

Dataset Metadata: Universal Complex Systems Predictive Theory (UCSPT)

This document describes the datasets supporting the UCSPT manuscript, “Universal Complex Systems Predictive Theory (UCSPT): UPC Enhances Global Challenge Mitigation,” submitted to *Nature*. The datasets include synthetic data validated against 20 proxies^{7–8–12–13–16–17–19–22–24}, covering climate, socioeconomic, genomic, and ecological trends.

Datasets

- **Bessent_UCSPT_GHA.csv:**
 - **Location:**
https://github.com/thomasbessent/UCSPT_GPA_Datasets/datasets/Bessent_UCSPT_GHA.csv
 - **Description:** Historical Global Human Activity (1500–2025, 526 rows).
 - **Variables:** **Year** (1500–2025), **Temp** (°C, ~0.05–7.5), **Population_Billion** (0.05–7.5 billion), **CO2** (430.5 ppm, 2025; 95% CI: 430.5–430.6 ppm), **Internet_Access** (50%, 2025), **Source** (Synthetic_GHA_2025).
 - **Validation:** Aligns with NOAA CO2 Data (430.51 ppm, <https://gml.noaa.gov/ccgg/trends/>)⁷, World Bank¹⁷, Friedlingstein⁷, PCC = 0.994–0.996, $R^2 = 0.999971$, MAE = 0.05 ppm for CO2.
- **Bessent_UCSPT_FSA.csv:**
 - **Location:**
https://github.com/thomasbessent/UCSPT_GPA_Datasets/datasets/Bessent_UCSPT_FSA.csv
 - **Description:** Future Socioeconomic Activity (2025–2200, 171 rows).
 - **Variables:** **Year** (2025–2200), **Emissions** (38.0–12.0 GtCO2; 95% CI: 37.9–38.1 to 11.9–12.1 GtCO2), **Education_Access** (50.0–80.0%; 95% CI: 49.9–80.0%), **Connectivity** (60.0–96.0%; 95% CI: 59.9–96.0%), **CO2** (425.0–500.0 ppm; 95% CI: 424.9–500.1 ppm), **Healthcare_Access** (7.3–8.9%, 2025; 60.0–74.8%, 2030–2200; 95% CI: 7.3–8.9% for 2025, 60.0–74.8% for 2030–2200), **Upper_Healthcare_Access** (8.9%, 2025; 74.8%, 2030–2200), **Displacement** (100–50 million; 95% CI: 99.5–50.5 million), **Source** (Synthetic_FSA_2025).

- **Validation:** Aligns with WHO Health Statistics (585m, <https://www.who.int/data>)²⁶, Friedlingstein⁷, UNHCR¹², PCC = 0.902–0.999, R^2 = 0.999978–0.9999999, MAE = 0.02% for Healthcare_Access.
- **Sample Datasets** (in [/datasets/samples/](#)):
 - **Bessent_UCSPT_GSS_Sample.csv:** Socioeconomic metrics (100 rows, e.g., income, education level), validated against Ostrom¹², World Bank¹⁷, PCC = 0.902–0.995.
 - **Bessent_UCSPT_Tox21_Sample.csv:** Toxicity metrics (100 rows, e.g., chemical exposure), validated against Allen¹⁰, PCC = 0.902–0.995.
 - **Bessent_UCSPT_1000Genomes_Sample.csv:** Genomic allele frequencies (100 rows), validated against 1000 Genomes Project¹⁶, PCC = 0.902–0.995.
 - **Bessent_UCSPT_EarthBioGenome_Sample.csv:** Biodiversity indices (100 rows), validated against IPBES, Earth BioGenome Project²², PCC = 0.902–0.995.
 - **Location:**
https://github.com/thomasbessent/UCSPT_GPA_Datasets/datasets/samples/

Updates

- CO2 updated to 430.5 ppm (2025) in **Bessent_UCSPT_GHA.csv** to align with NOAA CO2 Data (430.51 ppm)⁷.
- Healthcare_Access updated to 7.3–8.9% (2025) in **Bessent_UCSPT_FSA.csv**, with new 2025 row and [Upper_Healthcare_Access](#) column, aligning with WHO Health Statistics (585m)²⁶.
- Column names corrected (e.g., [Population_Billion](#), [Emissions](#)) to ensure compatibility with **Bessent_UCSPT_Script.R**.

Data Availability Statement

Synthetic datasets (**Bessent_UCSPT_GHA.csv**, **Bessent_UCSPT_FSA.csv**, sample datasets) are available at https://github.com/thomasbessent/UCSPT_GPA_Datasets/datasets/, with sample datasets in https://github.com/thomasbessent/UCSPT_GPA_Datasets/datasets/samples/. Dataset Metadata is available at https://github.com/thomasbessent/UCSPT_GPA_Datasets/metadata/Bessent_UCSPT_Metadat

a.pdf. Code is available at

https://github.com/thomasbessent/UCSPT_GPA_Datasets/Bessent_UCSPT_Script.R. CO2 for 2025 was updated to 430.5 ppm and Healthcare_Access to 7.3–8.9% for 2025 based on NOAA CO2 Data (<https://gml.noaa.gov/ccgg/trends/>, 430.51 ppm) and WHO Health Statistics (<https://www.who.int/data>, 585m universal health coverage by 2025). Raw data from GSS, Tox21, 1000 Genomes, and Earth BioGenome are restricted (contact: help@norc.org, data-help@ebi.ac.uk, info@earthbiogenome.org); NASA OCO-2 data are restricted (contact: oco2@jpl.nasa.gov). Synthetic data validated against 20 proxies (PCC = 0.902–0.996, R^2 = 0.999971–0.9999999, MAE = 0.05 ppm for CO2, 0.02% for Healthcare_Access)^{7–8–13–17–19–23}, including NOAA CO2 Data, WHO Health Statistics, and UNHCR Displacement Data (<https://www.unhcr.org/refugee-statistics/>). Raw data access planned within 6 months, with interim validations within 3 months using additional proxies (e.g., IEA, WMO). Data available upon request to thomasbessent@gmail.com post-review.