

ACS

Accelerated Safety Analysis Protocol (ASAP) Tool

First Used: May 2018

Computation Type: Ranking
Identifying Information: Yes

Purpose Type: Performance evaluation
Updated in 2024: No

Tool Description

Predictions of Severe Harm (identifying likelihood of substantiated allegations of physical or sex abuse within the next 24 months) are based on a machine learning methodology and are calculated for all children involved in active investigations early in the investigation (day 10). An investigation is assigned a numeric likelihood of this outcome based on the child in the case with the highest likelihood. The ACS Quality Assurance unit in the Division of Child Protection reviews about 3,000 active investigations annually, selecting those with the highest likelihood of severe harm.

Tool Purpose

The Quality Assurance Unit in the Division of Child Protection at ACS has the capacity to review about 3,000 investigation cases out of about 50,000 investigations annually. ACS developed this predictive model to support the selection of cases for Quality Assurance review. Open investigations involving children with the greatest likelihood to experience future severe harm – substantiated allegations of physical or sex abuse in the following 24 months – are selected for review. The tool does not support decisions about services or interventions for individuals or families involved with ACS, beyond the selection of the case for this additional Quality Assurance review.

If the Quality Assurance review team identifies gaps in routine, required documentation or practice, the team speaks with the field office conducting the investigation and follows up to make certain these gaps have been addressed. Scores are not shared with staff in the Quality Assurance unit or the investigative unit. The model only supports the decision about which investigations are prioritized for review by the Quality Assurance (QA) unit.

Populations Impacted

Individuals

Data Analyzed

Training Data

ACS trained the model on ACS historic administrative data about closed investigations from April 2014 to April 2016. The training set included about 142,026 observations. The model was tested on closed investigations from April 2016 to April 2017 with 53,477 observations.

Input Data

Predictions are based on administrative data about prior and current child welfare involvement including investigations triggered by a New York State Central Register (SCR) call and time spent in foster care. Only ACS administrative data are used in the model.

Output Data	Rank ordered list of open investigation cases involving children with the highest likelihood to experience future severe harm, defined as substantiated allegations of physical or sex abuse in the following 24 months to be reviewed by a special QA Review Team.
--------------------	---

Caseloads Projection

First Used: Jul 2024

Computation Type: Forecasting	Purpose Type: Resource allocation
Identifying Information: No	Updated in 2024: Tool was created in CY2024

Tool Description

The model estimates weekly caseloads by area for the next 12 months from the run date (day 14 of the prevention case).

The model is a time series model that forecasts caseloads Citywide. The forecast analysis was performed at a more granular geography including Borough and Zone levels. Borough level analysis worked well but the data at Zone level is noisy for data modeling purposes. The model accounts for seasonality, trend, and variation in data to accurately project future caseloads.

Staffing decisions are made based on projected caseloads, historical attrition rate, and current team size.

Tool Purpose

Caseloads projections help identify how busy certain neighborhoods are likely to be in the next 12 months to optimize workforce distribution across different Boroughs in the City. As social workers complete their training program, they are assigned to field offices. The right number of Child Protective Service (CPS) Social Workers needs to be assigned to each Borough based on the workload.

Populations Impacted

Geographic space

Data Analyzed

Training Data	The model was trained on caseloads from January 2021 to July 2023 and tested on caseloads from July 2023 to July 2024.
Input Data	Predictions are based on administrative data on investigations and FSU involvement were extracted from Connections (CNNX). A snapshot of all open investigations was taken at the start of the week (every Monday). This value was treated as the average caseload for the week. Only ACS administrative data are used in the model.
Output Data	List of Boroughs with open investigation cases involving children.

Housing Prioritization

First Used: Apr 2023

Computation Type: Ranking
Identifying Information: No

Purpose Type: Resource allocation
Updated in 2024: No

Tool Description

The model estimates a risk score for a child receiving prevention services whose family will apply and be eligible for a homeless shelter within 12 months from the start of service (day 14 of the prevention case).

The model helps predict the risk of application for homeless shelters among families receiving prevention services. With a limited number of vouchers available, the risk model helps ACS prioritize housing assistance for those families at greatest risk of becoming homeless.

The Service Provider meets with the family to conduct a qualitative assessment of the family's housing needs and vouchers are offered based on their findings. This is one of multiple ways in which ACS can identify families potentially eligible for housing assistance.

Tool Purpose

The City has allocated 100 housing vouchers to families receiving ACS Prevention Services. The shelter application model identifies the likelihood of a family in prevention services applying for homeless shelter within 12 months beyond the current prevention case. The model uses a machine learning methodology and is calculated for all children in a prevention case. ACS Prevention Services uses the results as one of several ways of identifying possible families that service providers can assist in applying for shelter.

Populations Impacted

Individuals

Data Analyzed

Training Data

ACS trained the model on ACS historic administrative data regarding preventive services started between 2014 and 2020. An 80/20 split of data to train on 80% and test on 20% ensuring that no family appears in both sets. The training set contains 140,242 observations between Jan 2014 and December 2020. The test set consisted of 34,508 observations between Jan 2014 and December 2020. Predictions are based on administrative data about prior and current child welfare involvement at the start of a case. This includes SCR investigations and time spent in foster care. Only ACS administrative data are used in the model.

Input Data

Output Data

Rank ordered list of open prevention cases involving children whose families have the highest likelihood of applying for a homeless shelter within 12 months of starting a prevention service.

Prevention Score Card

First Used: Sep 2021

Computation Type: Ranking
Identifying Information: Yes

Purpose Type: Performance evaluation
Updated in 2024: No

Tool Description

Predictions of Repeat Maltreatment (identifying the likelihood of being involved in a future indicated investigation within the next 24 months at the start of service) are based on a machine learning methodology and are calculated for all children receiving prevention services from ACS prevention service providers.

Tool Purpose

The Repeat Maltreatment model is used to make predictions on day 10 from the start of the prevention case to assess the risk of the family at the beginning of the service. A prevention case is assigned a numeric likelihood of an indicated investigation based on a New York State Central Register (SCR) within 24 months from the start of a prevention service.

The prevention providers are assessed for their performance based on the service needs/risk levels of the families they've served during the previous fiscal year.

The programs were sorted and ranked based on their average risk, and then divided into four quartiles by rank order: the top 25 percent of programs are classified as the Very High-Risk Cohort, the next 25 percent of programs as the High-Risk Cohort, the next 25 percent as the Medium-Risk Cohort, and the lowest 25 percent as the Low-Risk Cohort. Assignment to a cohort is not a way of performance assessment of the program but to group prevention service providers for fair comparisons based on the risk level of families they served.

Populations Impacted

Individuals

Data Analyzed

Training Data

ACS trained the model on ACS historic administrative data about closed investigations from July 2009 to June 2016. Training set included about 158,787 observations. The model was tested on closed investigations from July 2016 to June 2018 with 46,969 observations.

Input Data

Predictions are based on administrative data about prior and current child welfare involvement at the start of a case. This includes SCR investigations and time spent in foster care. Only ACS administrative data are used in the model.

Output Data

The model is used for generating a scorecard of prevention service providers by categorizing prevention programs based on the average risk profile of the cases they served during the assessment year. These groupings of program cohorts provide context for understanding the Scorecard, as it allows for performance comparison of programs that accepted and served families with similar risk profiles.

DOE## Eureka! Chatbot

First Used: Aug 2023

Computation Type:	Purpose Type: NA
Identifying Information: Yes	Updated in 2024: No

Tool Description

The Azure Cognitive Services technology and chatbot (internally branded as “Eureka!” has been configured and deployed in August 2023 to be the first response to calls to the NYCPS IT Service Desk. It accesses scripts to handle four common reasons for a user to call or contact the service desk - Password Reset, Create a Ticket, Ticket Status, Request for Information. The chatbot accesses pre-defined scripts to respond to user voice or text input. The user’s request is either serviced, completed and closed by the chatbot, or the user is given the option (at any time) to connect to a live agent.

Tool Purpose

The tool is used to respond to common IT service desk requests - Password Reset, Create a Ticket, Ticket Status, Request for Information. Users can access the tool by phone, by computer through the DOE Support Hub application, and from links from MS Teams and other DOE systems, such as TeachHub.

Populations Impacted

Individuals

Data Analyzed

Training Data	Pre-defined scripts designed to respond to four common requests to the IT Service Desk.
Input Data	A voice call or text-based chat session initiated by a user and responded to by the Eureka! chatbot before being handled by a human Service Desk agent.
Output Data	The chatbot generates responses to user-entered prompts based on the training data, or forwards the inquiry to a human Service Desk agent. Since its launch in August 2022, the chatbot handles an average of 1,500 calls and 300 web-based inquiries each day. Approximately 30 percent of the voice calls and 10 percent of the web-based queries have been handled completely by Eureka! without being forwarded to a human Service Desk agent.

MySchools

First Used: Aug 2018

Computation Type:	Purpose Type: NA
Identifying Information: Yes	Updated in 2024: Yes

Tool Description

The tool utilizes the Gale-Shapley deferred acceptance algorithm to match applicants to schools. This algorithm has been in existence for many years, used internationally for various purposes. Perhaps most common is its use in the National Resident Matching Program for medical school students.

Deferred acceptance works as an iterative series of steps: students and programs are tentatively matched in each step, but nothing is finalized until the algorithm terminates (hence the deferred). 1. Each student “proposes” to their first choice • Programs assign seats to students one at a time • When all seats are filled, programs may reject previously accepted students in favor of new applications from students they prefer (e.g., students with a better lottery number) • Remaining students are rejected 2. Students rejected in the last step “propose” to the next choice on their list 3. The algorithm terminates when all students are matched or have proposed to all the programs they listed

Tool Purpose

MySchools is an application used to house online school directories, collect application choices, and run the admissions matching algorithm that is used for all centralized admissions processes (3K, pre-K, Gifted & Talented, middle school, and high school). The tool encompasses a family-facing portal, a school-facing portal, and an administrative portal.

Populations Impacted

Individuals

Individuals impacted include: Students

Data Analyzed

Training Data

The algorithm was already widely recognized for its advantages prior to adoption in New York City. The DOE consulted with a team of researchers at MIT who had been closely involved in its initial creation when we adopted it.

Input Data

Student biographical information (e.g., home address, poverty status, home language), student academic information (e.g., course grades, state test scores), and student school records (e.g., sending school).

Output Data

The algorithm outputs a school match for each student.
