IRIS Incorporated Research Institutions for Seismology

2022 IRIS Evaluation of Project/Program Evaluations

Prepared for IRIS
Dr. John Taber
Mr. Michael Hubenthal

August 2022

By
Hilarie B. Davis, Ed.D.
Technology for Learning Consortium, Inc.

Table of Contents

I. Background and Purpose

The EPO assessment strategy relies on the following three core principles.

- Each activity in its portfolio, regardless of its scale, will be assessed.
- The rigor of assessment will correspond to the context and scope of each product or program.
- A combination of both internal and external evaluations will be employed to optimize EPO operational flexibility and value vs cost.

These principles underpin the EPO program's use of primarily internal evaluations, plus a periodic external "evaluation of evaluations". This approach provides important oversight of the quality and rigor of the evaluations employed, while also providing flexibility, local knowledge (context and operations), availability, etc. afforded by the internal evaluation process. The process is based on the Collaborative Impact Analysis Method (Davis and Scalice, 2015) and involves consultations, where the project lead for each project or program, or the internal elevator, reviews the evaluation plan and data currently in place jointly with the external evaluator. Together they score the project's evaluation using a qualitative rubric (e.g. Figure 1) based on best practices. The outcome of each consultation is a snapshot rating of the evaluation currently in place, a judgment on how well EPO has balanced the degree of rigor of the evaluation relative to the scope and scale of the project, and recommendations, as appropriate, for ways to enhance or expand the evaluation if deemed necessary. IRIS first employed this approach in 2015 when it embarked on an effort to place a greater emphasis and focus on evaluation of its work. The goal was to make evaluation an integral part of IRIS EPO staff's work and equip staff to be able to state why they do the activities they do (needs assessment), and make evidence-based claims about their work (impact) through evaluation processes. This forward-looking, capacity-building approach was particularly impactful for IRIS because the state of evaluation within EPO varied widely at the time. The consultation process was repeated again in 2017 building on the prior work in 2015. This time the consultations were conducted with all projects and also included having staff actively involved in the analysis of the quality and rigor of their products/projects beginning with the reason for the project or needs assessment, the goals and objectives (SMART=specific, measurable, action-oriented, realistic, time-based), a research-based design, the fidelity of the implementation, and the evidence of impact on behavior, attitudes, skills, interests, and knowledge (BASIK).

IAM Project Lifecycle Rubric

Indicating rigor level of evaluation practices for each project phase

Project Phase	Fair (1)	Good (2)	Very Good (3)	Excellent (4)
Needs Assessment What is the evidence of need?	Prior experience; "Seems like a good idea"	Research on what works; Literature review on similar programs/ products/ populations/ goals	Conversation with and/or direction from stakeholders (Focus Group); Experts review the ideas/plan	Survey of or pilot with potential audience/ users about the draft program
Goals and Objectives How measurable are the goals and objectives?	General direction; Understood by team; Agenda substituting for objectives	Explicit, written; For a target audience	Objectives are SMART: Specific, Measurable, Action-oriented, Realistic, Time-bound	Logic model of inputs, outputs, and outcomes in place
Design of Project How evidence- or research-based is the design?	Series of activities; Uses what has worked before	Based on objectives; Connects to standards; Includes contingency plans for emerging needs	Thematic; Has continuity; Participatory, personalized, responsive; Uses advanced organizers	Developmental; Embeds evaluation/ reflection
Implementation How true to the design is the implementation? (fidelity)	Facilitators prepare to implement the design	Collect and use feedback during implementation	High fidelity to design OR implements contingency plans to meet objectives if needed	Participants able to monitor their own progress against objectives
Outcomes Assessment/ Methods What is the evidence of impact on BASIK?	Post only survey or reflection; Follow up survey or interview; Web stats; Anecdotes; Facilitator reports	External evaluator observes, or does case studies; Pre/post self-report survey, reflections; Post only measure (test, retrospective survey, task)	Pre/post measures (tests, performance tasks, observation); Pre/post follow-up	Comparison group studies (quasi-experimental); Experimental study (random assignment)

Davis and Scalice, 2015

Figure1: Impact Analysis Method Levels of Rigor.

Results from the 2015 Evaluation – The process began at the annual staff retreat where EPO staff discussed the current state of the evaluations and rated their comfort level in four areas: formulating claims for my project (6.1/10), incorporating evaluation into my work (6.4/10), using evaluation data to make plans and decisions (6.9/10), and supporting evaluation efforts for my projects (7.1/10). Consultations, as described above, were then conduced for a total of 22 projects from the EPO program's portfolio of work. Using the rubric, the average rating for these projects/programs was a 1.6/4. Importantly, each project received recommendations for improving their evaluation methods and measures to make them more robust and rigorous.

Some examples of the evaluations developed post consultation included the following. The Seismic Waves website, a browser-based tool to visualize the propagation of seismic waves from historic earthquakes through Earth's interior and around its surface, was developed with expert review and beta-testing with potential users. Users of Teachable Moments which provides slide shows and information within 24 hours of an earthquake, were surveyed about how they use it and its effects. The Field XP experience had a full external evaluation to determine best practices for supporting field experiences in seismology in the future. The PIs, students, and seismologist mentors were interviewed. This provided data on the structure, content, and overall experience and recommendations for improving the experience. The GSA booth was evaluated for general awareness of IRIS, IRIS resources, and the IRIS login. For the Research Experience for Undergraduates, a plan was developed to validate a self-reflection tool.

Results from the 2017 Evaluation – To measure change since the 2015 consultations, the evaluations of 25 products/projects were similarly reviewed in 2017. This effort found that the overall robustness and rigor of the evaluations had improved, with an average rating of 2.5 out of

4.0. Again, each project received additional recommendations for improving their evaluation methods and measures to make them more robust and rigorous

Some examples of enhancements made to program/project evaluations based on the results of the consultations included the following. At the National NSTA meeting a brand recognition survey was developed and deployed, and unique URLs used to track follow-up engagement by attendees. These efforts provide key feedback to shape future booth design and activities at future NSTA events. For InClass, an educational resource web portal, new tools were developed to track download data to judge usefulness to educators. The Early Career Investigators program developed and implemented a needs assessment, and surveys for IRIS sponsored events (luncheons, webinars, colloquium series). Animations were evaluated based on newly available download data, along with the development of surveys to measure educators, attending IRIS workshops, intent to use them and their ability to integrate them into core undergraduate classes. YouTube statistics on likes, shares, and subscribers were also formally tracked.

II. Design

This evaluation builds on prior work to examine the evaluation practices of the EPO program across its portfolio. Specifically, it has been designed to address the following questions:

- What is the state of evaluation across the projects (robustness)? Are project objectives designed and evidence provided for needs assessment and design? Are outcome objectives measurable and appropriate measures of impact used?
- Are there any gaps or places where elements of evaluation are lacking? How can they be improved?
- Are best practices in evaluation employed when designing new products and implementing projects?
- How has EPO continued to improve its evaluation practices and incorporate evaluation into its work? What will help them improve moving forward?

In 2022, IRIS staff identified 34 projects/products to be evaluated and the phase of each product/project (see Project Phases Below). The phases are useful as a way to identify appropriate and rigorous evaluation methods and measures. Products/projects in the idea, prototype, pilot phases were identified for consultations, as described above.

Project Phase Definitions

- *The idea phase describes the early steps in an innovation process that involves brainstorming and filtering a large number of ideas until the best ideas emerge.*
- *The prototype phase is the creation of a model designed to solve a problem, meet a need, or validate ideas that can then be tried out, or piloted.*
- *The pilot phase is an initial small-scale implementation \ used to test the viability of a project idea. The project is often refined based on feedback collected in this phase.*
- *The implementation phase is the time during which a tested (piloted) design for an activity, project, or program is put into practice and data collected to establish if the implementation has the intended effects.*
- *The scale-up phase is the effort to expand the utilization of an implemented activity, project, or program, approaches are often based on what was learned in implementation phase.*

Products that were in the implementation or scale-up phases did not receive a consultation. In an attempt to streamline the evaluation process in 2022, the evaluations of these projects were reviewed through the use of a survey (see Appendix A), completed by all project leads. This survey asked about project objectives, the extent to which they were met, outcome objectives, evidence of outcomes and how the evidence was collected, as well as what NSF should know about this project. Projects leads were also invited to provide post reports, pictures, and other supporting documentation in a folder online for review. Follow-up questions were sent to project leads, as needed, to clarify information or to fill in any information gaps. Based on the information generated, the evaluation for each project was analyzed for the quality and rigor of the objectives, implementation, and collection of evidence.

III. Findings from 2022

In this section, there is a brief description of the 34 products/projects, an overview of the phases of the projects, ratings of their evaluation rigor, and recommendations.

III.A. Brief description of each project reviewed in 2022

Website	

InClass https://www.iris.edu/hq/inclass/search#language[]=1
An online portal to IRIS resources (e.g. activities, professional development workshops, posters, one-pagers, animations/video clips, data, software), in a consistent format, searchable by category and keyword, and offering related resources.

Teachable Moments https://www.iris.edu/hq/retm
Regular earthquake summaries within hours of the event with resources. Weekly publications on seismology during the pandemic.

Animations https://www.iris.edu/hq/inclass/search#type[]=1&language[]=1
IRIS has developed over 100 animations to illustrate fundamental concepts of seismology and earth science.

Women in Geoscience video series
https://www.iris.edu/hq/inclass/video/ifthen_women_in_geoscience_series Short videos/animations featuring women from around the world describing their work, interests, and desire to inspire young women

Seismology Skill Building	

Seismology Skill Building Workshop https://www.iris.edu/hq/inclass/course/ssbw
Workshop designed for undergraduate students (e.g.... computer science, geophysics, geology, math, physics, engineering) or recent graduates who will be starting graduate school in the fall of 2022 that want to develop scientific computing skills within a seismological context.

ROSES tech skills grad students https://www.iris.edu/hq/inclass/course/roses
Lessons are targeted towards advanced Ph.D. students and include video lectures and guided python notebooks (8 weeks).

Curriculum Development	

IGUaNA Modules https://serc.carleton.edu/iguana/index.html
Introducing geophysics for urban and near-surface applications - teaching materials feature urban environmental, engineering, and forensic questions that can be answered, at least in part, through geophysics.

Careers Module
To provide practical career resources for those who have an interest in or for those who are already studying the geosciences.

Summer REU undergraduates https://www.iris.edu/hq/internship/
Students spend 8 to 10 weeks working on a seismological research project with researchers at an IRIS member intuition deploying seismic instruments and/or analyzing seismic data to produce publishable results.

Anti-Harassment/ Discrimination Curriculum
https://www.iris.edu/hq/internship/anti_harassment_curriculum
The curriculum for undergraduate students about the terminology and concepts, policies and procedures regarding harassment, discrimination, and fraternization, how to respond and report an incident of discrimination or harassment

Determining and Measuring Earth's Interior (Online lab) https://www.iris.edu/hq/inclass/software-web-app/layered-earth
In this learning app, students compare observed seismic data to predictions they make from a model, to determine that the Earth must have a layered internal structure and to estimate the size of Earth's core.

InSight curriculum/Mars Monitor
IRIS is an Educational Partner on the InSight Mission, with the goal of engaging students with seismic data from Mars.

Professional Development	

Science Communication Workshops
A series of professional develop workshops aimed at building science communication capacity and confidence within the IRIS community.

Short format - NSTA, Smithsonian, AGI, STANYS, Mars InSight
IRIS facilitates at least 3 to 5 hour-long or longer workshops for educators annually.

U. of Alaska Online Course
In partnership with UA, IRIS developed and facilitated a semester long course for teachers to learn about Alaska relevant seismology and have activities they could use in their classrooms.

jAmaSeis - new web-based platform
The new web-based version will be web-based so users don't have to download the software.

Software/Web/Mobile Apps	

Seismic Monitor http://ds.iris.edu/seismon/index.phtml
The latest earthquakes on a map with news, lists, and links.

IRIS Earthquake Browser (IEB) http://ds.iris.edu/ieb/

Explore 5.4 million earthquakes on a Google map with ability to rotate thousands of quakes in 3D and export to Excel.

Seismic Waves http://ds.iris.edu/seismon/swaves/
An earthquake simulator that lets you watch how seismic waves radiate on the surface and bounce around inside of Earth while you drag to rotate.

Station Monitor https://www.iris.edu/app/station_monitor/
Provides access to continuous, real-time ground motion from hundreds of locations around the globe.

Global and Local Seismogram Viewer http://ds.iris.edu/gsv/ http://ds.iris.edu/lsv/
The Global Seismogram Viewer automatically creates clear plots of seismograms of large earthquakes from stations around the world, displayed by distance …

Quake Catcher Network http://www.iris.edu/app/qcn/stations
Interactive software developed for hands-on education about Earthquake Seismology. There are several modules within QCNLive for learning about earthquake vibrations and where earthquakes occur. The software measures and plots real-time motions acquired from several types of internal and/or external Micro-Electro-Mechanical System (MEMS).

EQLocate/EQ Trilateration https://www.iris.edu/app/eq-locate/method
Interactive app that allows users to locate earthquakes (latest and historic) using real seismic data. Users conduct quality control on the waveforms, pick seismic arrivals, and apply one of the 3 location methods to find a solution they they then compare to the USGS accepted solution.

Social Networking/Science Communication

Social media (Twitter, FB, Instagram, Pinterest) https://www.iris.edu/hq/programs/epo/social_media
https://twitter.com/IRIS_EpO https://m.facebook.com/IRISEarthquake/
https://www.instagram.com/iris_epo/?hl=en https://www.pinterest.com/IRIS_EPO/
Twitter is https://twitter.com/IRIS_EpO FB sites – Earthquake science, undergraduate interns, early career investigators, and research/facilities. Follow IRIS on Pinterest for ed resources, animations, and lessons. Instagram is @iris.epo.
TikTok https://www.tiktok.com/@terraexplore
This project is reaching a critical demographic for work-force development - 13-19 year-olds. Additionally, the reach and impact for the amount of time spent is impressive.

Science Ambassador Program
The science ambassador program will replace the Distinguished Lecture Series with regional communicators who are trained and matched with media, institutions, conferences, and other opportunities.

Classroom (K-16) presentations/visits
IRIS offers talks upon request on content, science communication, or as a role model

Public Display Support (Earthquake Channel)
https://www.iris.edu/hq/programs/epo/museum_displays/eqc/
Shows earthquakes on high resolution maps, suitable for display in a museum, lobby, visitor center or school setting. Earthquake updates are every 10 minutes, while the map changes every 30 seconds.

IRIS/SSA Distinguished Lectureship https://www.iris.edu/hq/programs/epo/distinguished_lectureship

The program offers on-technical presentations on earthquake science topics to general audiences across the US. Each speaker is an expert in his/her specific research area and is skilled in effectively communicating new and exciting findings to the public

NSTA
IRIS has attended NSTA every year since at least 2002.

Booths at GSA, AGU, SSA
The IRIS booth in the exhibit hall at professional geoscience meetings (AGU, GSA, SSA) and at outreach events provides an IRIS presence and serves as a source of information about seismology, earthquakes and programs implemented by IRIS and funded by NSF.

Enhanced Dissemination of Materials
The goal was to expand the impact of IRIS in the K-12 education community by producing a monthly e-newsletter, distributing press releases, and surveyed middle and high school teachers.

Early Career Researchers (ECR)
Support for ECR through a Facebook page, skill building workshops, networking at conferences/meetings

Webinar series https://www.iris.edu/hq/webinar/
IRIS offers interesting and informative webinars for the community.

III.B. Overview of Project Phase
Thirty-three projects were evaluated by the external evaluator. The number of projects by phase is shown in the table below. The majority of projects are in the implementation phase (63%).

	#	%
Idea	0	0
Prototype	3	9%
Pilot	5	14%
Implementation	22	63%
Scale-up	4	11%

Products/Projects by Phase
Idea – none currently

Prototype
- EQLocate/EQ Trilateration
- jAmaSeis - new web-based platform
- Science Ambassador Program
- ROSES tech skills grad students
- Seismology Skill Building Workshop Undergraduate - Research

Pilot
- Careers Module
- Enhanced Dissemination of Materials
- TikTok

Implementation

- Animations
- Anti-Harassment/Discrimination Curriculum
- Booths at GSA, AGU, SSA
- Classroom (K-16) presentations/visits
- Determining and Measuring Earth's Interior
- Early Career Investigators
- Global and Local Seismogram Viewer
- IGUaNA Modules
- InSight curriculum/Mars Monitor
- IRIS Earthquake Browser (IEB)
- IRIS/SSA Distinguished Lectureship
- NSTA
- Public Display Support (Earthquake Channel)
- Quake Center Network
- Science Communication Workshops
- Seismic Waves
- Seismology Skill Building Workshop Undergraduate
- Short format PD - NSTA, Smithsonian, AGI, STANYS, Mars InSight
- Social media (Twitter, FB, Instagram, Pinterest)
- Summer REU Site
- Teachable Moments
- University of Alaska Online Course
- Women in Geoscience Video Series

Scale-up

- InClass
- Seismic Monitor
- Station Monitor
- Webinar series

III.C. Ratings for Project Objectives and Outcome Objectives

The table below rates each project on a scale of 1-4 with 4 being the highest for defining objectives and measuring their impact.

- Project Objectives Defined (POD) – What will the project do? and met (POM) Is appropriate and rigorous evidence collected of how well the objective was met?
- Outcome Objectives Defined (OOD) - Does each project have measurable outcome objectives?
- Outcome Objective Evaluation (OOE) – Is the design appropriate for the project/product? Is there a data collection process in place to measure the outcome objectives and is the scale of this process appropriate for the scope of the project/product?

Ratings 1-4, 4 highest

	POD	POM	OOD	OOE	Ave
Website					
• InClass	4	4	4	4	4.0
• Teachable Moments	4	4	4	4	4.0
• Animations	4	4	4	4	4.0
• Women in Geoscience Video Series	4	4	4	2	3.5
Seismology Skill Building					
• Seismology Skills-Building Workshop	4	4	4	4	4.0
• ROSES tech skills grad students - Prototype	2	0	2	3	1.8

Curriculum Development					
• IGUaNA Modules	4	3	4	4	3.8
• Careers Module - Pilot	4	3	4	3	3.5
• Summer REU Undergraduates	4	4	4	4	4.0
• Anti-Harassment/Discrimination Curriculum	4	4	4	4	4.0
• Determining & Measuring Earth's Interior	4	4	4	4	4.0
• InSight curriculum/Mars Monitor	4	3	4	3	3.5
Professional Development					
• Science Communication Workshops	4	4	0	0	2.0
• Short format PD - NSTA, Smithsonian, AGI…	4	4	3	4	3.8
• University of Alaska Online Course	4	4	4	4	4.0
Software/Web/Mobile Apps					
• jAmaSeis - new web-based platform -Prototype	3	1	4	4	3.0
• Seismic Monitor	4	4	4	3	3.8
• IRIS Earthquake Browser (IEB)	4	4	4	4	4.0
• Seismic Waves	4	4	4	2	3.5
• Station Monitor	4	4	4	4	4.0
• Global and Local Seismogram Viewer	4	4	4	4	4.0
• Quake Catcher Network Viewer	4	4	4	2	3.5
• EQLocate/EQ Trilateration - Prototype	4	0	4	4	3.0
Social Networking/Science Communication					
• Social media	4	4	4	3	3.8
• TikTok - pilot	4	1	4	4	3.3
• Science Ambassador Program - Prototype	4	0	4	4	3.0
• Classroom (K-16) presentations/visits	4	4	4	3	3.8
• Public Display Support (Earthquake Channel)	4	4	4	4	4.0
• IRIS/SSA Distinguished Lectureship	4	4	4	4	4.0
• NSTA	4	4	4	4	4.0
• Booths at GSA, AGU, SSA	4	4	4	4	4.0
• Enhanced Dissemination of Materials	4	4	4	4	4.0
• Early Career Researchers - Pilot	4	4	0	0	2.0
• Webinar series	4	4	4	3	3.8
Average rating by element	3.9	3.4	3.7	3.4	3.6

III.D. Overall Ratings and Observations by Project

An average was obtained from the ratings for identifying objectives and measuring the extent to which they were achieved. Recommendations are given in detail in the summaries in Appendix B.

Project	Rating	Observations
Website		
• InClass	4.0	Suggest user study to identify how they find out about it
• Teachable Moments	4.0	Suggest case studies of frequent users to develop a range of use cases
• Animations	4.0	Better documentation of needs assessment of ideas for animations – who will use it, for what, what need does it fill?

• Women in Geoscience Video Series	4.0	Suggest pilot studies to establish impact and develop use cases
Seismology Skill Building		
• Seismology Skill-Building Workshop Undergrad	4.0	None This project is moving into a research phase to deepen an understanding of what works
• ROSES tech skills grad students	1.8	Prototype, design and measures tbd; Faculty doing need assessment now
Curriculum Development		
• IGUaNA Modules	3.8	Student data coming summer 2022, 4th module under development, outside evaluator, use by more professors
• Careers Module - Pilot	3.5	Expect to complete module and pilot this summer, have external evaluator
• Summer REU Undergraduates	4.0	None. Continue using alumni census approach to long-term tracking
• Anti-Harassment/Discrimination Curriculum	4.0	Suggest follow-up with faculty on effects on their programs through surveys or interviews
• Determining & Measuring Earth's Interior	4.0	Suggest a user study.
• InSight curriculum/Mars Monitor	3.5	Expert reviews, pilot other two modules, finish 4th module
Professional Development		
• Science Communication Workshops	2.0	Need headcounts for workshops and panels and evaluation feedback from participants
• Short format PD - NSTA, Smithsonian, AGI…	3.8	None. Continue collecting head counts and survey feedback
• University of Alaska Online Course	4.0	Suggest follow up survey with 2017 and 2021 participants
Software/Web/Mobile Apps		
• jAmaSeis - new web-based platform Prototype	3.0	Prototype phase; evaluation is planned of how hosting stations can stream and how educators use the data
• Seismic Monitor	3.8	None
• IRIS Earthquake Browser (IEB)	4.0	Need measurable outcome obj; need more data on effects on different audiences
• Seismic Waves	3.5	Need Webstats and classroom study of effects
• Station Monitor	4.0	None
• Global and Local Seismogram Viewer	4.0	Suggest user studies on effects
• Quake Catcher Network Viewer	3.5	Need evaluation of upcoming Sept teacher workshop
• EQLocate/EQ Trilateration - Prototype	3.0	Will be piloted to collect evidence of effects
Social Networking/Science Communication		
• Social media	3.8	No recommendations
• TikTok - pilot	3.3	Pilot phase – appropriate evaluation to date
• Science Ambassador Program	3.0	Prototype phase - appropriate to date
• Classroom (K-16) presentations/visits	3.8	Define model, embed questions in presentations to see changes
• Public Display Support (Earthquake Channel)	4.0	Survey sites on how they use the displays
• IRIS/SSA Distinguished Lectureship	4.0	Phasing out – no recommendations
• NSTA	4.0	Follow up with booth visitors to see what resources they have used, if any, every other year would help refine approach.

• Booths at GSA, AGU, SSA	4.0	Webstats before and after a conference could also provide evidence that the booth is driving people to the website.
• Enhanced Dissemination of Materials	4.0	Continue to track newsletter recipients and media releases; follow up with some newsletter recipients to identify nature and extent of use
• Early Career Researchers - Pilot	2.0	Continue to track participation and survey participants about they learned.
• Webinar series	3.8	Ask for characteristics of participants in sign up form. Use a short exit survey value of webinars for them.

IV. Discussion and Recommendations

Based on the analysis of above, each type of product and project has best practices for evaluation to inform it throughout the phases of idea to scale-up. IRIS EPO has committed to using these best practices in its work to offer the most effective resources and services possible and has offered evaluator support and information for staff to do it.

Idea Phase
For products and projects in the idea phase, a needs assessment is critical. How did this idea arise? Which stakeholders need it (a new resource or a revised old one)? What needs to be changed? Why does this need to be added to IRIS EPO offerings? How do you know? For example, IRIS/SSA's Distinguished Lectures program has reached a fairly large audience in face-to-face terms. However, these sorts of public talks seem to becoming less popular, and offered by fewer venues. Thus, the program had been slated for sunsetting. The need that emerged from evaluations of the lectureship program was to reach even more people with Subject Matter Experts (SME) across the US and through media by region. From this the Science Ambassadors program was conceived and explored to develop seismologists' communication skills and contacts. From this need, a concept paper was written, discussed, and approved by the IRIS's oversight committee. The project is now just beginning to transition out of the idea phase and into the prototype stage in which the model will be developed and again reviewed for efficacy.

Prototype Phase
Evaluations of products/projects in the prototype phase will be designed to answer questions such as, "Will it meet the identified need? Is it a reasonable and powerful approach to achieve the desired impact?"

For example, the EQLocate self-directed online tool was built to respond to the need to update the 20-year old Virtual Earthquake tool for current technology and the method of locating earthquakes. The design was reviewed by teachers and was piloted this summer (2022). This exemplifies the concept of evaluation in the prototype phase etc. where feedback is sought from end users and preliminary data collected on impact.

Several projects are making use of feedback from users in design as well literature reviews to identify best practices.

Implementation Phase
Evaluation of projects in the implementation phase focus on the expanded use from pilot testing and the effectiveness demonstrated in the original testing. Most (63%) IRIS EPO efforts are in the implementation phase. These include longstanding efforts that could be considered infrastructure, such as creating animations across programs, hosting booths at conferences to increase awareness of IRIS resources, and Teachable Moments that produce slide shows and identifies resources right after an event. Evaluation of these is based on use or users, but could include periodic follow up with users to get feedback on they are using the resources and ideas they have for improvement. Use cases not only provide more in-depth evidence of effective use, but can serve as models for new users. Having an advisory board of users for longstanding programs can provide ongoing feedback and identify pivot points for upgrading or revising programs.

Also in the implementation phase are professional development or curriculum such as the University of Alaska online course for teachers which has now been offered twice with modules on the InSight mission and using the IRIS tool Mars Monitor. The online course environment has built in evaluation of what teachers are learning. An end-of-course survey asks teachers about their intent to continue to use what they learned (they did activities with their students throughout the course). This approach should be applied even in 'short-form" professional development or Science Communication Workshops, building in reflection and feedback is important to refining the presentations and providing evidence of effects. What ah-has did they have? How will this experience affect their professional activities going forward? What would they like to know more about?

The IGUaNA modules used best practices from idea to implementation. There is gap in applying geophysical techniques to societally relevant, real-world problems in introductory courses. Faculty were asked about how to fill this gap from which a rubric and goals were written to guide the development of four modules. Each of the modules was peer reviewed and tested in undergraduate classrooms, then revised. Pilot testing showed effects on student engagement, interest, and knowledge of geophysics.

Scale-Up Phase
Projects in the scale-up phase are projects being expanded to reach more audiences. Best practices in evaluation in this phase require identifying potential users or audiences and determining from them what they need to implement the program or use the product effectively – to get the same results as the original implementers? What are the essential elements for success? How will IRIS help them determine their capacity for implementing? What support will they need to implement the program or product in the way it was intended and has previously been successful? Ongoing evaluation can provide evidence that scale up can work and under what conditions. Consider Seismic Monitor, which is a webtool that allows users to monitor global earthquakes in near real-time, visit seismic stations around the world, and search the web for earthquake headlines or region-related information. Although Webstats show over 4 million unique pageviews with an average time on page of two minutes, what brings people to the site and keeps them there long enough to benefit has not yet been studied. For scale-up of this long standing tool, user studies need to be conducted more consistently. An example of this that IRIS has implemented was the recent evaluation of current IEB users in which they

collected data on why visitors come to IEB (to learn about earthquakes in general or see local effects of interest), the ease of use, and the most useful feature.

V. Conclusions

Comparisons between 2015, 2017, and 2022 evaluation of evaluations
Improvements in evaluation design and their associated results were seen between 2015 and 2017 and from 2017 to 2022 In 2015, 22 projects were analyzed and had an average, baseline rating of 1.6/4.0. In 2017, the evaluations of 25 projects and programs were assessed. Here, it was found that the average rating had increased to 2.5/4.0. This included nineteen of the 2015 projects which showed a positive change of 37% .

As reported above, in 2022, 34 projects were analyzed. Nearly half of these projects (N=15) were evaluated in each 2015 and 2017. The average rating for all projects was 3.6/4.0 or a 44% increase over the overall average of 2017. This is notable as this average includes 8 projects in prototype and pilot projects (N=8 or 23%). These alone had an average of 3.5/4 and would be expected to have a lower score since their data on effects is limited.

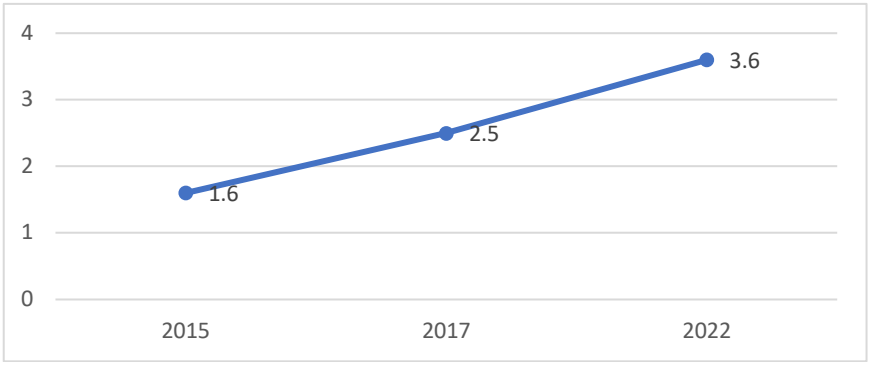

Figure 2: Overall average evaluation rating for the EPO portfolio as reviewed in 2015, 2017, and 2022.

As evidenced by the overall trends in each of these three evaluation reviews, many of the recommendations made each year are being implemented, such as headcounts and quick surveys for short events, more formal evaluations of effectiveness of longer or more intense activities, and evaluation of presentations for effects on interest, intent to use something from the presentation, or interest in further opportunities.

Overall, how is the quality/rigor of the evaluations being implemented now?
As discussed above, overall, the quality of evaluations is very high. All the projects have measurable objectives and have collected evidence of how well those objectives have been met. Getting feedback from an advisory board of users, or super users, or newbies could provide important additional evidence of what works, how, and why that informs upgrades/revisions and scale-up.

How is the rigor balanced with the scope/scale of the projects across the portfolio?
The Outcomes Objective Evaluation rating is about the appropriateness of the evaluation. It takes into account the duration and intensity of programs and balances that against the resources required to implement an evaluation. Shorter programs or use of products will not have as great of an effect, or require as rigorous an evaluation. Almost all (29/34) projects have appropriately designed evaluation efforts. Two examples where evaluations could be improved are software applications where there is not currently an evaluation plan in place.

How could EPO improve evaluation across the portfolio?
While the EPO program has shown continued growth in its approach to evaluation, there are places and processes that the program should consider focusing on to continue this journey.

Recommendation 1: These multi-faceted efforts need to be documented to understand how, or if, they come together to make IRIS resources impactful for each of their audiences. Just as InClass provides a portal for accessing IRIS resources, perhaps bundles of lessons and tools could be assessed by audience, subject, level, or course to allow for a better understanding of how users are combining multiple resources and what impact such combinations are having. Such an enhanced evaluation could be conducted in a setting where IRIS tools are used consistently, say in a classroom over a semester, so the knowledge and skills of the students could be assessed.

Recommendation 2: IRIS EPO has gotten good at identifying needs and responding to them with prototypes that are then piloted to test their effectiveness. Evaluation in the implementation phase is a place where EPO should look to invest additional effort in the future, especially over time for long-lasting projects. There could be more user studies over time, follow-ups with super users, or one-time users, and a survey of the array of products and services in the field that are evolving in the field to keep IRIS EPO current and expanding. For example, these would help answer questions such as: How can users rate their experience with a tool? Where do they tell their stories of implementation and effects? How are they recruited for more in-depth studies of impact? Do they feel part of an IRIS community and so are invested in using, improving, and recommending products and programs?

Recommendation 3: The review of the evaluation plans showed there is room for improvement with some project leads in some areas. It is not uncommon for people who are very good at implementation to be less focused on measurable objectives and data collection for evidence of success. They often value and use informal feedback but are not systematic about collecting data on the impact of their efforts. The first step is to have measurable objectives and support for non-intrusive, effective ways to measure success. In this review, as in 2017, project leads were given feedback and coaching on how to improve their objectives (make them SMART) and how to collect data of effects. Yearly internal reporting on each product/project's evaluation methods and results would support building further capacity of the staff as they learn from each other. On-demand consultations and targeted trainings with an external evaluator could continue to help individual staff improve their evaluation designs and use of results. An external evaluator could review the yearly reports every two years. Ultimately, the most effective evaluation is embedded in the work or the product, so it informs revisions and collects evidence of effects.

Appendix A: Evaluation Data Collected on Each Product/Project

Email
We'd like to establish the extent to which each project has met its project and outcome objectives and the evidence for results. We'd also like to hear about the progress you have made since the last evaluation in 2017. The results will be used for the NSF review in September as well as the merger talks.
Description (optional)

Project Name
Your Name

Project Stage: The stage of your project (check one - idea, prototype, pilot, implementation and scale-up)

In this next section, you will have the opportunity to describe what you are producing or offering (process objectives) and what you get from it (outcome objectives).

PROCESS OBJECTIVES are defined as what you do, develop, or present. This includes modules, presentations, websites, workshops, courses, etc. State your project objective with the audience, deliverable and timeline:

Describe the extent to which you accomplished this process objective and the EVIDENCE of that accomplishment (link to the product or documents in the project folder on the google drive - https://drive.google.com/drive/folders/1XFbit0UY6t-Bux3NdVEpwlfOi6k4Rbs)

OUTCOME OBJECTIVES are defined as what you get for what you create, i.e. the impact it has on audiences, IRIS (or other organizations) or the field. State your outcome objectives with a description of the extent of the impact and the evidence you have of that impact. For example, if you created a program for undergraduates, how many did it reach (compared with how many you hoped to reach) and to what extent did it affect them and how do you know what they effect was?

OUTCOME OBJECTIVE 1

Describe the extent to which you accomplished the objective and the evidence of that accomplishment (with link or doc)

OUTCOME OBJECTIVE 2

Describe the extent to which you accomplished the objective and the evidence of that accomplishment (with link or doc)

OUTCOME OBJECTIVE 3

Describe the extent to which you accomplished the objective and the evidence of that accomplishment (with link or doc)

OUTCOME OBJECTIVE 4

Describe the extent to which you accomplished the objective and the evidence of that accomplishment (with link or doc)

What should NSF know about the nature and value of this project?

Additional comments you would like to make