# Privately Evaluating Untrusted Black-Box Functions

Ephraim Linder* Sofya Raskhodnikova* Adam Smith*† Thomas Steinke†

August 9, 2024

**Abstract**

We provide tools for sharing sensitive data in situations when the data curator does not know in advance what questions an (untrusted) analyst might want to ask about the data. The analyst can specify a program that they want the curator to run on the dataset. We model the program as a black-box function $f$. We study differentially private algorithms, called *privacy wrappers*, that, given black-box access to a real-valued function $f$ and a sensitive dataset $x$, output an accurate approximation to $f(x)$. The dataset $x$ is modeled as a finite subset of a possibly infinite set $\mathcal{U}$, in which each entry $x$ represents data of one individual. A privacy wrapper calls $f$ on the dataset $x$ and on some subsets of $x$ and returns either an approximation to $f(x)$ or a nonresponse symbol $\perp$. The wrapper may also use additional information (that is, parameters) provided by the analyst, but differential privacy is required for *all* values of these parameters. Correct setting of these parameters will ensure better accuracy of the privacy wrapper. The bottleneck in the running time of our privacy wrappers is the number of calls to $f$, which we refer to as *queries*. Our goal is to design privacy wrappers with high accuracy and small query complexity.

We introduce a novel setting, called the *automated sensitivity detection* setting, where the analyst supplies only the black-box function $f$ and the intended (finite) range of $f$. In contrast, in the previously considered setting, which we refer to as the *claimed sensitivity bound* setting, the analyst also supplies additional parameters that describe *the sensitivity of* $f$. We design privacy wrappers for both settings and show that our wrappers are nearly optimal in terms of accuracy and locality (i.e., the depth of the local neighborhood of the dataset $x$ they explore, which is a proxy for query complexity). In the *claimed sensitivity bound* setting, we provide the first accuracy guarantees that have no dependence on the size of the universe $\mathcal{U}$. We also re-interpret and analyze previous constructions in our framework, and use them as comparison points. In addition to addressing the black-box privacy problem, our private mechanisms provide feasibility results for differentially private release of general classes of functions.

*Boston University. {ejlinder,sofya,ads22}@bu.edu
†Google DeepMind. {adamdsmith,steinke}@google.com