

# Firefox Request for Data Collection

New Firefox-related data collection (for client (e.g. telemetry) and services (e.g. Firefox Accounts) must be reviewed prior to deployment of collection code. Please use this template to submit for data review.

## Purpose of data collection review

- Ensure that the proposed data collection meets our data and privacy policies relating to the specifics of the measurement specifications
- Ensure that there is sufficient documentation for all data collection in the Product

Once you have completed this form, please send to your Data Steward for review.

Before answering this form, think through the following questions:

- What data-driven questions am I trying to answer?
- What data do I actually need to collect to answer these questions?
- How long do I need to have the data?
- How identifiable does the data need to be? (E.g. Can it be aggregated upon collection?)

The following questions are mandatory for all new data collection. **You must have answers to the following questions before you proceed with any work on new data collection.**

1) What are the motivating questions you want to answer with this data?

To determine and resolve issues causing Windows sandboxed child process launch failures.

2) Why does Mozilla need answers to this question? Are there benefits for users? Do we need this information to address business requirements?

If the child process fails to launch for the content process this makes Firefox unusable. Failure to launch NPAPI or GMP affect flash and certain media (generally media delivered using widevine).

We already collect the chromium error code, this will change that collection to be keyed by the process type and Windows error code, which will give us more information to be able to diagnose the issues.

3) List all proposed measurements. Refer to [data collection types for Firefox](#) on the Mozilla wiki and indicate types of measurements for each one. General characterization of your data collection request will fall the highest type of data.

High-level Description	Measurement type	Tracking Bug #
<b>Chromium sandbox error code when the child process fails to launch. Keyed on combination of Gecko process type and Windows error code.</b>	Category 1 “Technical data” <b>Opt-out Keyed histogram.</b>	<b>1395952</b>

4) How long will this data be collected? Choose one of the following:

- This is scoped to a time-limited experiment/project until Firefox 62.
- I want this data to be collected for 6 months initially (potentially renewable).
- I want to permanently monitor this data.

Failure to launch the child process results in a partially or fully unusable browser and is something for which we probably won't get crash reporting. So, it is important for us to always be able to track these.

5) What populations do you need to these measurements for?

- All Windows users running Firefox who have sandboxed child process launch failures.

6) Please provide a general description of how you will analyze this data.

Telemetry will provide the sandbox error code to help identify the failing source code and the Windows error code to help diagnose the reason for failure.

Gecko process type will give an idea of the impact, with failures for the content process being most serious as they result in a non-functioning browser.

7) What alternative methods did you consider to answer this question? Why were they not sufficient?

We could crash the parent process and annotate the crash data, but this would result in a worse experience if the failures are intermittent or not the content process. Crash reporting also potentially captures more privacy sensitive data.

8) Can current instrumentation answer this question?

Not fully, this extends the current collection.

9) Where do you intend to share the results of your analysis?

Within the Content Isolation team and possibly wider depending on the failures.

## Supplementary details

The following questions are mandatory by the time of shipping your instrumentation code.

**Describe the analysis you will perform on the collected data to answer the motivating question.**

- The product manager/**engineering lead**/other will monitor telemetry.mozilla.org dashboards (circle one or more).
- The product manager/engineering lead/other will receive alerts from telemetry.mozilla.org
- The product manager/engineering lead/other will monitor a custom dashboard at sql.telemetry.mozilla.org
- Other (please describe in extensive detail below)

**What data platform are you using to collect this data?**

- Unified Telemetry

**If using Unified Telemetry, please indicate how your measurements are implemented.**

- Telemetry histogram (main ping)

**Provide an example of the measurement artifact (e.g. JSON blob) that will be produced**

- Unified Telemetry