

PP20 CDISC Analysis Results Data with the R Package {cards}

Including Integrations with {gtsummary}

Daniel D. Sjoberg, sjobergd@gene.com

Senior Principal Data Scientist, Genentech



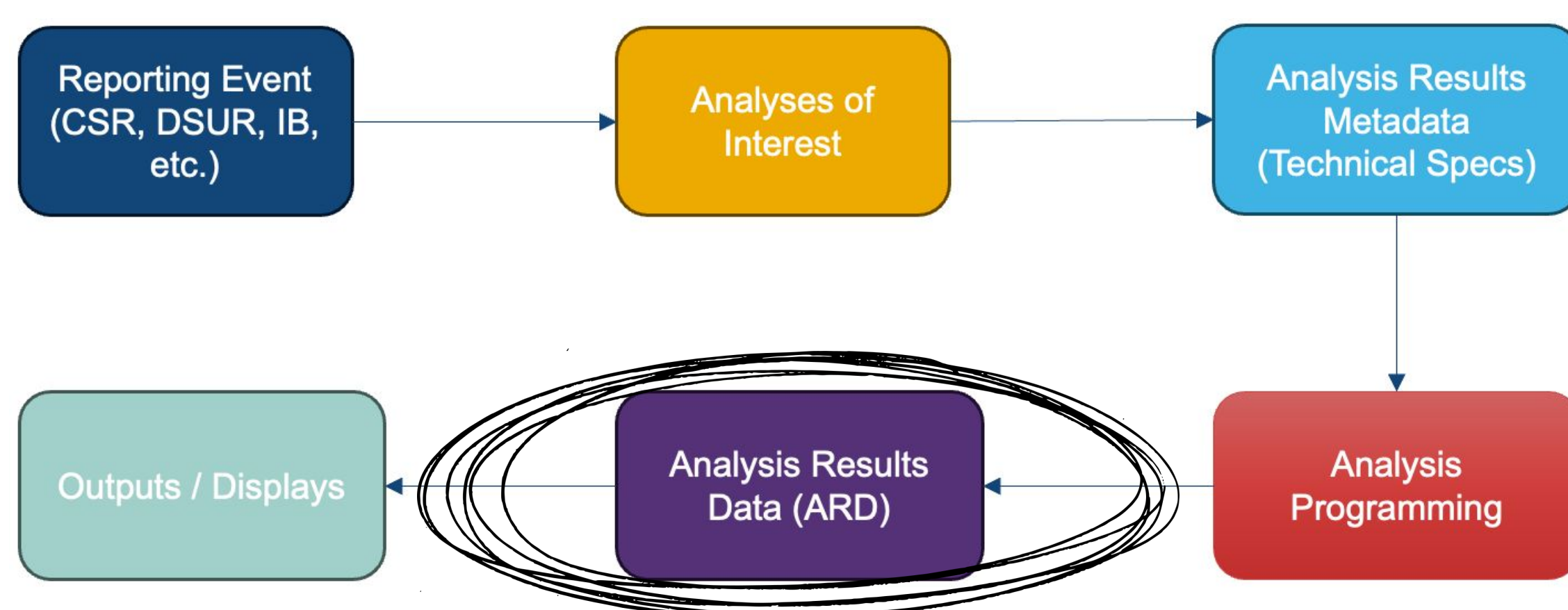
Genentech

A Member of the Roche Group



Analysis Results Standard (ARS)

- Emerging standard for encoding statistical analysis reporting pipeline in a machine-readable format.
- Primary objectives are to leverage analysis results metadata to drive the automation of results and support storage, access, processing, traceability and reproducibility of results.
- Logical model that describes analysis results and associated metadata.
- The **{cards}** package focuses on the creation and manipulation of Analysis Results Data (ARD).



Simple Data Summaries

- The **{cards}** package exports functions for simple continuous and categorical data summaries.

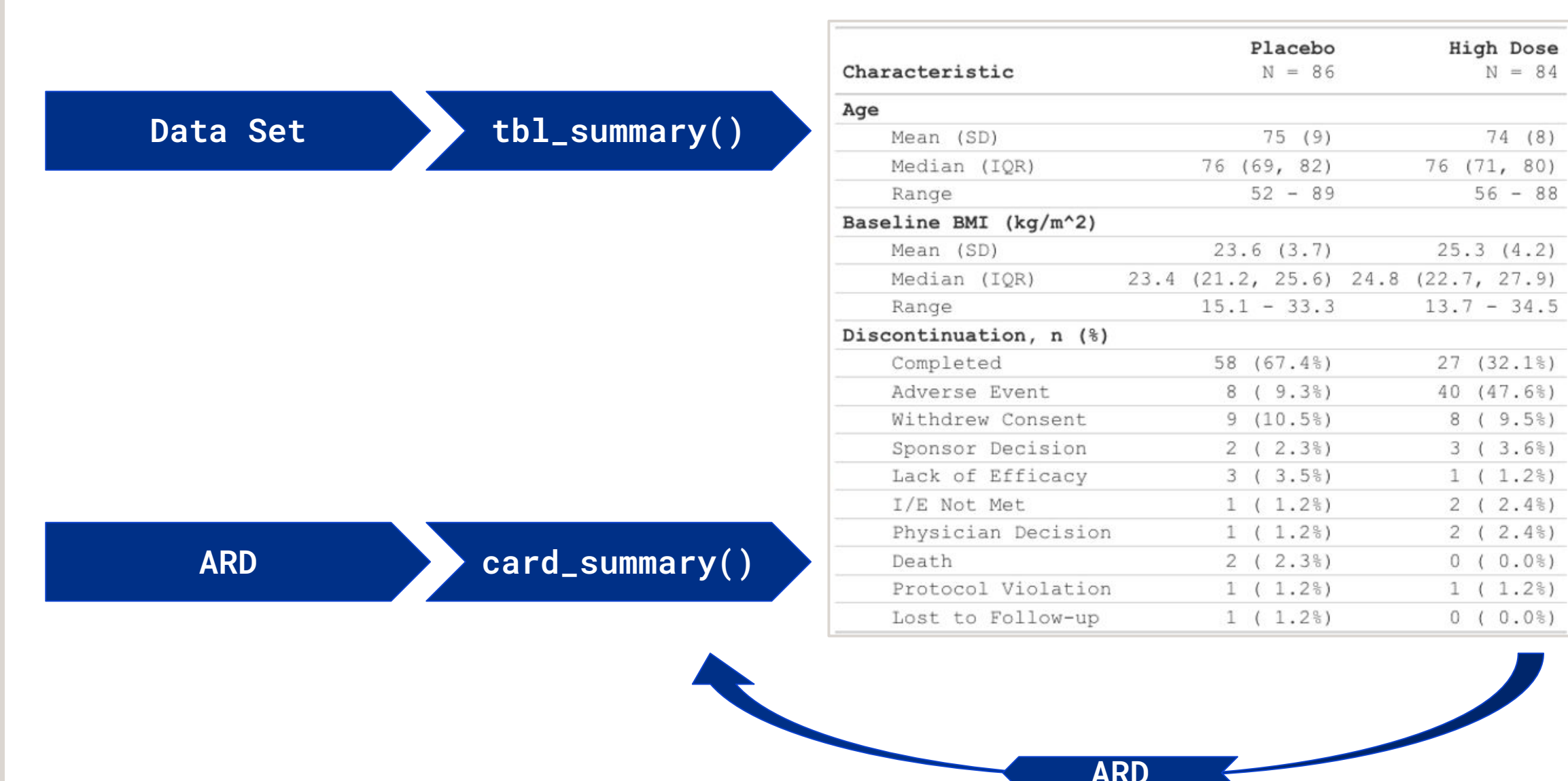
- Easily customizable for any continuous summary and bespoke denominators for categorical data.

```
# install and load {cards} package
remotes::install_github('insightengineering/cards')
library(cards)

ADSL |>
  ard_continuous(variables = c(AGE, BMIBL), by = ARM)
#> {cards} data frame: 24 x 10
#>   group1 group1_level variable stat_name stat_label statistic
#> 1     ARM      Placebo    AGE      mean      Mean      75.209
#> 2     ARM      Placebo    AGE      sd        SD        8.59
#> 3     ARM      Placebo    AGE      median    Median     76
#> 4     ARM      Placebo    AGE      p25       25th Per... 69
#> 5     ARM      Placebo    AGE      p75       75th Per... 82
#> 6     ARM      Placebo    AGE      min        Min     52
#> 7     ARM      Placebo    AGE      max        Max     89
#> 8     ARM      Placebo    AGE      N          N      84
#> 9     ARM      Xanomeli... AGE      mean      Mean     74.381
#> 10    ARM      Xanomeli... AGE      mean      Mean     74.381
#> [i] 14 more rows
#> [i] Use `print(n = ...)` to see more rows
#> [i] 4 more variables: context, statistic_fmt_fn, warning, error
```

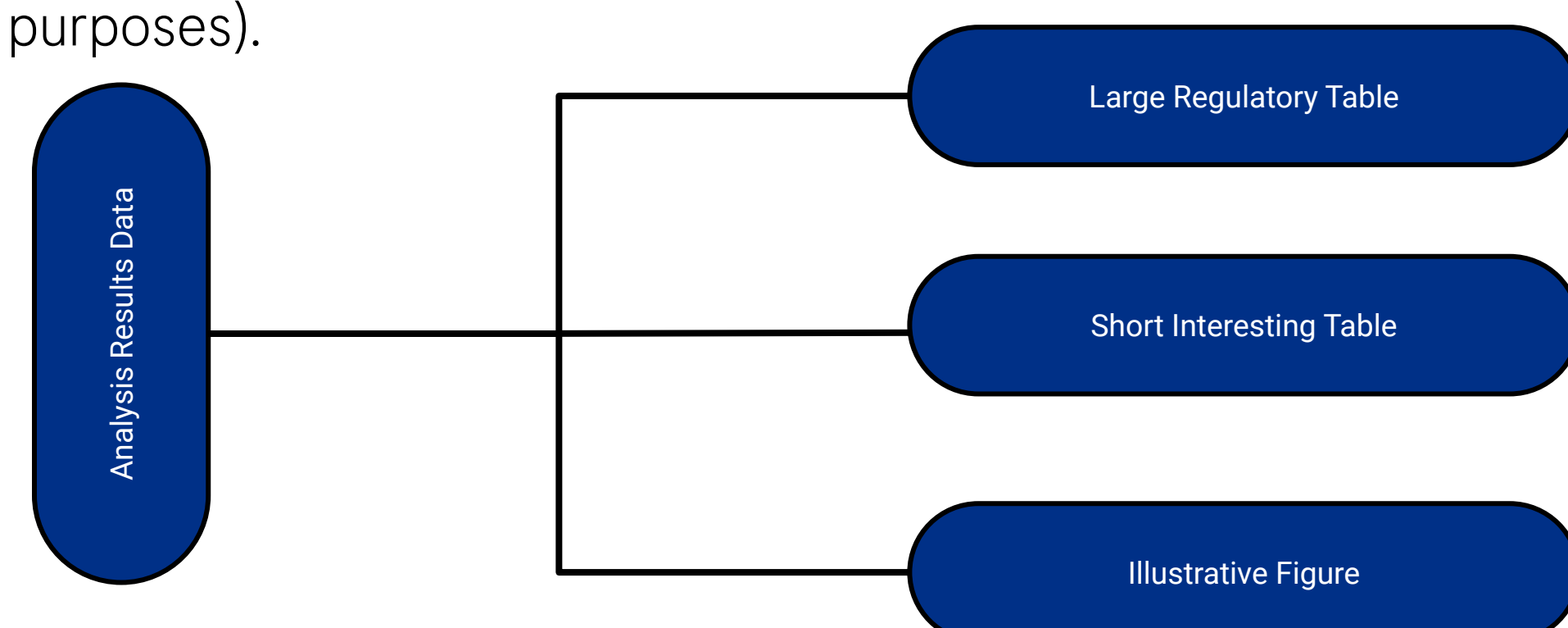
Integration with {gtsummary}

- The **{cards}** package does not present results and this is where the **{gtsummary}** package shines.
- The **{gtsummary}** package offers a modular framework to construct summary tables.
- The **{gtsummary}** package is the most widely used package for summary tables in the R ecosystem, and won the American Statistical Association's 2021 award for Innovation in Statistical Programming and Analytics and it's currently being refactored with a **{cards}** backend.



Analysis Results Data (ARD)

- Encodes statistical analysis outcomes in a machine-readable format.
- Primary objective is to streamline the processes of automation, ensuring reproducibility, promoting reusability, and enhancing traceability.
- The ARD model specified how statistical results are saved into a structured format.
- The ARD can be used to to subsequently create tables and figures.
- After the initial creation of an ARD, the results can later be re-used again and again for subsequent reporting needs; for example, re-creating a figure with only two of the three treatment arms shown, or pruning a large complex table to show only the most interesting results for a presentation (when the initial results may been prepared for regulatory purposes).



Extending {cards}

- The **{cards}** package includes functions that make it simple to extend, and many of these extensions live in the **{cardx}** package (extra cards).

- Leverage the enormous effort in the **{broom}** package to build new ARD functions. Optionally, include function arguments in the results.

```
# The cardx::ard_ttest() function is simple to create using the utilities.
ADSL |>
  dplyr::filter(ARM %in% c("Placebo", "Xanomeline High Dose")) |>
  ard_ttest(by = ARM, variable = AGE)
#> {cards} data frame: 14 x 10
#>   group1 group1_level variable stat_name stat_label statistic
#> 1     ARM      Placebo    AGE      estimate Mean Dif... 0.828
#> 2     ARM      Placebo    AGE      estimate1 Group 1 ... 75.209
#> 3     ARM      Xanomeli... AGE      estimate2 Group 2 ... 74.381
#> 4     ARM      Placebo    AGE      statistic t Statist... 0.655
#> 5     ARM      Placebo    AGE      p.value   p-value    0.513
#> 6     ARM      Placebo    AGE      parameter Degrees ... 167.362
#> 7     ARM      Placebo    AGE      conf.low  CI Lower... -1.668
#> 8     ARM      Placebo    AGE      conf.high CI Upper... 3.324
#> 9     ARM      Placebo    AGE      method    method Welch Tw...
#> 10    ARM      Placebo    AGE      alternative alternat... two.sided
#> 11    ARM      Placebo    AGE      mu         H0 Mean     0
#> 12    ARM      Placebo    AGE      paired    Paired t... FALSE
#> 13    ARM      Placebo    AGE      var.equal Equal Va... FALSE
#> 14    ARM      Placebo    AGE      conf.level CI Confi... 0.95
#> [i] 4 more variables: context, statistic_fmt_fn, warning, error
```

YAML, JSON, and REST API

- The **{cards}** package creates a data frame with the results, but the object is full representable in other formats.
- Using the **{plumber}** R package, we can create a REST API with a handful of lines of code
- If an organization prefers a YAML representation of the results, simple call `as_nested_list(x) |> yaml::as.yaml()`

```
# Below is an example creating a JSON representation
ard_continuous(ADSL, variables = AGE, by = ARM) |>
  as_nested_list() |>
  jsonlite::toJSON(pretty = TRUE)
#> {
#>   "variable": {
#>     "AGE": {
#>       "group1": {
#>         "ARM": {
#>           "group1_level": {
#>             "Placebo": {
#>               "stat_name": {
#>                 "N": {
#>                   "statistic": [86],
#>                   "statistic_fmt": ["86"],
#>                   "warning": {},
#>                   "error": {},
#>                   "context": ["continuous"]
#>                 }
#>               }
#>             }
#>           }
#>         }
#>       }
#>     }
#>   }
```