



Context-Aware AI Search with UAS data and LLMs

February 2026
Bas Weerman
weerman@usc.edu



Aim

- To create a modern AI-driven search platform that delivers precise, context-aware answers by combining UAS data and metadata with advanced LLM reasoning.



UAS AI Search: Unified, Intelligent Search

- Ask any natural-language question about UAS data and get precise, relevant answers.
- One hub for all surveys, visualizations, responses, and publications. No more searching across pages.
- Automatically aggregates and retrieves the most meaningful content from all our data sources.
- Scales as new data arrives through a single, unified retrieval framework.
- Provides summarization and insights to help users quickly understand complex survey data.



Examples of questions

- What health conditions does the UAS ask about?
- Can I look at the relation between health and income when using the data?
- Are there any questions related to pensions?
- What is the response rate for your surveys on cognition?
- Are there any papers on retirement savings that use your data?
- How many of your respondents are wearing fitbits?



How UAS AI Search works (1)

- Open-Source LLMs for transparency, flexibility, and ‘offline’ deployment.
- Semantic Search identifies the most relevant historical UAS survey questions using embeddings and vector search.
- RAG Pipeline grounds LLM answers in real survey data by retrieving matching questions and responses before generation.
- Unified Data Layer pulls related documents, publications, and visualization data across all UAS systems.
- Smart Summaries extract key sections from PDFs and surface only the most relevant insights—no document scrolling required.



How UAS AI Search works (2)

- Uses open-source LLMs for cost efficiency, adaptability, and full customization.
- Understands user intent to surface the most relevant questions, variables, and insights that go far beyond keyword search.
- Combines all content into one view: documentation, visuals, micro data and publications from a single query.
- Pulls from all key sources and uses AI to highlight only the most relevant sections with clear, concise summaries.



Current input data sources UAS

- Survey codebooks: survey number, survey title, questions, and sections
- Website info: content available on the UASDATA website
- Comprehensive File and Panel Dataset: includes all variables and descriptions
- Publication and visualization data: includes the publication title, author, abstract etc. and our visualizations
- Exposome data: Information from our Exposome project

Understanding America Study | USC Dornsife Center for Economic and Social Research

ABOUT | DOCUMENTATION | DATA PRODUCTS | FOR CLIENTS | PUBLICATIONS | SEARCH

This repository is under review for potential modification in compliance with Administration directives.

WELCOME TO THE Understanding America Study

GETTING STARTED GUIDE

WHAT'S NEW

INTERACTIVE DATA EXPLORER

FOR PARTICIPANTS

FOR CLIENTS

Come to CIPHER 2026 in D.C.! February 26-28, 2026. Free to attend, just RSVP. [Click here for details](#)

Interested in collecting data for free? [Click here to learn more](#). If you do not have a google account, please follow these alternative instructions.

Want to quickly learn more about the UAS? [Click here to download our brochure or read the UAS cohort profile](#).

WELCOME
The Understanding America Study (UAS), housed at the University of Southern California, is a nationally representative, probability-based internet panel of approximately 15,000 U.S. residents, including a

UAS Exposome Database | exposome.usc.edu/?p=domains

Domains

This website acts as a centralized hub for accessing and linking exposome data across various research disciplines. We emphasize three interconnected areas: the physical environment, social environment, and policy environment.

- Physical Environment**
Currently we have 4 categories: Pollution, Food Environment, Street Connectivity and Nature Quality
- Social Environment**
Currently we have 4 categories: Crime Reports, Census Survey, Housing and IPUMS
- Policy Environment**

UNDERSTANDING AMERICA STUDY | THE USC DORNISIFE ELECTION "DAYBREAK" POLL

FULL WAVE ESTIMATE | ALL GRAPHS | 2016 | SURVEY METHODS | CONTACT US

Probabilistic voting questions | Predicted percentage of votes for President

by None | window 7 Day | Go

Predicted percentage of votes for President (7-day window)

based on probabilistic voting questions

Interactive Data Explorer | usvis.usc.edu/monthly/monthlyvisualization.php?v=a4199de&d=073a4d5a&g=trends

Welcome to the Interactive UAS Monthly Panel Data Explorer

For quick guide on how to use this dashboard, please click [here](#)

Show 95% Confidence Interval

Life satisfaction

Category: Completely or very satisfied

Understanding America Study | USC Dornsife Center for Economic and Social Research

ABOUT | DOCUMENTATION | DATA PRODUCTS | FOR CLIENTS | PUBLICATIONS | SEARCH

This repository is under review for potential modification in compliance with Administration directives.

PUBLICATIONS > ACADEMIC PAPERS

2026

- Purpose in life and depressive symptoms: An individual-participant meta-analysis in regions** - *Journal of Affective Disorders*, Volume 398, 1 April 2026, 120881 ; Angelina R. Sut Karakose , Norma Mansor, Antonio Terracciano
- When Political Pivots Shift Behaviors but Not Beliefs** - *American Sociological Review*, C
- Why don't eligible workers receive unemployment insurance?** - *Applied Economics Le*
- Side Effects: The Social Ecology of Adverse Drug Reactions** - 978-0231217798, New York
- Are high neuroticism and low financial well-being mutually-exacerbating risk factors in a process longitudinal study** - *Journal of Affective Disorders*, Volume 394, Part B, 1 February
- Navigating child SSI: A qualitative study of systemic, family, and provider-level inf** Lila Rabinovich, Tabasa Ozawa

Physical Activity Data Explorer | rfbt.cesusc.org/index.php

This page serves as the home for exploring the Understanding America Study (UAS) Physical Activity Data.

The UAS is a panel of households at the University of Southern California (USC). Currently comprising of approximately 14,700 respondents, the panel is expected to grow 20,000 by end of 2025. As an Internet-based panel, participants complete surveys using a computer, tablet or smartphone - anytime and from anywhere they choose.

To date, we have collected 1,169,151,705 (~1.17 billion) minutes of activity tracker data across 7,271 unique participants.

For quick guide on how to use this website, please click [here](#)

Our Aims

- Share activity and fitness data collected by various trackers with the wider research community.
- Create guidance for utilizing physical activity data.
- Enable seamless download and interactive visualizations of monthly aggregated physical activity data.

UAS AI x +

dev.cesrusc.org/uasai/demo/index.php?p=home

UAS ? AI Understanding America Study USC Dornsife Center for Economic and Social Research USC University of Southern California

UAS Search Engine

Q Are there any questions related to pensions? Search

Some questions you can ask:

- What questions does the UAS ask about healthcare?
- Are there any questions related to pensions?
- Have you asked about buy now pay later?
- Which surveys ask about depression?

Copyright © 2017 The University of Southern California. All Rights Reserved. // Terms and Conditions

UAS AI

dev.cesrusc.org/uasai/demo/index.php

UAS ? AI

Understanding America Study

USC Dornsife
Center for Economic and Social Research

USC University of Southern California

← Ask another question

Are there any questions related to pensions?

Top relevant results generated by the LLM:

AI Overview

The Understanding America Study (UAS) surveys include questions related to pensions. Specifically, the UAS Health and Retirement project covers pensions within its six modules. The UAS HRS Wave 4 surveys include section J-J2/J3/M1-M2 which covers employment, pensions, and workability due to health. The UAS HRS surveys in general include questions about pensions as part of the financial wellbeing domain.

[Health And Retirement](#) [Health And Retirement Wave 4](#) [Health And Retirement Wave 4](#)

Surveys found

Survey Number	Title	Survey Home
UAS 22	Health and Retirement Study Sections J-M	↗

Relevant Survey Sections

UAS 22

[sectionL](#), [sectionJ2](#)

Questions and variables information

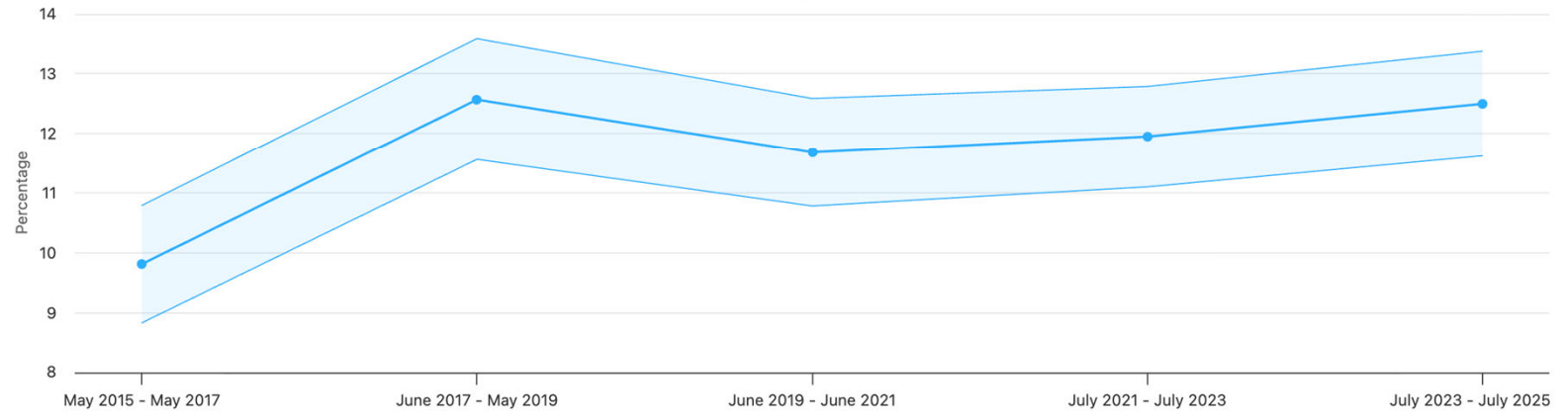
Questions and variables information

Variable	Question	Surveys	Section
LW069	The next few questions relate to retirement plans from other current or former jobs from which you are not currently receiving any income payments, and that you have not already told us about. Thinking first about the type of plans that accumulate a balance, what is the sum of the balance in all the remaining pension plans where you expect to receive benefits in the future?	UAS 22	sectionL
JW401	Do you have any other retirement or tax deferred savings accounts or pensions with your current job or union?	UAS 22	sectionJ2

Data Trends

Currently receiving pension income

- % Yes



About Visualization:

Who receives that income from retirement pensions?

Relevant Publications

3 Publications Found

Factors Influencing the Choice of Pension Distribution at Retirement

2022

Authors: Robert L. Clark, Olivia S. Mitchell

Outlet: Wharton Pension Research Council Working Papers

Abstract

One of the most important financial decisions that pension participants make concerns how they access their pension assets when they terminate employment with their plan sponsor. Their choices depend both on own preferences and the options offered by their retirement plan. This paper examines both past and future pension withdrawal choices for those with defined benefit and defined contribution pensions, separately. Our data are drawn from a set of pension distribution questions we fielded in the Understanding American Study. Results show significant differences in distribution choices based on the type of retirement plan, with individuals covered by defined benefit plans significantly more likely to select annuities compared to similar employees covered by defined contribution plans. We also find differences in how higher annual income affects annuity choices based on coverage by DB plans. Individuals with lower levels of financial literacy and lower annual income have less knowledge of basic pension characteristics.

[Read full publication](#)

How Do Older Workers Learn from Social Security about Their Benefits?: Characterizing Variation in my Social Security Account Access and Social Security Statement Receipt in the 2010s

0

Authors: Philip Armour

Outlet: NBER Center Paper

Abstract

Effective communication about Social Security benefits is vital to inform workers about their entitlements, aiding in claiming decisions and retirement planning. SSA communicates with future and current beneficiaries with a range of tools, but the two largest in terms of scope are the Social Security Statement and my Social Security online accounts. There has been substantial recent variation in both of these communications: prior to 2011, paper Statements were automatically sent to all workers age 25 and over; in 2011, budget constraints led to a cessation of automatic mailings. In the six years that followed, there were three separate reintroductions and two additional cessations, leading to the current policy, whereby only eligible individuals age 60 or over receive an automatic mailing every year. This variation in

UAS AI x +

dev.cesrusc.org/uasai/demo/index.php

Relevant Exposome Variables

Percent of Medicare decedents hospitalized at least once during the last 6 months of life

Variable Name: **d266** Source Name: **Dartmouth University**

Category: **public health** Dataset Name: **Dartmouth Atlas of Health Care**

Unit: **percent** Domain: **policy**

Citation: **Jennifer Ailshire and Matthew M. Young. 2018. Contextual Data Resource (CDR): Dartmouth Atlas of Health Care by Hospitals, HSAs, HRRs, Counties and States, 1992-2011. Version 1.0. Los Angeles, CA: USC/UCLA Center on Biodemography and Population Health.**

[Source Link](#)

Percent of beneficiaries filling at least one prescription for a dementia medication

Variable Name: **d555** Source Name: **Dartmouth University**

Category: **public health** Dataset Name: **Dartmouth Atlas of Health Care**

Unit: **percent** Domain: **policy**

Citation: **Jennifer Ailshire and Matthew M. Young. 2018. Contextual Data Resource (CDR): Dartmouth Atlas of Health Care by Hospitals, HSAs, HRRs, Counties and States, 1992-2011. Version 1.0. Los Angeles, CA: USC/UCLA Center on Biodemography and Population Health.**

[Source Link](#)

Proportion of residents present on the 1st Thursday in April who have Alzheimer's disease or related dementia.

Variable Name: **pctalz_dem_sta** Source Name: **LTCFocus**

Category: **LTCfocus: Health and Retirement Study Contextual Data Resource (HRS-CDR)** Dataset Name: **LTCFocus Public Use Data**

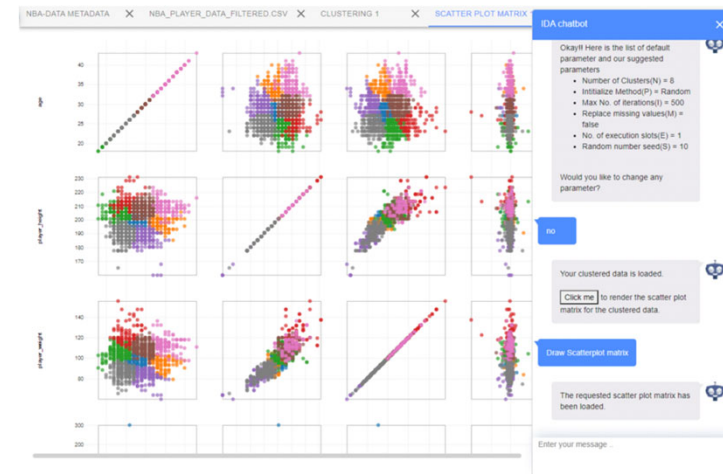
Unit: **percent** Domain: **policy**

Citation: **LTCFocus Public Use Data sponsored by the National Institute on Aging (P01 AG027296) through a cooperative agreement with the Brown University School of Public Health.**



Future plans

- Release current version in coming months
- Chatbot
- Tool for analysis, visualization, and automated reporting





Thanks!



weerman@usc.edu