

Bringing your ECR questions to SafeInsights

JP Slavinsky

Executive Director

July 1, 2026

10:00 AM – 11:15 AM ET

NSF ECR PI Meeting, Regency B

Acknowledgements



Debshila Basu Mallick
Scientific Director



Brittany Bradford
Research Scientist

Today's Agenda

01 Setting the stage

The gap and the SafeInsights model

02 End-to-end walkthrough

See SafeInsights in action

03 Governance & IRB

How a secondary analysis study moves through the SafeInsights system

04 Scope your own study

Map your research question to educational data sources

What question about student STEM learning would you dive into with rich platform and institutional data?

01.

Setting the stage

Education research: Aspirations vs. reality

What education researchers want

- Representative samples
- Authentic behavioral and learning data
- Longitudinal trajectories
- Subgroup analyses without re-identification risk
- Reproducible research at scale

What the access landscape allows

- Convenience samples
- Surveys and self-reports
- Cross-sectional snapshots
- Stripped-out subgroup detail to preserve privacy
- One-off IRB arrangements

Safelnsights is...

- A research platform that enables large-scale research about learning
 - ...**doesn't let researchers see the data**
 - ...where the data stays put
 - ...built for **education and related data**
 - ...that can **convert research to effective practice**
- Technical and logistical infrastructure
- A tech company within Rice University
- A nationwide team of experts, edtech companies, & educational institutions
- NSF's biggest investment in education research infrastructure

What SafeInsights Enables

Large-Scale Data Analysis

Secure secondary analyses,
educational data mining

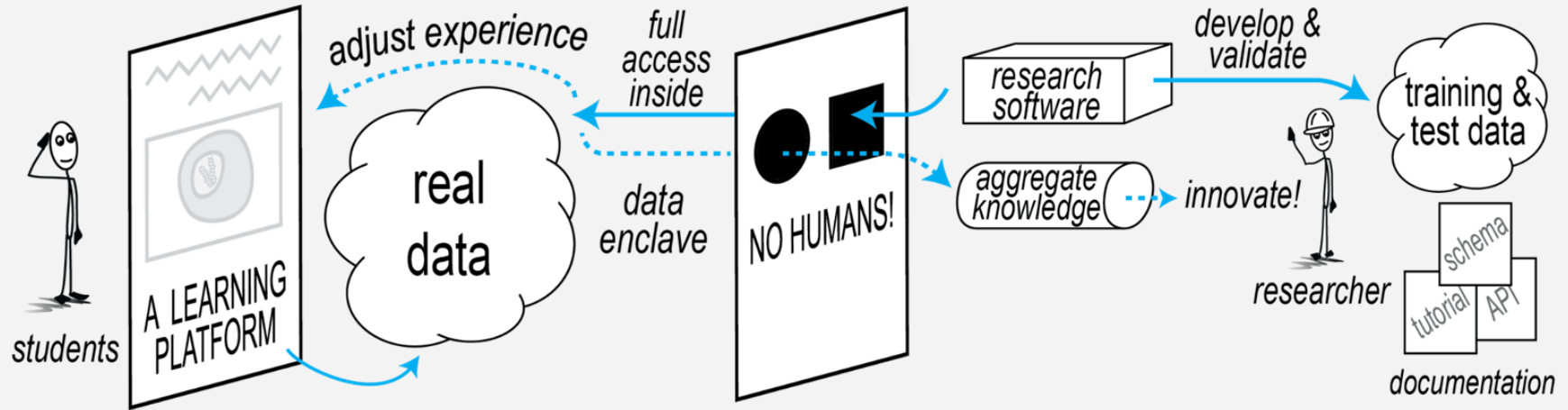
Embedded Interventions

Embedding interventions within
digital learning platforms

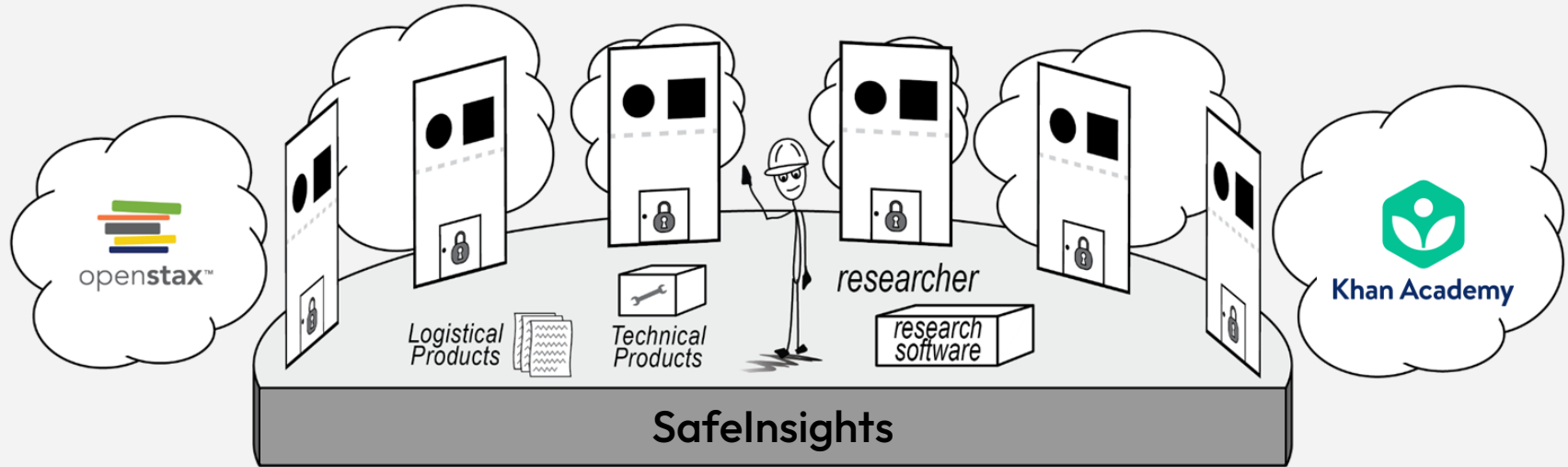
Enclave Fusion

Research with multi-platform
data sources

Safelnsights brings questions to the data

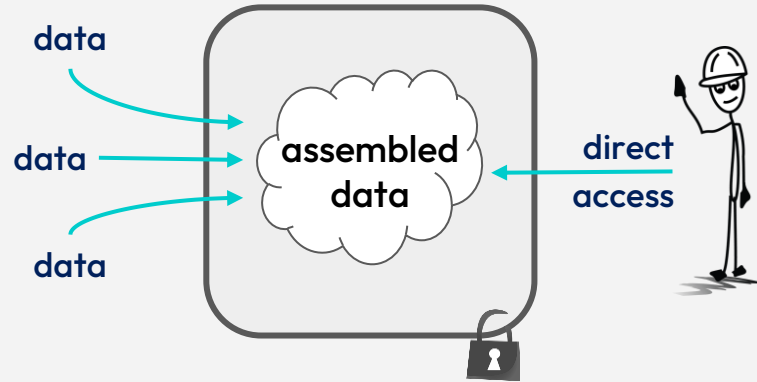


Safelnsights is a growing network of learning platforms



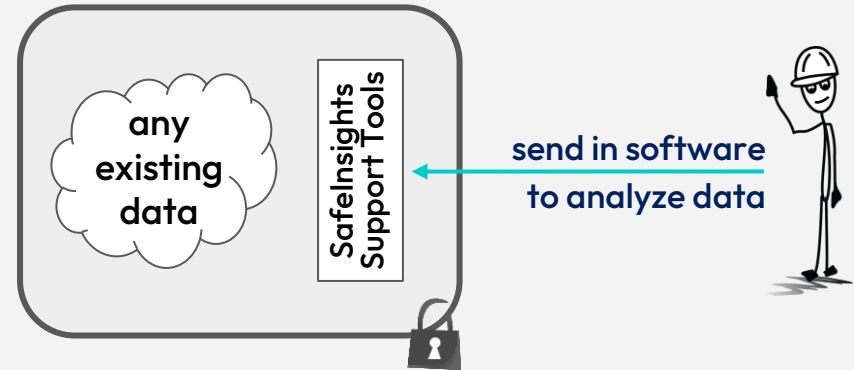
Other Enclave Tech

Data brought to a central, protected environment, shared out from there. ADRF, NORC, etc.



Safelnsights

Safelnsights is installed where data is protected today; researchers do not have direct access.



Safelnsights can be used wherever there is existing data, including other enclaves

The core idea:

SafeInsights brings research questions to the data

Built by the field, for the field to:

- Enable **large-scale** research about learning & education
- Preserve student **privacy**
- Make **sensitive data** available for research
- Keep the **data** where it is collected
- Convert research to **effective practice**



Baraniuk



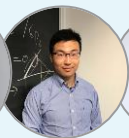
Roediger



Antón



Lan



Heffernan



Nosek



McNamara



Baker



Basu Mallick



Roschelle



Gosha



Castro Fernandez



Your code moves and the data stays in place

person

You

Bring your research question

Write analysis code against example data

security

Trust boundary

Allows reviewed code to cross in

Allows reviewed outputs to cross back out

lock

Secure enclave

Protects and secures data in place

Runs reviewed code against secure data

SafeInsights offerings when it opens vs. in the future

At opening

- Secondary analysis of authentic learning platform and institutional data from eight data partners
- Working with a single data partner per study

Coming up next

- Primary data collection (e.g., surveys, interventions)
- More data partners
- Cross-partner data merging

Data available on SafeInsights for secondary analyses

Fine grained and proximal

Broader and distal

Data range from

- **fine-grained learning processes to**
- **class and assessment performance to**
- **long-term educational outcomes**

Learning Process Data & Proximal Outcomes

OpenStax

Open textbooks & courseware

ASSISTments

Math homework platform

CourseKata

Interactive statistics courses

Khan Academy

K-12 & college STEM learning

Quill

Writing & grammar practice

TERC

STEM curriculum & research

Clickstream, problem attempts, hint usage, time-on-task, mastery progression, ...

+ More edtech partners joining

Distal Outcomes & Administrative Data

Institutions of Higher Education (IHE)

Arizona State University

Enrollment, grades, degree completion

+ More IHEs joining

State and Local Education Agencies

South Carolina State Education Agency (SEA) with Education Analytics

Statewide assessments, graduation, accountability

+ More SEAs joining

+ More Local Education Agencies (LEAs) joining

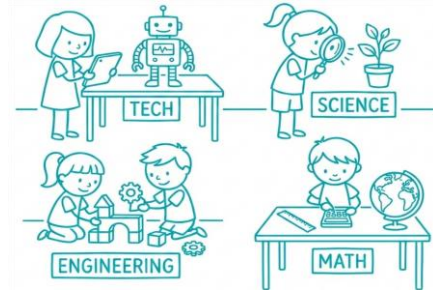
Research questions for NSF ECR and SafeInsights: Aligned by design

Is this a quest for
fundamental
understanding?

Is understanding **real-world**
use a priority?

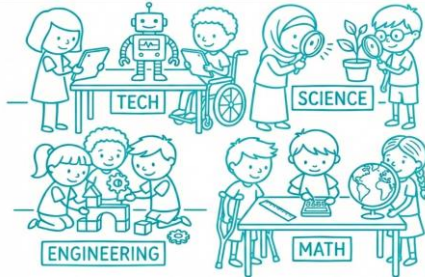
Response	No	Yes
Yes	Pure basic research	Use-inspired research
No		Applied research

Built to support the three ECR pillars



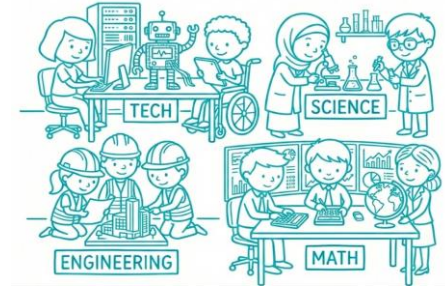
STEM learning at scale

The SafeInsights ecosystem includes data partners spanning kindergarten through post-secondary and lifelong learning levels, and representing different STEM learning contexts.



Broadening participation

For researchers: SafeInsights will support emerging researchers across different disciplines with scaffolded supports. **For students:** SafeInsights will enable researchers to complement learning process data with administrative data and investigate how students furthest from opportunities learn and thrive.



STEM workforce development

SafeInsights will represent data partners spanning full education-to-workforce trajectories.

Where we are and why now

May 2024

→ Awarded!

July 2026

→ Technical and governance builds in full swing for an October beta

October 2026

→ Operations ready
→ **Beta opens!**

02.

The walkthrough

What you will see in the walkthrough



1. Explore partner data and documentation



2. Propose a study to a data partner



3. Analyze data with AI support



4. View reviewed results

Welcome to Mars University



Welcome to Mars University ⁽²⁾

**Fully synthetic data that is realistically messy.*



SIS Records*

Enrollment, GPA, degree paths, and academic history.



LMS Activity*

Real-time clicks, logins, and engagement patterns.



Homework*

Problem-level attempts and cognitive step data.

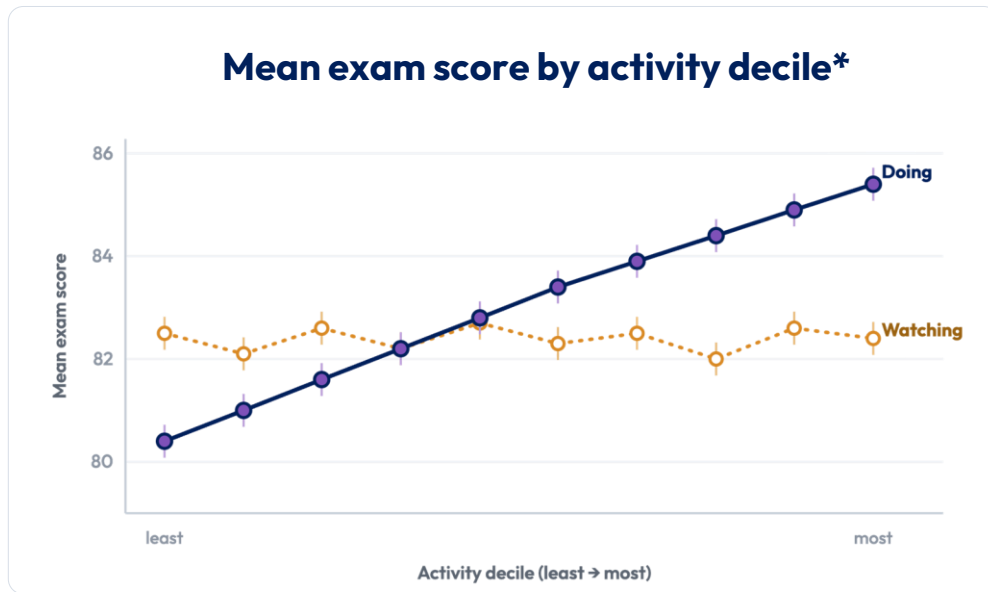


Census*

Neighborhood income, education, and social context.

The question: Does doing beat watching?

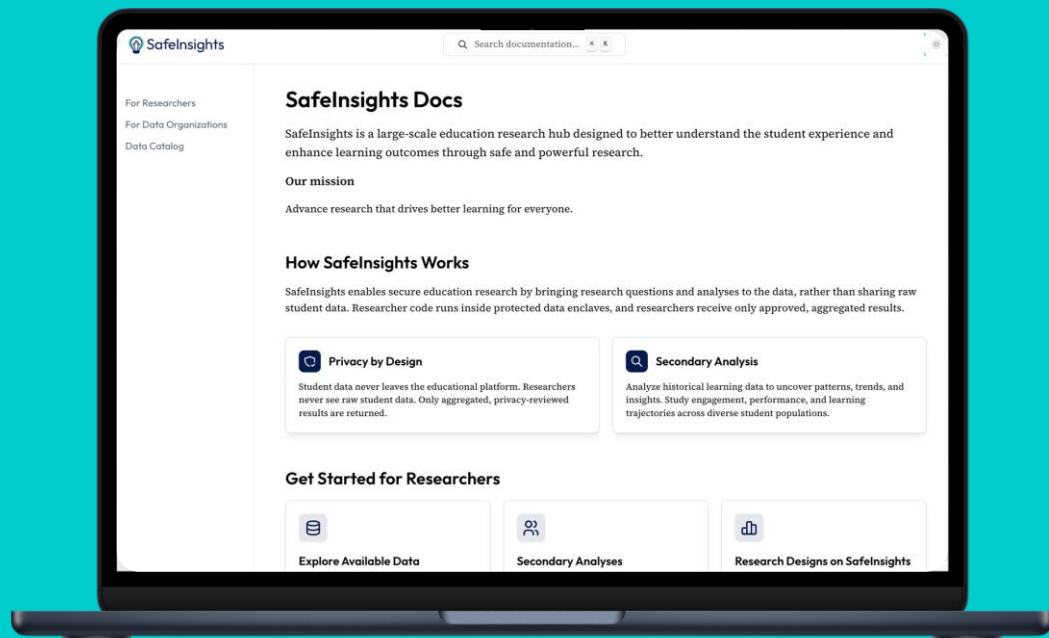
- This well established phenomenon in learning sciences was demonstrated by Koedinger et al., (2015). They found that **active practice** drives learning vs. passive video watching or reading.
- This research question is answerable by joining the practice engagement in the homework platform to the LMS gradebook



**Simple example of the Doer Effect*

Let's begin the demo

- [Documentation](#)
- [The App](#)



03.

Governance, agreements, and IRB

Two parties, one workflow

Researcher side

STEP 1: Submit research **proposal**

Your research question(s)

Proposed abstract, including hypotheses, variables required, and planned analyses

[View Data Documentation](#)

Submit Proposal

STEP 2: Develop **analysis code** (AI-assisted)

```
...  
mutate(date_completed =  
  date(first_completed_at)  
) |>  
  group_by(book_title,  
    teacher_id, task_type,  
    course_id |> ...
```

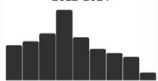
**AI-Assisted
Code Review**

**Test Code on
Example Data**

Submit Code

STEP 3: Receive **aggregated outputs**

**Final Exam Grades
2022-2024**



Instructor Experience	Mean	SD
0-5 Years	84.4	4.5
6-10 Years	87.5	4.9
11-15 Years	86.2	8.9
21-25 Years	88.9	11.4

**Download
Outputs**

- Explore data
- Propose a study
- Develop and test code
- Receive reviewed results

Data partner side

STEP 1: Review research **proposal**

Research Question: How does instructor experience impact final exam grades?

Abstract: We propose examining years of experience as a...

✓ Approve/Revise

✗ Decline

STEP 2a: Review **analysis code** (AI-assisted)

```
date(first_completed_at)) |>  
group_by(book_title, teacher_id,  
  task_type, course_id |> ...
```

Possibility of
small subgroup
sizes

✓ Approve Code & Run

Request Code Revisions

STEP 2b. Run approved code within DP enclave

Approved
Analysis Code

Data Partner
Enclave

STEP 3: Review **aggregated outputs**

**Final Exam Grades
2022-2024**

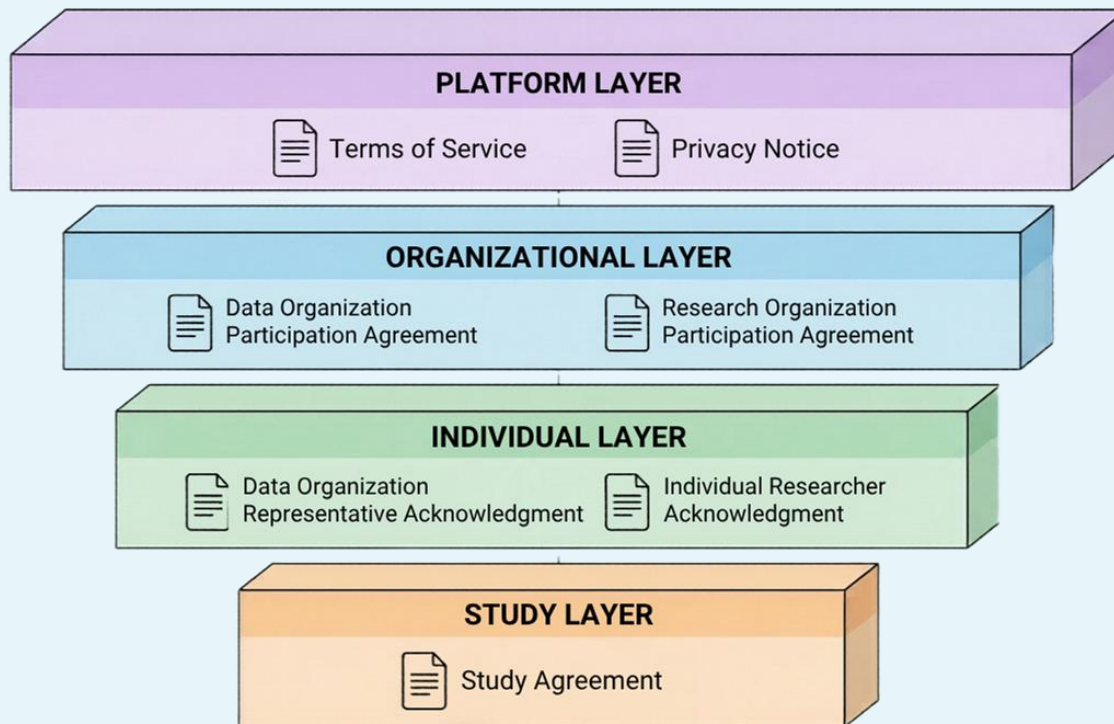


**✓ Approve & Release
to Researcher**

**Request Code
Revisions**

- Review proposal and make decision
- Approve code to run in the enclave
- Review outputs and release results

Tiered agreements ensure responsible data use



IRB pathways that work with your institution



The enclave is the privacy control

Researchers never access identifiable records; only aggregate, reviewed outputs leave the platform.

This architecture streamlines IRB oversight requirements.



Rice serves as the IRB of record...

Secondary analyses run under Rice's umbrella protocol.

You are not writing a protocol from scratch.



Reducing your home institution's IRB burden

Rice's umbrella protocol reduces the burden on your own institution.

We are putting resources in place to streamline the flow.

04.

Join the Beta

The beta at a glance

Dates

- **Oct–Dec 2026:** Internal researchers
- **Jan 2027:** External researchers

Who is the secondary analysis beta for?

- Researchers at different career levels and from a variety of institutions
- Researchers familiar with secondary analysis and quantitative methods
- Intermediate coding proficiency in R and/or Python

What you get from participating in the beta, and what we ask



What you get

- Agreements and IRB support
- AI coding assistant
- A direct line to the SafeInsights team
- No cost during the Beta



What we ask of you

Now

- Fill out interest form

During the Beta

- Run a study end-to-end
- Respond to surveys
- Participate in interview(s) for a modest honorarium

Three ways to engage



Express interest

Visit safeinsights.org

Fill out the interest form now



Follow along

Keep up to date with SafeInsights updates via our newsletter



Apply to participate

The Beta application will open in September

Decisions made on a rolling basis

05.

Scope your own study

Scope your own study



bit.ly/4gAseza

Let's take a few.

The QR code takes you to a Google doc that has a few questions before we get to our biggest question for the day.

“What question about student STEM learning would you dive into with rich platform and/or institutional data?”

Think about which data types appeal to you most, your questions, and the subgroup(s) you would most want to explore.

Wrap up

Privacy by design

With SafeInsights, we are introducing a privacy-preserving approach to conducting secondary analyses on educational data sets.

Beta access

We encourage researchers to express interest and apply for beta access, which will begin in January 2027.

Community alignment

SafeInsights is aligned with the ECR community's research needs and enables exploration of STEM-focused research questions.

Coming up next

Ability to conduct interventional research on learning platforms, as well as cross-platform research.

Thank you.

Questions?



Beta interest
form

<https://bit.ly/4p41UA1>

safeinsights.org