

# NSF guidelines for research security analytics

Last updated February 2023

## Table of Contents

- Table of Contents..... 1
- 1. Summary..... 2
- 2. Foreword by the chief of research security strategy and policy..... 3
- 3. Review ..... 5
- 4. Relevant authorities and supporting documentation..... 5
- 5. Definitions ..... 5
- 6. Research security responsibilities and process of the Office of the Chief of Research Security Strategy and Policy..... 7
  - 6.1 OCRSSP research security responsibilities ..... 7
  - 6.2 Process for notification and communication with institutions..... 10
- 7. Monitoring and reporting by NSF offices and staff..... 11
  - 7.1 Terms and conditions compliance-monitoring responsibilities of program staff..... 11
  - 7.2 Vetting for employment..... 12
- 8. Permissible and prohibited practices for research security-related analytics by the CRSSP ..... 12
  - 8.1 Permissible approaches for research security analytics ..... 12
  - 8.2 Prohibited practices for research security analytics ..... 13
  - 8.3 Individual matching criteria for validation and information sharing activities..... 13
- 9. Data, services and methods used for research security analytics ..... 14
  - 9.1 Non-NSF data used in research security analyses..... 14
  - 9.2 Analysis criteria and purpose ..... 15
  - 9.3 Services used in research security analyses..... 15
- 10. Sharing guidelines for security-related information ..... 15
  - 10.1 Human oversight ..... 15
  - 10.2 Sharing of information with institutions ..... 15
  - 10.3 Sharing of information by OCRSSP with inspector general or federal agencies ..... 15

# 1. Summary

NSF guidelines for research security analytics is a public document describing NSF's internal guidance for research security data-related practices. It includes a breakdown of which agency personnel may conduct research security-related activities; what monitoring activities are allowed and with what resources they are conducted; how information will be validated to ensure accuracy; and how information may be shared within NSF and externally (summarized in Figure 1 below).

This report builds on extensive feedback from community stakeholders and establishes key principles for NSF's research security analytics activities:

1. Program staff are not permitted to use research security concerns as a determining factor in the merit review process.
2. All research security analytics activities at NSF will be conducted solely by the Office of the Chief of Research Security Strategy and Policy, or OCRSSP.
3. Program staff are not permitted to conduct intentional information querying activities related to research security. Concerns encountered during routine merit review activities (see "routine assessment" in definitions) are to be reported to OCRSSP.

The guidelines are consistent with information laid out in the [System of Records Notice on NSF-77 Data Analytics Application Suite, or SORN NSF-77](#), and are a part of NSF's efforts towards transparency of internal practices. The guidelines also relate OCRSSP and NSF requirements as described in the "[CHIPS and Science Act of 2022](#)," [National Security Presidential Memorandum 33, or NSPM-33](#), and its accompanying [implementation guidelines](#).

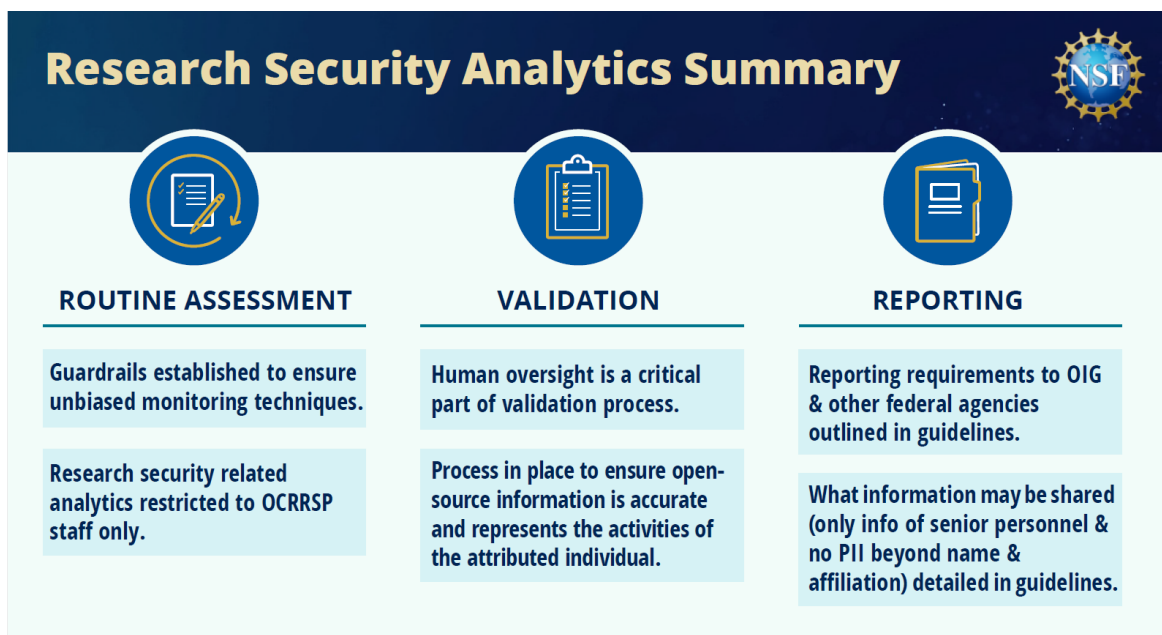


Figure 1. Research security analytics summary.

## 2. Foreword by the chief of research security strategy and policy

The U.S. National Science Foundation continues to uphold a centuries-long understanding that the expansion of scientific knowledge requires the world's talent. Reinforced by rigorous study and community experience, the outcomes of collaborative ecosystem are critical for excelling in NSF's mission of furthering the scientific enterprise for the United States' benefit. These gains are driven by tens of thousands of scientists who bring their expertise to our research labs to train, teach and contribute to our knowledge and economy.

In parallel, select practices by some governments have conflicted with our values and culture of open science. This has required the U.S. scientific enterprise adjust its policies and programs to ensure that international science continues to benefit the nation while minimizing risks to our economic and security interests. These efforts have resulted in thorough analyses being conducted by the community with a recurring conclusion reaffirming that **the United States continues to benefit from fostering a collaborative, welcoming environment for international science**. It is a leading priority for the agency to enable and support participation in collaborative research when conducted according to the principles described in the [National Security Presidential Memorandum 33, or NSPM-33](#), implementation: "openness, transparency, honesty, equity, fair competition, objectivity, and democratic values."

With these principles in mind, how can we prioritize and benefit from an open, collaborative system, while still safeguarding our activities from abuse? **Direct engagement and transparency with research community stakeholders is critical to enabling the honest dialogue needed to work as partners towards a harmonized goal**. A key first step in this process is assisting all participants in the research and grant management ecosystem as they work to adhere to the above principles. NSF has been a dedicated voice in the development of NSPM-33, providing guidance on the types of information required from researchers to support an open, transparent and honest research environment. Program staff and leadership at grant-awarding agencies will be able to rely on the NSF guidelines for research security analytics to help ensure they review proposals with attention to objectivity, equity and fair competition.

A key purpose of this document is to ensure that NSF program staff — the stewards of the merit review process — can conduct their work without the burden of geopolitics. We depend on program staff diligence to ensure we select the best science through a process that adheres to our research integrity standards. At the same time, given the complexity of research security issues, **NSF is requiring that no personnel beyond the Office of the Chief of Research Security Strategy and Policy engage in any research security analytics activities** (section 8). Agency staff encountering research security concerns during routine activities are expected to communicate concerns to OCRSSP and, when appropriate, to the Office of General Counsel, or OGC, and the Office of Inspector General. Verification of information is also a critical dimension of all research security monitoring activities; human oversight and verification standards are required for all monitoring activities by the CRSSP (section 10) and must be conducted before any information is shared (section 12).

These activities continue a legacy of openness and inclusivity around which NSF's mission was built. In 2019, NSF [commissioned JASON](#), an advisory body known for tackling difficult questions on matters of national security, with evaluating key concerns of the research security debate. Of note was the

continued validity of National Security Decision Directive 189, a policy directing federal agencies to ensure "the products of fundamental research remain unrestricted." This classification should be used when national security concerns require better control. JASON's finding that "NSF should support reaffirmation of the principles of NSDD-189, which make clear that fundamental research should remain unrestricted to the fullest extent possible," confirms that the tenets of NSF's activities still stand strong today.

There is a shared responsibility to maintain a framework that upholds the values of honest, transparent science. NSF continues its commitment to supporting open science and creating an inclusive, collaborative research enterprise that welcomes the immeasurable talent of the international community.

Rebecca Lynn Spyke Keiser  
Chief of Research Security Strategy and Policy  
U.S. National Science Foundation

### 3. Review

These guidelines will undergo iterative improvement based on the needs of the scientific community, the federal government and the national security community. Updates made to the guidelines will be published on NSF's [research security page](#).

### 4. Relevant authorities and supporting documentation

The NSF guidelines for research security analytics were created in response to the following policies requiring proper disclosure of appointments, affiliations and current and pending support for external funding sources:

- [National Security Presidential Memorandum 33, or NSPM-33](#): Published in January 2021, NSPM-33 directs Federal agencies to "strengthen protections of United States Government-supported Research and Development (R&D) against foreign government interference and exploitation." Its [subsequent guidance document](#), published in January 2022, provides further direction to departments and agencies regarding disclosure of conflicts of interest, digital persistent identifiers, consequences for violation, information sharing and research security programs.
- [The "CHIPS and Science Act of 2022"](#): Signed in August 2022, the act boosts federal R&D investments to strengthen U.S. scientific research and technological leadership. It authorizes and defines required activities by NSF's OCRSSP (section 10331; 42 U.S.C. 19031), covering analytics, information sharing and research security training.
- [Office of Management and Budget Memorandum 22-04](#): This memorandum provides guidance regarding agency actions needed to support the important role of agency inspectors general. Further guidance to NSF staff on compliance with OMB M-22-04 is detailed in NSF's Office of the Director Memorandum 22-03.
- NSF's [terms and conditions](#): Each NSF award notice specifically identifies certain conditions that apply to, and become part of, that award. The [award conditions](#) are available electronically on NSF's website.
- [Proposal & Award Policies & Procedures Guide](#): The PAPPG is comprised of information relating to NSF's proposal and award process for the agency's assistance programs. The PAPPG is designed to set forth NSF's proposal preparation and submission guidelines, as well as NSF policies and procedures regarding the award, administration and monitoring of grants and cooperative agreements.
- [Dear Colleague Letter from former NSF Director France Córdova](#): Released in July 2019, this Dear Colleague Letter addressed to the academic community identifies emerging risks to the nation's science and engineering enterprise and actions NSF is undertaking to uphold the values of "openness, transparency, and reciprocal collaboration."

### 5. Definitions

\* Denotes a definition taken from [NSPM-33](#) or the [NSPM-33 implementation guidance](#).

- Advanced monitoring: Aggregate portfolio monitoring activities required for managing NSF programs and offices. Advanced monitoring often relies on data and analytics to collect, analyze and extract information to understand the progress of a program, division or directorate.
- Conflict of commitment\*: A situation in which an individual accepts or incurs conflicting obligations between or among multiple employers or other entities (i.e., conflicting commitments of time and effort or obligations to improperly share or withhold information from an employer or funding agency).
- Conflict(s) of interest\*: A situation in which an individual, or the individual's spouse or dependent children, has a financial interest or financial relationship that could directly and significantly affect the design, conduct, reporting or funding of research.
- Digital persistent identifier\*: A unique digital identifier that permanently and unambiguously identifies a digital object or an individual.
- Foreign country of concern: As stated in Section 10612 the "CHIPS and Science Act of 2022," "the term 'foreign country of concern' means the People's Republic of China, the Democratic People's Republic of Korea, the Russian Federation, the Islamic Republic of Iran, or any other country determined to be a country of concern by the Department of State."
- Human oversight: The review and validation of information detected by analytic methods to confirm that potential inconsistencies reflect verified inconsistencies, and then whether these verified inconsistencies reflect previously unknown conflicts of interest."
- Investigation: Both the agency and the Office of Inspector General, or OIG, gather and analyze information, draw conclusions and consider courses of action thereon. However, OIG possesses specialized investigative techniques, including searches, seizures, subpoenas, arrests and access to federal prosecution authorities which ensure an allegation is brought to a defensible conclusion. The *NSF Personnel Manual*, Section 143, requires that NSF employees report "allegations of misconduct, fraud, waste, abuse, or corruption involving NSF, NSF employees, NSF-funded research or education, or proposals for NSF funding" to OIG for investigation.  
**Within the context of research security, NSF does not define OCRSSP activities as investigative.** Rather, such activities are related to the identification of potential compliance inconsistencies. If OCRSSP identifies a matter within OIG jurisdiction, the matter will be referred to OIG for investigation. For a detailed description of OIG's scope of investigation, visit [link].
- Potential inconsistency: A possible inconsistency between information provided to NSF and other external sources (e.g., publication affiliation) that has yet to be verified by human oversight, communication with an institution or the Office of Inspector General.
- Research security\*: Safeguarding the research enterprise against the misappropriation of research and development to the detriment of national or economic security, related violations of research integrity, and foreign government interference.
- Routine assessment: The regular monitoring activities required by NSF subject matter experts to uphold NSF's mission and grant distribution responsibilities. Activities include reviewing grant recipient activities, engaging with the research community and reviewing scientific literature to keep abreast of scientific advances.

- **Verification:** The act of comparing disparate data sources (e.g., internal NSF data and open-source information) to confirm that potential inconsistencies reflect accurate information regarding the individual's identity and activities. While verification activities might require engagement with third parties (research institutions, principal investigators and federal agencies), they are not considered investigative activities and do not include any punitive actions.
- **Verified inconsistency:** A potential inconsistency whose conflicting information has been verified by human oversight to accurately reflect a difference between information reported to NSF and information from other sources.

## 6. Research security responsibilities and process of the Office of the Chief of Research Security Strategy and Policy

NSF is committed to safeguarding the integrity and security of science while also keeping fundamental research open and collaborative. This requires coordinated, transparent initiatives managed by NSF's Office of the Chief of Research Security Strategy and Policy. **OCRSSP is the only office within NSF approved to conduct advanced monitoring and verification activities related to research security.** These unique responsibilities of routinely assessing, verifying and reporting inconsistencies are summarized in [Figure 1](#).

NSF's OCRSSP is firmly committed to conducting its research security-related advanced monitoring and verification activities in strict accordance with federal law. As stated in the 2022 "CHIPS and Science Act" Section 10637, "In carrying out requirements under this subtitle, each Federal research agency shall ensure that policies and activities developed and implemented pursuant to this subtitle are carried out in a manner that does not target, stigmatize, or discriminate against individuals on the basis of race, ethnicity, or national origin, consistent with title VI of the Civil Rights Act of 1964 (42 U.S.C. 2000d et seq.)." OCRSSP will ensure these requirements are upheld while carrying out its research security responsibilities.

### 6.1 OCRSSP research security responsibilities

**Policy Development:** The CRSSP is responsible for leading NSF's research security strategy and policy development to ensure that research security priorities and federal guidance are implemented throughout the agency. For example, NSPM-33 4.b.iii requires policies for disclosure of specific information and is implemented through NSF's [pre-award and post-award disclosure requirements](#). The CRSSP also oversees the development of training resources, meeting NSPM-33 4.f.'s guidance that "Federal agency personnel conducting R&D activities or participating in the process of allocating Federal R&D funding receive research security training."

**Advanced Monitoring:** NSPM-33 4.b.viii requires that NSF works with partners in the enterprise "to identify and investigate potential violations of agency disclosure requirements." Additionally, Section 10331.6 of the "CHIPS and Science Act of 2022" requires OCRSSP to "[perform] risk assessments, in consultation, as appropriate, with other Federal agencies, of Foundation proposals and awards using analytical tools to assess nondisclosures of required information." Leveraging advanced monitoring techniques (see [Definitions](#)) is critical to such efforts; OCRSSP is expected to employ transparent,

unbiased and accurate techniques to ensure that NSF can understand and respond to research security developments. Dedicated data scientists work with the CRSSP to ensure that all analyses meet these goals, employing the methods (see Sections [9.1](#) and [9.3](#)) and verification criteria (see [Section 9.2](#)) described in the guidelines.

**Verification:** To safely monitor the scientific enterprise and respond to potential inconsistencies (see [Definitions](#)) in information disclosed to NSF, the CRSSP must ensure that open-source information is accurate and represents the activities of the attributed individual. OCRSSP staff employ computational match requirements (see [Section 9.2](#)) and directly engage with institutions (see Sections [10.2](#) and [10.3](#)) to validate information and determine whether a potential inconsistency can be verified (see [Definitions](#)). **OCRSSP verification activities are not considered "investigative" activities, as defined by the Office of Inspector General** (see [Definitions](#)).

**Reporting:** Should a potential inconsistency or other activity be validated, thus becoming a confirmed inconsistency (see [Definitions](#)), the CRSSP may be required to report the incident to the Office of Inspector General. Furthermore, NSPM-33, 4.(e) requires that agencies "share information about violators ... across Federal funding institutions and with Federal law enforcement agencies, the DHS, and State, to the extent that such sharing is consistent with privacy laws and other legal restrictions and does not interfere with law enforcement or intelligence activities." SORN NSF-77 further aligns with NSPM-33 4.e by establishing routine uses that constitute "provisions that allow for such information sharing."

**Outreach and Education:** OCRSSP works to foster understanding about research security concerns across academia, industry and government. It is responsible for engaging with research security stakeholders, ensuring transparent communication of NSF's activities and concerns, and providing the informational resources to act upon them. **In addition, the CRSSP is responsible for coordinating activities required by NSPM-33 and the "CHIPS and Science Act of 2022":**

- "CHIPS and Science Act of 2022" Sections 10331 and 10332 require OCRSSP to serve "as a resource at the Foundation for all issues related to the security and integrity of the conduct of Foundation-supported research," which includes conducting outreach and education activities.
- NSPM-33a 3.(ii) Requires NSF to "cooperate with organizations receiving Federal funds to ensure that the organizations have established and administer policies and processes to identify and manage risks to research security and integrity, including potential conflicts of interest and commitment."
- NSPM-33 4.(g) Requires that "research institutions receiving Federal science and engineering support in excess of 50 million dollars per year certify to the funding agency that the institution has established and operates a research security program."



# NSF's Steps for Handling Potential Inconsistencies



OBSERVING, REPORTING, VALIDATING & TAKING ACTION

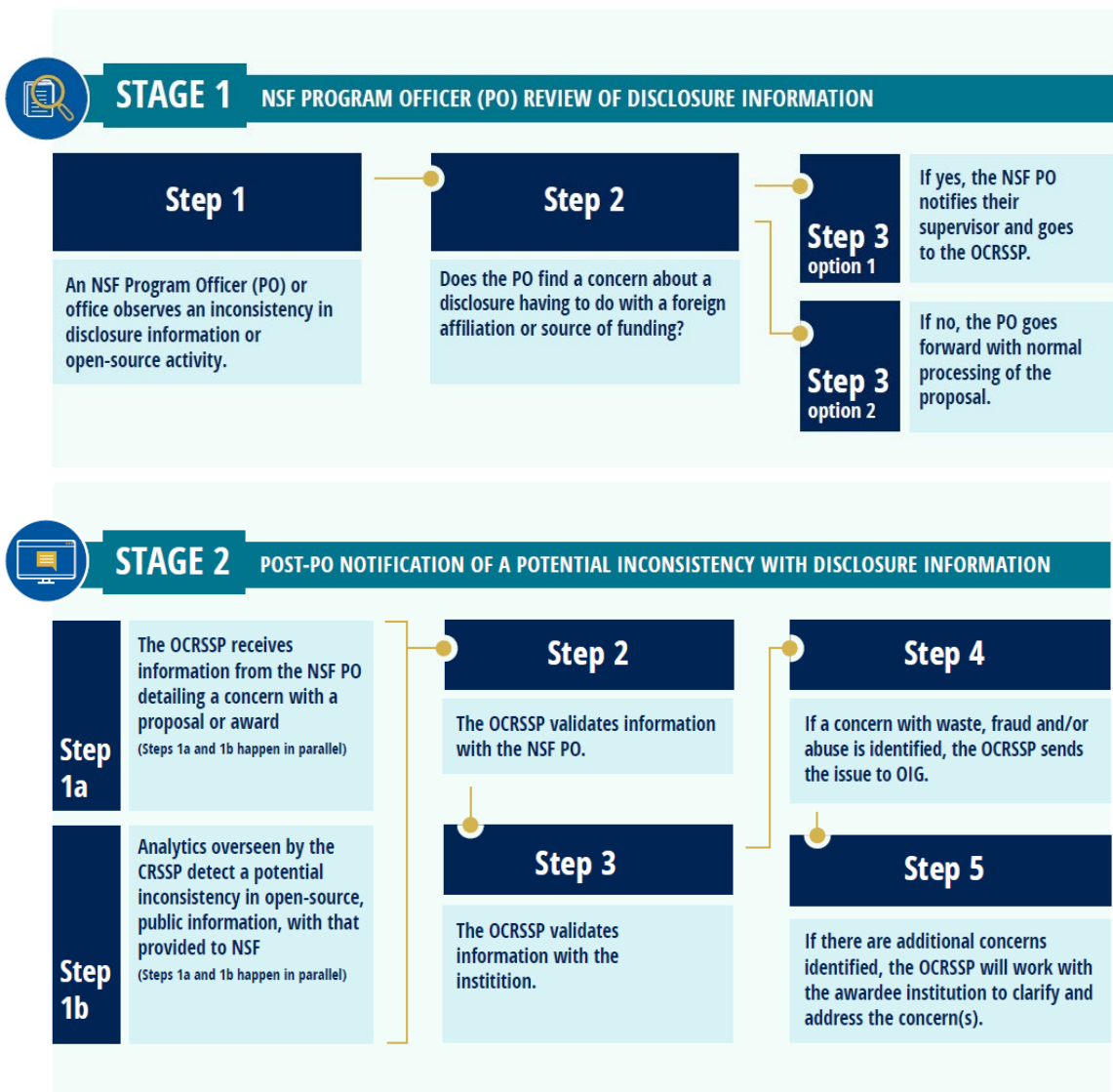


Figure 2. NSF's Steps for Handling Potential Inconsistencies

## 6.2 Process for notification and communication with institutions

Transparency and consistency are necessary pillars for NSF to build trust with stakeholders and uphold its principles of fair, unbiased research security practices. Contractual obligations are between NSF and institutions, and the latter is the legally responsible entity for NSF to engage with. When OCRSSP engages with institutions on potential inconsistencies, it will employ the process below, devised to comply with key guardrails in NSF's practices:

1. All research security communications to proposing and/or awarded institutions are conducted solely by OCRSSP.
2. OCRSSP activities are considered to be verification related. OCRSSP does not conduct investigative activities (see [Definitions](#) for the difference between verification and investigation).
3. Program staff are not allowed to discuss research security concerns with a principal investigator and are not permitted to incorporate research security concerns into the merit review process.

Institutions are the entity OCRSSP will primarily engage with when a potential inconsistency relates to a lack of compliance with NSF's proposal submission requirements or award [terms and conditions](#). As such, OCRSSP may communicate any concern over non-compliance with an institution's Sponsored Research Office, or SRO.

### **OCRSSP Process for verification of potential Inconsistencies:**

1. OCRSSP may contact an institution's SRO to inquire about a potential inconsistency.
  - The notice will clearly state that this is a verification activity, and that no preemptive actions should be taken other than providing NSF with the relevant information.
  - NSF will only engage with the SRO to verify a potential inconsistency. Neither the OCRSSP nor the program officer will engage with the PI on the matter.
2. OCRSSP will request a response from the institution SRO by 30 days from notification.
  - SROs are encouraged to report a non-resolved dispute between the SRO and the PI on the validity of the inconsistency. Such incidents may be discussed with the Office of General Council and other offices at NSF as appropriate before action is taken.
3. Depending on the inconsistency reported, OCRSSP will take actions that include, but are not limited to, the items listed below. A verified inconsistency may also be shared with directorate leadership.
  - Sharing the verified inconsistency with the OGC.
  - Should there be a concern of non-compliance with NSF's award terms and conditions, coordinating with OGC and the Division of Grants and Agreements, with consultation with the Policy Office, to determine appropriate action.
  - Sharing the verified inconsistency with the Office of the Inspector General should there be a concern of waste, fraud, or abuse.
  - Sharing the verified inconsistency with other federal agencies engaged in research security activities, as appropriate (see routine uses 13 and 14 from [SORN NSF-77](#)).

## 7. Monitoring and reporting by NSF offices and staff

### 7.1 Terms and conditions compliance-monitoring responsibilities of program staff

As part of NSF's grant review and management process, program staff are responsible for checking proposal disclosure information and overseeing their award portfolio to ensure awardees complete their obligations (see routine assessment in [Definitions](#)). They are also expected to engage with the research community and keep abreast of scientific developments. Through these activities, program staff may encounter information that suggests an awardee is non-compliant with NSF's [PAPPG](#) disclosure requirements and NSF's [terms and conditions](#), which span multiple topics, including budget expenditures, project reports, etc.

In January 2022, the White House Office of Science and Technology Policy released the [NSPM-33 implementation guidelines](#). While the implementation guidelines provide clarity to the federal government on what disclosure information is required, the scope of NSF's program staff responsibilities remains unchanged.

**During routine assessment and compliance monitoring responsibilities, program staff are expected to adhere to the following guidelines:**

- Program staff should not conduct data searches and analyses focused on research security concerns beyond an initial check of potential inconsistencies observed during the grant review process and routine assessment.
- Program staff should not conduct any research security-related activities that actively search for inconsistencies with NSF's terms and conditions, such as NSF's disclosure requirements. The program officer may propose to their supervisor to submit a request to OCRSSP.
- Program staff should not engage directly with PIs on any matter related to research security without guidance from OCRSSP.
- Potential research security concerns observed during proposal review and routine assessment that reflect non-compliance with NSF's terms and conditions are to be communicated to OCRSSP.
  - *Example of routine assessment incident: During a review of an annual project report, a program officer sees that the PI has published a new paper as part of the research funded by the NSF award. While browsing the paper, they notice that the PI acknowledges funding from a non-U.S. funding agency they don't recall seeing in the original proposal application. Upon checking NSF's records, it is apparent that the other source of funding had not been disclosed in the proposal application. The program officer consults with their supervisor and sends OCRSSP the relevant information.*
  - *Example of proposal review incident: Prior to an award, a program officer asks the principal investigator to update their Current and Pending Support information. After an update is made by the principal investigator, the program officer is suspicious that the update is left intentionally incomplete. After consulting with their supervisor, the program officer notifies OIG and OCRSSP and continues to process the awarding of the proposal.*

- Directorate leadership may request research security-related analyses from OCRSSP. Human oversight and verification will be conducted by OCRSSP in coordination with said office.
- Potential research security concerns observed during other activities such as routine assessment and advanced monitoring (see [Definitions](#)) are to be communicated to OCRSSP.
  - Human oversight and validation should be conducted before information is shared with OCRSSP (see Section 10.3 for validation requirements).
  - As appropriate, the program officer overseeing the award should be informed in advance of concerns.
- NSF recognizes that program staff routinely conduct literature searches and utilize data-driven tools as part of their responsibilities, such as keeping abreast of the scientific literature and finding reviewers with certain expertise. Potential research security concerns identified unintentionally while conducting such activities are to be communicated to OCRSSP.
  - *Example: While browsing a journal relevant to their subject matter expertise, a program officer notices a publication by a PI that has an active award from NSF. The PI has listed an affiliation with an international university that the program officer does not recall knowing about. Upon checking NSF's records, it is apparent that the affiliation had been undisclosed in the proposal application. The program officer consults with their supervisor and sends OCRSSP the relevant information.*

## 7.2 Vetting for employment

NSF may use analytics for vetting employment. The mechanism for that vetting and the responsibility of different NSF organizational elements is a current topic of discussion in NSF's "Intergovernmental Personnel Act" Vetting Working Group. Once the working group's efforts are complete, this section will be updated to reflect the agreed-on approach for vetting for employment.

# 8. Permissible and prohibited practices for research security-related analytics by the CRSSP

## 8.1 Permissible approaches for research security analytics

As part of the responsibility to monitor and report research security-related concerns, the CRSSP is approved to apply the data-driven approaches listed below. **OCRSSP staff are prohibited from conducting any analysis that selects for a particular national origin or racial identity** (see Sections [8.2](#) and [8.3](#) for complete list).

1. Conducting analyses that compare **all** participants in agency programs.
2. Comparing of self-reported information of PIs with open-source information by the same PI (see [Section 10.3](#) for matching requirements).
3. Filtering and analyzing NSF's portfolio based on institutional characteristics (e.g., the institution's Carnegie Classification, amount awarded in funding, etc.) and research topics (e.g., quantum computing, artificial intelligence, biological sciences, etc.).

The examples below include permissible practices for research analytics by OCRSSP staff:

- *Example 1: OCRSSP staff conduct a search of all awards given in 2020 to see who has published academic papers in collaboration with institutes from a foreign country of concern.*
- *Example 2: OCRSSP staff have searched all awards given towards a particular topic (e.g., "AI"), filtered by R1 institutions and conducted a portfolio analysis detailing which institutes have research security policies.*

## 8.2 Prohibited practices for research security analytics

As stated in the [NSPM-33 implementation guidance](#), "It is essential that policies and consequences must be applied without discrimination in any way, including with respect to national origin or identity." For research security-related analytics, data queries and analyses that are explicitly or implicitly designed to return the identities of individuals of a specific national origin or racial identity are **prohibited**. This includes querying NSF data and public information by:

- Nationality.
- Citizenship.
- Common names for specific countries/ethnicities.

The examples below show **prohibited** practices for research analytics, as they select for community members based on their national origin or identity.

- *Example 1: An OCRSSP staff member has decided to filter a search based on PIs who obtained their doctorate from an institution in France to check for further inconsistencies in reporting. Because this selects for a group of individuals based on their previous academic training, this search is not allowed.*
- *Example 2: An NSF office has asked the CRSSP to search through their grants to find PIs collaborating with scientists with common Israeli names (Tal, Renan, Dafna, etc.). The intent was to then compare this list of published collaborations to check for inconsistencies in reporting in NSF proposals. Because this activity would select for a particular demographic due to their name, this is not allowed.*

## 8.3 Individual matching criteria for validation and information sharing activities

Human oversight is a critical safeguard of the CRSSP's monitoring, validating and reporting responsibilities described in Section 9.

**Any validation and reporting activities of individuals with conflicting information in the public domain will first be checked against matching criteria listed below.** NSF staff reporting inconsistencies to the CRSSP are instructed to apply similar oversight. A potential inconsistency will be considered "verified" if the institution, or OIG, confirms the inconsistency detected by NSF.

- The first names without the middle names **and** the email addresses match in both datasets.
- The last names or a portion of the last names (in case of hyphenated names) **and** the email addresses match in both datasets.

- The first names (without the middle names) and the last names (or a portion of the last names, in the case of hyphenated names) match with either a) department affiliation, or b) the institutional affiliation in both datasets
- Email addresses **and** department affiliation match in both datasets.
- Persistent ID (e.g., ORCID) match in both datasets.

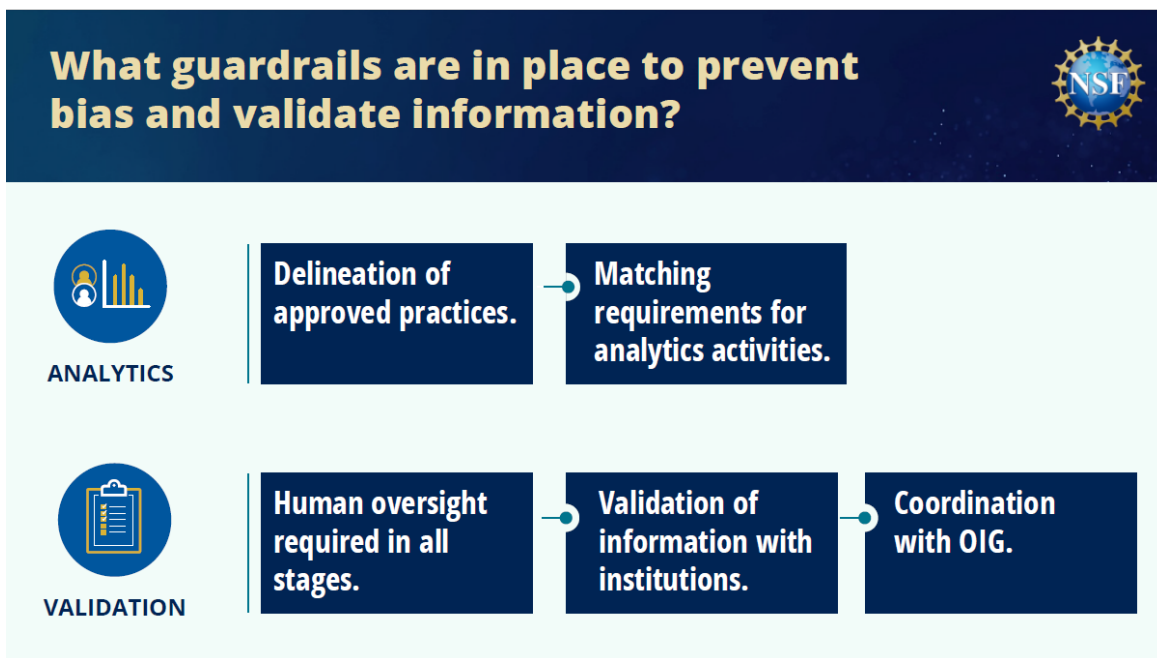


Figure 3. Guardrails in Place to Prevent Bias and Validate Information

## 9. Data, services and methods used for research security analytics

### 9.1 Non-NSF data used in research security analyses

From 2021 to 2022, OCRSSP staff have used the following data sources to conduct research security-related analyses. These are widely used industry standards and are maintained and updated by dedicated staff. This list will be updated when and if new data sources are used.

1. Elsevier SCOPUS.  
<https://www.elsevier.com/>  
Transfer method: private encrypted AWS download.  
Update rate: monthly, with option for weekly updates.
2. Web of Science.  
<https://clarivate.com/webofsciencegroup/solutions/web-of-science/>  
Transfer method: private secure SFTP download.  
Update rate: weekly.
3. U.S. Patent and Trademark Office Patent Database.  
<https://bulkdata.uspto.gov/>

Transfer method: public bulk weekly downloads.  
Update rate: weekly.

## 9.2 Analysis criteria and purpose

NSF uses several criteria to detect reporting inconsistencies between internal information and data that is published and/or otherwise in the public domain. These criteria include, but are not limited to:

1. Mismatches between institutional affiliations in published papers and disclosed/reported institution affiliations within proposals to NSF.
2. Mismatches between funding sources in published papers and disclosed/reported Current and Pending funding sources within proposals to NSF.
3. Mismatches between filed patents in the USPTO and self-reported intellectual property filings in NSF annual reports

## 9.3 Services used in research security analyses

**NSF currently does not employ any off-the-shelf analysis tools or services for research security-related activities.** Current analysis activities of the above databases are conducted using Apache SOLR, Carrot Search Lingo4G, Python and R. Information is shared through Excel spreadsheets and visualized through dashboard interfaces created with the [Lucidworks fork](#) of the [Kibana framework](#).

# 10. Sharing guidelines for security-related information

## 10.1 Human oversight

Safeguarding NSF's analytics process is a top priority for the agency, and human oversight is a paramount requirement for all information sharing activities. No information on individuals may be reported and no adverse action may be taken based solely on a potential inconsistency without human verification of the matching criteria.

## 10.2 Sharing of information with institutions

To clarify inconsistencies in reported information, NSF may share information with the organization or institution that originally submitted the proposal to help cross-reference and verify information (see [SORN NSF-77](#) Routine Use 12).

**Verification requirements:** When sharing an individual's reporting inconsistencies with an institution (see [SORN NSF-77](#) Routine Use 12), OCRSSP must attempt to verify the information according to the matching criteria listed in [Section 8.3](#). Should any item not be successfully verified, this will be clearly communicated to the organization or institution.

## 10.3 Sharing of information by OCRSSP with inspector general or federal agencies

In accordance with the "CHIPS and Science Act" Section 10331 and NSPM-33 4.(e), information on validated incidents of violations of NSF's disclosure requirements for submitted proposals, the terms and conditions of an award, and project reports may be disclosed by OCRSSP to the appropriate federal agencies (including but not limited to OIG, law enforcement, intelligence agencies, and other relevant

agency components) to inform efforts related to national and research security (See [SORN NSF-77 Routine Use 13](#)).

- All reporting must have documentation showing that the verification process included human oversight.
- Personally identifiable information, or PII, shared with agencies will be restricted to information about senior personnel only and coordinated with OIG, unless otherwise required by applicable law.
- To the extent allowed by law, NSF will not share any PII information beyond an individual's name and affiliated institution. Self-defined ethnicity, gender, etc. will not be included. Relevant administrative information, such as PII-redacted proposal and award history and open-source information, may be shared. Only proposal content of verified inconsistencies may be shared with appropriate points of contact at other agencies.
  - *Example of information sharing: An intelligence agency is investigating a PI and has asked for information regarding NSF-funded activities, including proposal information. CRSSP staff prepare the relevant information but will omit any PII-content from the proposal, including gender and ethnicity.*

The CRSSP is responsible for approving all information shared by OCRSSP with OIG or federal agencies. Prior to consideration of further action, OCRSSP may consult with host institutions to verify potential inconsistencies, and when appropriate, share the verified inconsistencies with the appropriate federal agencies.