TripleBlind™

EXECUTIVE SUMMARY

# DR. BRADLEY MALIN (PHD)'S **ASSESSMENT OF TRIPLEBLIND**

# INTRODUCTION

Dr. Bradley Malin, Accenture Professor of Biomedical Informatics, Biostatistics, and Computer Science, as well as Vice Chair for Research Affairs in the Department of Biomedical Informatics at Vanderbilt University, is one of a handful of independent experts with the authority to review technologies and determine whether or not they meet HIPAA compliance. While a majority of determinations focus on a singular dataset, Dr. Malin's expert determination method covers the entire process of de-identification.

## WORK ASSESSED

Dr. Malin assessed multiple elements of the TripleBlind product.

1. Deployment of existing models on data that may contain PHI through TripleBlind
2. TripleBlind's Blind Learning and query solutions to train models and attain insights, both against one entity and multiple entities' PHI containing data
3. TripleBlind's ability to support the above activities without the need for de-identifying data prior to engaging with TripleBlind (i.e. the ability to enable PHI-containing data including unstructured data through TripleBlind)

## FINDINGS

1. There is no need to "de-identify" data prior to deploying TripleBlind. De-identification occurs at the speed of installation. Consequently, this allows for a continuous stream of new data to be instantly made available for collaboration, rather than requiring processing to remove PHI in advance.
2. Any and all model types, including but not limited to statistics, queries, and prediction models including neural networks can be trained and run on PHI containing data safely given the technological approach used by TripleBlind. The opinion is not limited to tabular or structured data but broadly applies to unstructured data as well (images, etc)

As a result, data nodes can be brought onto TripleBlind within days to weeks of them being ready to be onboarded.

This also means any data user can do any type of analytic work on data from anywhere while remaining fully compliant with HIPAA and variable state by state regulations.

# CONCLUSIONS

TripleBlind's current architecture, cryptographic approach, and technological makeup meets de-identification requirements as set forth by the HIPAA Privacy Rule and meets the requirements of all state-level data protection acts. Dr. Bradley Malin's de-identification opinion states that use of TripleBlind against PHI-containing data holds in any and all US jurisdictions.

Consequently, TripleBlind's product can directly "touch" a data lake containing PHI data, without the need for pre-de-identification. TripleBlind's software renders data de-identified for the purpose of analytic and data collaboration operations, such as conducting statistical analyses, training prediction models, and deploying prediction models.

# KEY QUOTES

**Note:** MAP/DAP refers to the access points. MAP is the access point installed with the "data user", while DAP is the access point installed with the "data owner".

## Regarding deployment of an existing model

> "... it is evident that at no point does DAP, MAP, or TripleBlind have access to any information that, by itself, would communicate directly, or through inference, any raw data ... It should further be recognized that the data could include what would traditionally be considered identifying, or potentially identifying (e.g., date of birth), information.
>
> Under the way that the data is processed, this information would not be disclosed in the multiparty prediction protocol."

## Regarding Blind Learning

> "From a privacy perspective, the important part to recognize is that the information passed from the data holder to the algorithm learner is, in effect, a compression of the data.... Even in the event that a MAP could bound the values in the records of the underlying resource maintained by the DAP, the MAP is still not able to discern any specific underlying record with sufficient certainty to render the data into an identifiable state. Similar to the prediction setting ..., the blind learning process could include identifying or potentially identifying information in the inputs."

## Regarding Blind Learning against multiple parties

> "As a result, it is evident that this scenario is no less secure than the one party split learning scenario. This is because the information contributed by one DAP does not directly insinuate what the data in the other DAPs corresponds to. As a result, since the one party learning scenario is considered to be very small risk of disclosure of the DAP records, so too is the multiple party scenario."