharing agreement be in place?
- 1.9. Who is authorized to share this data (organizational authority)?

2. MINIMIZE SENSITIVE DATA COLLECTION AND ACCUMULATION

- 2.1. When collecting this data, have we been transparent about how the data will be used? Is there an expectation of privacy on the part of the data subjects?
- 2.2. Do any statutory or policy restrictions or obligations exist when collecting this data?
 - 2.2.1 Does everyone in our organization who has access to this data understand and agree to these restrictions or obligations?
 - 2.2.2 If statutory or policy restrictions or obligations are unclear or non-existent, then the government entity should default to caution.
- 2.3. If the data we have collected is shared with a vendor, is there any ability for a vendor, its employees, or its affiliate to download/save data, preserving it beyond normal/approved data retention limits?
 - 2.3.1 If yes, are such events recorded to create an audit trail? How serious is this risk? How serious are the consequences?
- 2.4. Clearly delineate what data the vendor is receiving and specify the appropriate use of that data.
 - 2.4.1 Have the vendor clearly describe its data retention policies for all key data types as well as for "metadata" (data derived from source data such as video, audio, biometrics, etc.).
- 2.5. See also section 5.2 Access Provisioning and Management.

3. VALIDATE TECHNOLOGY CLAIMS - INCLUDING CAPABILITY REVIEW

- 3.1 How will vendor's marketing claims be validated by the government entity during the RFP process and on an ongoing basis? Example claims that warrant particular caution include:
 - 3.1.1 Asserted use of AI/ML
 - 3.1.2 Proposed use or integration of disparate data sources, including commercially available data, social media or integration of government and private sources
 - 3.1.3 Real-time capabilities, especially real-time data intelligence or analysis.
 - 3.1.4 Surveillance activities or systems
 - 3.1.5 See Section 6. Perform In-Depth Review of AI/ML Algorithms:

4. RELY ON OBJECTIVE, REPEATABLE METRICS

- 4.1 Are we clear about the problem that we intend for this system or solution to solve for our entity and for our constituents?
- 4.2 How do we currently solve this problem? What are the strengths and weaknesses of our current solution?
- 4.3 How do we measure the performance of our current solution? Are we comfortable with sharing our metrics? If not, why not?
- 4.4 What ongoing metrics will we use to evaluate the effectiveness of this new solution over time? Which of these metrics will the vendor provide and which will the government entity obtain independently?
- 4.5 How will we decide if the proposed solution is not performing to our expectations or to the expectations of our constituents?

5. ASSESS THREAT MODELS.

System Architecture and Practices to Protect Against Internal or External Malicious Actors or Attacks

- 5.1 What policies and safeguards exist to protect against a third party actor from intercepting sensitive data?
- 5.2 Require Vendor to explain their Access Provisioning and Management processes related to:
 - 5.2.1 Authorization of access
 - 5.2.2 Provisioning

- Does the vendor operate on a least privilege model (minimum required access to accomplish the task)?
- Does the vendor have role-based access with hierarchical permissions?
- What is the vendor's criteria for authorizing administrative/back-end access?

5.2.3 Access Reviews/Auditing

- Does the vendor perform regular access validations and reviews?
 - If yes, what is the review cadence or timing?
- What auditing/logging exists for end-user system activity?
- Does the vendor perform audit/log reviews?
 - If yes, what is the review timing and frequency?

5.2.4 Access Revocation

- How is access removed and under what circumstances?

- 5.3. What prevents someone from deliberately introducing bad data (or deliberately skewing the timing or amount of input) to "game" the system or distort outcomes?

General Best Practices in System Maintenance

- 5.4. Have the vendor describe its baseline system maintenance processes and procedures, describing any third party providers (such as cloud providers or open source solutions) that they use in their system.

Key elements include:

- 5.4.1 System hardening
- 5.4.2 Security standards and key management
- 5.4.3 Data link security
- 5.4.4 Backup and Recovery
- 5.4.5 Third party vendors including license compliance
- 5.4.6 Upgrade, patch, and vulnerability management processes

- 5.5 Has the vendor had a security breach of its system in the last 5 years? If so, what data was involved and what were the mitigation actions? If so, was there a requirement for regulatory or customer notification?

- 5.5.1 Have there ever been any simulations to practice a response to a security breach?

6. PERFORM AN IN-DEPTH REVIEW OF AI/ML ALGORITHMS

Many vendors promise that the “magic” of AI/ML will solve complex problems. These technologies are just emerging and generally only specialists are familiar with the details. At the same time, AI/ML can be used poorly, or even abused, and a public entity must evaluate the potential impact of these technologies.

- 6.1 Have the vendor provide data that validates the claims of their AI/ML system’s performance? Where possible, external data or validation should be used.
- 6.2 Have the vendor provide an overview of the AI algorithms used in the software application and the rationale behind the choice of those algorithms. It is important to remember that AI algorithm choice is closely related to the problem at hand. There is no “single best” AI solution for all problems.
- 6.3 Have the vendor identify how the outputs of the model are used in different parts of the solution.
- 6.4 Have the vendor explain the model training and evaluation procedures used in developing the system.
- 6.5 Have the vendor identify the source and composition of the training, validation, and test datasets. For example, are these commercially available datasets (identify source) or privately developed or proprietary datasets?
- 6.6 Have vendor demonstrate that it has a license to use datasets.
- 6.7 Have vendor demonstrate whether the source of training and model data has been evaluated by subject matter experts?
- 6.8 Have vendor explain pre- and post-processing methods applied to data processes?
- 6.9 Have vendor explain model lifecycle management, including on-going evaluation
 - 6.9.1 Have vendor discuss whether their model has evolved over time. Does it evolve with system use? How is this managed?
- 6.10 What data does the vendor have about false positives and false negatives?
- 6.11 See Section 8 for details on preventing discrimination in the use of AI/ML techniques.

7. DEMONSTRATE PRIVACY COMPLIANCE: PRIVACY SPECIFIC ITEMS AND PROTECTION

- 7.1 What evidence does the vendor have, and can provide to the government entity, that the privacy protection model is accurate?

- 7.2 Contracts with government entities sometimes define PII very narrowly (e.g., government-issued identification numbers, financial account information, and personal health information). Have the vendor define PII under the contract. The government entity should also assess whether there are other elements that the entity or the public also consider as PII. Highly sensitive items might be address, phone number, email address, social media handles, demographic information etc. (see broader definition of PII at beginning of document)
 - 7.2.1 What PII is protected in interactions with this vendor? Which elements of PII are removed from which data sources?
 - 7.2.2 If PII about the target person is protected, what about others who might be identified? What about EXIF tags/geotags on images, for example? How is this “non PII” data stored and protected?
- 7.3 Describe the process for the government entity to verify the vendor's safeguarding of data.
 - 7.3.1 Have any such checks been performed? If so, who performed the verification and how often are they performed? Is this cost built into the contract?
- 7.4. Have the vendor describe the process by which the vendor anonymizes PII and other relevant data. Where is this done and what are the protections and validations of this process?
 - 7.4.1 Has the vendor considered the problem of de-anonymization of data (see for example, <https://www.eff.org/deeplinks/2009/09/what-information-personally-identifiable>)
- 7.5 What certifications does the application require and what limitations do these certifications have?

8. REVIEW STEPS TAKEN TO MITIGATE THE POTENTIAL FOR DISCRIMINATION

- 8.1 What evidence does the vendor have, and can provide to the public entity, that the preventing discrimination model is accurate. For example, has a sensitivity analysis (routinely performed on deterministic systems) been performed to know what types of inputs this system is sensitive to?
- 8.2 What steps has the vendor taken to consider the question of introducing bias that might lead to discrimination within their software application? What mechanisms, such as audit results, does the vendor have to demonstrate that their software application does not disproportionately affect various categories of individuals, particularly any legally protected class (federal, state, or local).

- 8.3 Does the vendor's system use, rely on, or consider any sources of data that may include implicit or historic bias (e.g., distribution of video cameras by region or neighborhood)? What analysis has the vendor performed?
- 8.4 Has the vendor evaluated how model choice and training may introduce bias? Have the vendor share such evaluations and conclusions.
- 8.5 What type of interpretation does the vendor apply to the model output? How might this introduce bias?
- 8.6 What models has the vendor used to evaluate the risk of discrimination, particularly in the case of biometric analysis or facial recognition?
 - 8.6.1 As an example, NIST provides evaluations of the accuracy of facial recognition based on demographic differences. If such a third party evaluation is available for the vendor's application, has the vendor had its application tested? What are the results?

Note: The output of an AI-based software application may still have issues of bias or lack of fairness, even if the inputs and system are reasonably judged not to include such failings. The output of the software application should be monitored to ensure protection of privacy and avoidance of improper or unexpected bias.
- 8.7 See also section 6. Perform In-Depth Review of AI/ML Algorithms.

9. DETERMINE ONGOING VALIDATION PROCEDURES TO DEMONSTRATE DATA HANDLING TO PROTECT PRIVACY AND PREVENT DISCRIMINATION

- 9.1 Are the vendor's data handling practices self-asserted, or are they audited independently? If self-asserted, then the government entity should evaluate the data handling practices. If audited, the government entity should review audit results.
- 9.2 Is there any mechanism for auditing the quality of anonymization of sensitive data? What is this mechanism? Who conducts these audits? How often are such audits conducted?
- 9.3 If a particular vendor's system is used by multiple public entities in the State, are all state agencies or public entities that provide or contribute data allowed to see/use one another's data, or is the aggregate view of all these sources illegal or disallowed? Is it or can it be prohibited by contract?
 - 9.3.1 What mechanism exists at the government level to mediate these issues or concerns?

10. REQUIRE VENDOR TO OBTAIN CONSENT OF INDIVIDUALS CONTAINED WITHIN TRAINING DATASETS

10.1 Does vendor have the permission of every individual whose information is contained within its training, validation and test datasets?

10.1.1 Is there any risk that data in its dataset(s) has been “scraped” or otherwise gathered from online sources without the permission of those whose information is included and/or without permission of the owners (who otherwise have permission to use the data)? Have vendor provide credible confirmation of these permissions.

11. VET KEY VENDOR PERSONNEL

The relevance of this section is dependent upon the sensitivity of the data and the implementation model of the technology.

Employee and Founder Background Checks

11.1 Has the vendor conducted intensive background checks of all key employees and founders? What are the factors that would result in employment disqualification?

11.1.1 Are background checks updated on a regular basis? How often?

11.1.2 If key founders or employees have elements in their background that are not consistent with protecting privacy and preventing discrimination, what is our process to address these issues with the vendor?

11.2 Have the vendor provide the “standard” (i.e., template) used by their background check provider.

Vetting Partners and Related Entities

11.3 What other partners help the vendor acquire needed data for the desired capabilities?

11.3.1 Vendor should provide a detailed list of all government entity and private entity partners, along with information on the relevant data sharing agreements. For example, some AI companies have “scraped” social media sites to build their image databases, in violation of social media agreements. All data must be acquired with full permission/cooperation of the source organizations.

11.3.2 See Section 10.

11.4 If vendor is working with or planning to work with multiple government entities within Utah, a continuously updated master list of all such relationships must be provided to all entities and also to State Purchasing and must be readily available.

11.4.1. "Working with" includes both customer/vendor agreements as well as any data sharing or data access agreements.

12. EVALUATE VENDOR CORPORATE MANAGEMENT, SOLVENCY, AND TRANSPARENCY

12.1 It is often the case with "newer" technology solutions that the company is a startup, privately held, or not yet profitable. Have the vendor provide financial statements and, if the company is not profitable, have vendor provide their financing plan, including the current state of investor commitment.

12.1.1. Review vendor's financial statements and financing history including independently gathering data on the company from "deal reporting" sources such as Crunchbase, Pitchbook, etc. If the company is a startup, privately held, or not profitable, an expert in startup capital should review the cap tables (capitalization tables including details of shareholders and their holdings), and "deal docs" or financing documents to ensure that there are no "poison pills," and no single shareholder that could terminate the company, as well as to ensure financial solvency during the contract period.

12.2 Larger companies should provide written assurances that the proposed product or service will be supported for the proposed project duration. Larger companies are frequently cutting off products and services even when contracts have been signed, leaving government entities without a solution, or required to pay for a new, essentially duplicate, implementation.

12.3 Ensure that a thorough background check has been conducted on each key member of the management team, including key system and technology architects to avoid contracting with a firm who may have malicious actors within the company.

12.3.1 See also Section 11. Vet Key Vendor Personnel

12.4 Conduct a general information search on the company to discover concerns that have been identified.

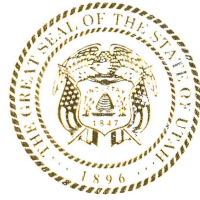
12.5 Is the vendor transparent about its use of data? For example, are there ways for members of the public to provide feedback and express concerns about the technology, data, and compliance to the governmental entity and/or vendor?

Contract Provisions

12.6 Some contracts include sections that provide for the addition of new modules, components, and other supplementary items being added to the contract after contract signature. The vendor should be required to notify the entity of proposed changes.

12.6.1 All elements of this document may be impacted by the addition of new features or capabilities and so the vendor should be required to address the elements of this document with each new version or upgrade.

STATE OF UTAH
OFFICE OF THE ATTORNEY GENERAL



SEAN D. REYES
ATTORNEY GENERAL

Spencer E. Austin
Chief Criminal Deputy

Ric Cantrell
Chief of Staff

Melissa A. Holyoak
Solicitor General

Brian L. Tarbet
Chief Civil Deputy

March 26, 2021

Dear Auditor Dougall and the Commission on Protecting Privacy and Preventing Discrimination,

Thank you for accepting the request of the Utah Attorney General's Office ("UAGO") to explore the bias and privacy questions regarding the State's previously contemplated use of Live Time technology as an investigation tool, and for going even beyond that request to provide state agencies with expert guidelines for the procurement process.

We are encouraged by your findings and feel validated that neither privacy intrusion nor racial or religious bias was inherent in the Banjo Live Time system. Your findings align with our experience regarding this company, its founder, priorities, work product, and ethics. We observed, and you have confirmed, that sensitive PII was never shared with Banjo. That protection was always a high priority for this office.

We further commend the work of the Commission, specifically in producing two excellent documents regarding procurement principles and questions. We will encourage state agencies to include the applicable questions in RFPs, RFIs, and proposal scorecards when evaluating new technology.

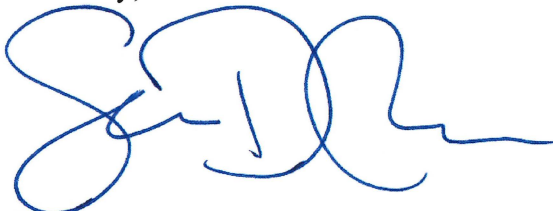
It is worth noting that some of these recommended protections were already integrated in the State's Live Time contract at Banjo's request. We appreciate the Commission's warning and insight on the possible exposure risks when combining government data with private systems.

The UAGO is convinced that if it were possible, a deeper study would reveal an even more accurate and indicative analysis of the capabilities and protections of the system in question. However, we recognize certain limitations to this review in terms of time, scope, and inaccessibility of particular witnesses and data. To wit, company leadership has changed, contracts have been suspended, and contemporaneous information was no longer available for review. Throughout the contracting and fulfillment process, it was well understood by both Banjo and the State that more privacy-safe inputs would come online before we realized the full capabilities of the Live Time system. While some preliminary dashboards were provided, Banjo was still in the system building, data-intake, programming, and test phases when the contract was suspended.

We wish to address the Commission's concern regarding the vetting of Banjo's founder and former CEO, Damien Patton. As a preliminary matter, in RFPs such as the one in which Banjo participated, there is no requirement in the state procurement process for the UAGO to investigate companies and particularly not their employees. The UAGO, however, went above and beyond what is normally done for contractors including conducting interviews with colleagues, technology experts, leaders of other companies familiar with the CEO, law enforcement officials, elected officials, etc. The subsequent negative information that came out about Mr. Patton was contained in records that were sealed and/or would not have been available in a robust criminal background check. Based on our first-hand experience and close observation, we are convinced the horrible mistakes of the founder's youth never carried over in any malevolent way to Banjo, his other initiatives, attitudes, or character.

The Attorney General's Office is committed, as always, to protect all the rights of Utah citizens. This includes civil liberties as well as freedom from predatory violence and other crimes. I will continue to use cutting-edge technology to keep Utah safe and will do so within the guidelines recommended by this Commission. Thank you for your careful attention.

Sincerely,

A handwritten signature in blue ink, appearing to read 'S. Reyes', with a stylized flourish extending to the right.

Sean D. Reyes
Utah Attorney General