RISKY SPACE BUSINESS: NASA AI RISK PREDICTION CHALLENGE

Predict project risks using NASA's lessons from the past. Design a project management tool that extracts past project risk information and uses AI/ML to predict risks on future projects.

If you're an AI/ML specialist or someone who is great at extracting data from past project reports, this challenge is for you! The Risky Space Business Challenge is looking for White Papers and algorithms to help NASA predict risks on future projects. If you're up to the challenge, you could win a share of the $50,000 purse.

To participate in this challenge, you must describe or show how you would: 1. Extract risk data from past project reports; 2. Create a template for future project reporting; 3. Create a taxonomy of risks for future use; and 4. Use AI/ML to identify potential risks in future projects using past project reports. Winning solutions will be used to help NASA, and potentially anyone managing a project, improve their ability to identify risks before they become real issues.

CHALLENGE OVERVIEW

All project plans are created with the hope that everything will go to plan. However, that rarely happens; schedules slip, costs increase, and components can't be built as planned. If only you could use all the knowledge from past projects to help you assess risk on future projects. That is what NASA's Game Changing Development (GCD) Program hopes to do through the Risky Space Business Challenge. Although NASA moves technology forward by standing on the shoulders of giants, the hard lessons that were learned often get lost. Along with collections of project documentation, NASA has developed a Lessons Learned System to capture those lessons, but often it is hard to correlate what is documented from the past with what is being worked on now. What if there was a tool that could help filter the knowledge of the past and help identify the relevant items to help avoid issues on a current or upcoming project?

GCD is looking for competitors who can solve one or more or all of the following:

- an algorithm to extract risk data from past project reports
- a template for reporting project data that provides data structures suitable for AI/ML
- new ways of categorizing risk and creating data structures to enable AI/ML
- a plan to put it all together to use AI/ML to identify project risks

For this challenge, whether you choose to address all or part of it, you will need to write a White Paper and provide algorithms, if available, describing or demonstrating your solution. If you can solve all or parts of the challenge, you could win your share of the $50,000 purse. While this solution is designed to help NASA identify risk on their projects, it could be applicable to many other businesses and situations where planning for risk is important. The results of this challenge may be implemented by NASA or followed by an implementation challenge to finalize the digital assistant.

GUIDELINES

Challenge Background

The Game Changing Development (GCD) Program manages well over a hundred technology development projects for NASA. These projects take technologies from the lab and show they can work in the expected environments of space, the Moon or Mars. They are run at NASA centers, by commercial industry, or a combination of both. The technologies may lead to entirely new approaches for the Agency's future space missions and provide solutions to significant national needs. The most promising ideas progress through analytical modeling, ground-based testing, and spaceflight demonstration of payloads and experiments. Teams are held accountable for ensuring that discoveries move rapidly from the laboratory to application.

Being able to identify and manage risk is essential to effectively run any development project. The goal of this Challenge is to move GCD from recognizing and reporting risk during a project and after the project is completed, to being able to identify potential risks by comparing project documentation on active projects to historical data and information available from other GCD projects. This data might include project plans, status reports, test reports, formal review charts, and closeout reports. Currently, risks are only included in these reports and in the Lessons Learned System. To make that happen, we believe the following steps must be addressed:

1. Define the data that needs to be collected and a structure to be used to output risk data for AI/ML use.
2. Extract past project risk data into new output.
3. Develop structures for categorizing and identifying risks.
4. Design an AI/ML system that uses all past data to identify known and unknown potential risks in new projects.

PROJECT REPORTING DATA

NASA has provided data in many forms to help participants approach the challenge. After registering for the challenge, eligible participants will be given access to the NASA files. These files include:

- Project reporting files from two of their past projects:
    - the Astrobee project, a free-flying robot that is part of Human Exploration Telerobotics 2 (HET2) project.

    - the Space Synthetic Biology (SynBio) project which develops technologies to biomanufacture valuable products on-demand such as vitamins and medicines to be deployed on long-duration missions.
- Data from the NASA Public Lessons Learned System reformatted into CSV format (lessons_learned.csv).
- NASA Project Document Summary describing documents reporting on risk, what documents are produced, and the order in which they are produced. Includes descriptions of all project documents provided (NASA_Project_Documents.pdf).

The available data will give you a sense of the kinds of reports and risks that GCD project managers use.

As a proof of concept, it may be beneficial to use other datasets that may be richer than NASA's when building your solution. If you are using external datasets, at a minimum you will need to provide a general description of the data and why it would be a good analog dataset to use for this Challenge.

RISK CATEGORIZATION

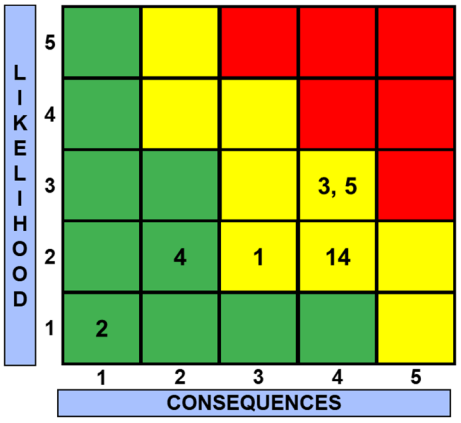
Currently, GCD project managers categorize risk in multiple ways, which can appear in documents in different ways. Risk can be reported in a 5x5 matrix based on severity and likelihood. A brief mitigation plan is provided for the top project risks. For technology development projects, identifying risks is challenging; often there are few analogs available in these areas. This leads to many "unknown-unknown risks" that cannot be anticipated. Risk can also be categorized into schedule, cost, technical, and programmatic risks; severity can be color coded, and trends (increased or decreased risk) can be indicated by directional arrows.

In the following diagrams from the Astrobee Project FY2016 Q3 Review PowerPoint presentation (included in the shared documents), there are several ways that risk is presented. In Figure 1, colors, based on severity of risk, are used in a table with rows for each time period (Q1, Q2, Q3) and columns for each performance area. The rationale for performance in the red or yellow state are described in a separate column.

Period	Summary Performance				Rationale
	Cost	Schedule	Technical	Programmatic	
Q1					**Technical** - Unresolved power fault with R2 on ISS **Schedule** - Lack of US crew time for R2 on-orbit activities **Programmatic**- Astrobee technology transition to customer will be incomplete if current project scope is not increased.
Q2					**Schedule** - Astrobee schedule requires overlapping cert and flight unit builds to achieve "on dock" at end of FY17. **Programmatic**- Astrobee technology transition to customer will be incomplete if current project scope is not increased.
Q3					**Schedule** - Astrobee schedule requires overlapping cert and flight unit builds to achieve "on dock" at end of FY17. **Programmatic**- Astrobee technology transition to customer will be incomplete if current project scope is not increased.

Figure 1 - Project Summary Performance

In Figure 2, a standard 5x5 risk matrix is used along with trend direction arrows, the approach to dealing with the risk, and the affinity, or risk area. Note that risk areas in Figures 1 and 2 are not labelled with the same terminology.



Risk ID Trend	Approach *Affinity*	Astrobee Risk
1 ↓	**M** *T*	Pose accuracy with vision-based navigation (FFREQ-42)
2 ↓	**M** *T*	Avionics and fan noise (FFREQ-47)
3 ↑	**R** *T*	Safety approval of autonomous operations (FFREQ-49)
4 ↓	**M** *T*	Performance trades: fans, batteries, size (FFREQ-44)
5 ↓	**W** *Sc*	Schedule is best case (FFREQ-46)
14 ↓	**M** ***T***	Technology transition is incomplete (FFREQ-1194)

Criticality
High
Med
Low

L x C Trend
↓ Decreasing (Improving)
↑ Increasing (Worsening)
→ Unchanged
□ New Since Last Period

Affinity
T - Technical
C - Cost
Sa - Safety
Sc - Schedule

Approach
M - Mitigate
W - Watch
A - Accept
R - Research

Figure 2 - Risk Summary

Description of the Challenge

Currently, GCD is dependent on its project managers to identify the risks on their projects and report them at periodic status reviews. GCD project managers have a wide range of backgrounds and experience levels so the types of risks that are reported vary greatly. Often, key risks that seem obvious to managers with one background are overlooked or not considered critical to those with different backgrounds. And, since project managers aren't always familiar with other projects, risks on one project may not be identified on a very similar project. Creating a tool to let project managers be aware of all potential risks would help manage risks across all GCD projects and increase the project success rates. This tool, or "digital assistant" would use AI/ML to identify risks. For example, the digital assistant might be given a project plan with the following information:

> *A new GCD project plan provides a 2-week duration for thermal-vacuum testing. The average duration for thermal-vacuum testing on 5 other projects was 4 weeks.*

In this case, the schedule for testing should be identified as a potential risk, point to the historical data for the rationale, and provide a confidence level. There may be other collateral risks that can be identified as well, for example, testing costs may be under budgeted based on the longer testing time. While solutions to this challenge may not get GCD all the way to their end goal of a digital assistant, solutions for each section of the challenge will be steps along the way.

For this challenge, you will present a White Paper and algorithms, if available, to help NASA extract known risks from past projects that can then be applied to identifying risks on future projects. The White Paper must detail how you would complete the following:

1. Propose data to be collected in reports including data fields and structure to support AI/ML to predict and categorize risk. Include recommended data fields/data structures based on past data and recommend new fields to improve risk analysis.
2. Populate either NASA project data or an external dataset you provide into the new format. Describe and demonstrate the data extraction process, approach, and methodology.
3. Propose a method for predicting risks on future projects using AI/ML and your new structured data format.
    a. Demonstrate and/or describe how to build a library of risks and classify risks based on past projects. If possible, risks should be classified using NASA's risk categories (schedule, cost, technical, and programmatic) and/or the 5x5 matrix. However, solutions with new, or more impactful methods of categorizing risk are preferred.
    b. Demonstrate and/or describe the process, approach, methodology using the NASA data or your own dataset.

After registering for the challenge, you will have access to the sample documents from two past NASA projects. The documents are in multiple formats including csv, ppt, pdf, and doc. These are representative of the types of reports created during a NASA project but are not the only documents or document types. You can also perform a proof of concept using your own dataset and explain how it would apply to NASA's requirements.

Schedule

LAUNCH	**October 22, 2021**
Webinar with Challenge Team SMEs	**Week of November 9th, 2021**
Submission Deadline	**February 7, 2022, 5:00 pm ET**
Judging	**February 8 – February 28, 2022**
Awards/Announcement	**March 1, 2022**

Prizes

Total Purse: $50,000

Prizes for this challenge will be monetary only.

Up to 3 prizes will be awarded to the top solvers that achieve the requirements of the challenge for a total of $35,000. The best solution will be awarded $20,000, second place will be awarded $10,000, and third place will be awarded $5,000. Another $15,000 in prizes may be awarded for innovative solutions in data

extraction, data formatting, or risk prediction with a minimum prize amount of $2,500. These additional prizes can be given to solutions that place in the top three, or to other submissions that are exceptional in these areas. If fewer solutions meet the criteria for a prize, the money will be redistributed among the qualified winners.

For this Challenge, ownership of all intellectual property rights, if any, in the idea or concept demonstrated by the proposed solution will remain with the solver. However, as a condition of receiving a cash prize, Each Winning Team (whether an individual or group of individuals) selected to win a prize under this Challenge must apply a permissive open source license to their solution. Also, NASA reserves the right to negotiate licenses for additional submitted solutions.

CHALLENGE

To be eligible to win this challenge, competitors must submit a White Paper detailing the solution to the challenge. Proposals that also include algorithms demonstrating their solutions have the ability to score higher during judging (see: Judging Criteria below). Your algorithm can work with the data provided by NASA or you can use your own dataset that demonstrates the requirements of the challenge.

White Papers must be submitted using the Challenge Submission Form. All code and code documentation must be uploaded to your personal GitHub repository, and you must share access with the Challenge team.

Submission Requirements:

Note: A team or individual may submit more than one solution to the challenge. However, each submission must be a substantially different solution. For example, if the solutions use a different algorithm for extracting data or assessing risk, they can be submitted separately. However, if two submissions from the same team are deemed to be too similar, only the last submission will be considered for judging.

All submissions must be written in English and cannot be handwritten.

All solutions **must:**

- Be from an eligible team.
- Include uploaded Registration Form, Submission Form. Submissions may also include algorithms demonstrating any or all portions of the solutions.
- Include solutions to all parts of the challenge.

WHITE PAPER

The White Paper information must be submitted using the **Challenge Submission Form**.

1. Team Lead Name
2. Team Lead Freelancer.com username
3. Submission Title

4. Executive Summary (500 words)
   Describe your solution and provide context for your solution for the judges.
5. Your solution:
   a. **Proposed data to be collected** (2-3 pages)
      i. Recommended data fields and data structures to be used based on past data.
      ii. Recommended new data fields to improve risk analysis.
      iii. Describe the proposed structure.
         1. Describe how the structure works with existing risk data (NASA's or another dataset) and how it will help predict risks and identify unknown risks on projects.
         2. Provide a visual of your data fields and structure. For example, you can create a wireframe, picture, data schema, or diagram. Upload a file or provide a link.
      iv. Describe how the structure supports program management needs, isn't overly constraining, and is easy to adopt.
      v. Describe how the output can be used by AI/ML to categorize and predict risk.
   b. **Proposed method to populate the existing data into the new format** (2 pages)
      i. Describe the data extraction process, approach to working with GCD documents or your own dataset, the methodology applied to the process
      ii. If you have an algorithm, demonstrate the data extraction process by showing the output from the algorithm. Provide the source code file name, instructions for use and compilation including any environmental requirements and dependencies, and sample outputs, if available. Upload the code to GitHub (instructions below).
   c. **Proposed method for predicting risks for each project given the new structured format** (3 pages)
      i. Describe the process, approach, and methodology using GCD data or your own dataset.
      ii. Describe the method of building a library of risks and classifying risks from past projects
      iii. Describe how you could use this to identify unknown risks.
      iv. If you created an algorithm, demonstrate the process. Provide the source file name, instructions for use and compilation including any environment requirements and dependencies, and sample outputs, if available. Upload the code to GitHub (instructions below)
6. If external data sources are used in your analysis, provide information about the source. For example, the type of data, how it addresses risk, and the output risk analysis. Provide as much information as you can about the data source.
7. If any algorithms are included, provide the instructions for accessing the code from your GitHub repository:

a. GitHub Username
b. GitHub Repository Name
c. Link to your GitHub Repository

CODE

Algorithms may be tested by the judges. All source code, compilation instructions, and operation instructions must be clear and accurate. Submissions must be uploaded to your private GitHub repository and access must be granted to the Challenge team. For each algorithm submitted you must include:

1. Source code for algorithms
2. Output from the algorithms

If an external open-source library is used, the libraries must have a permissive open-source license, you must be able to provide a full list of libraries used (e.g., requirements.txt), and the libraries must be accessible to the Challenge team and NASA.

GITHUB ACCESS

To submit your compiled solver and source code through GitHub, you will need to provide the Challenge Team access to your private GitHub repository:

- Upload your solution and code to a private GitHub repository.
- Add "enterprise-admin@freelancer.com" as a collaborator.
- Make sure "enterprise-admin@freelancer.com" has permissions to fork your repository.

Provide the following information in your Submission form:

- Your GitHub Username
- Your GitHub Repository Name
- Link to your GitHub Repository

Registration

Registration will be open from the launch of the challenge until the submission deadline of February 7, 2022, 5:00 PM ET.

To Register for the Challenge, Teams must:

1. Sign up for an account on the Freelancer.com platform
2. Complete and submit the Registration Form

Registration Forms will be reviewed for eligibility; ineligible teams will be notified and teams that meet the Challenge eligibility requirements, as listed in the Rules, may continue in the competition. Eligible teams will be provided with the sample project data from NASA to use for the solution.

It is important to note that the earlier a team Registers for the Challenge, the sooner they will have access to NASA's data.

Final Judging

Submissions will be prescreened for eligibility, completeness, and attention to the requirements of the challenge. A judging panel consisting of SMEs and representatives from the NASA GCD Program will judge the submissions based on the judging rubric below. Algorithms may be tested for accuracy against a different set of project documentation.

JUDGING CRITERIA

Criteria		**Weight**
DATA COLLECTION ● Data fields and structures effectively capture past risk data ● Designed for future AI/ML ● Designed to predict known and unknown risks on current and future projects ● Doesn't limit the information that is reported ● Allows for future modifications	White Paper Description	20
	Wireframe, mockup or data structures	5
BRINGING PAST RISK DATA INTO THE NEW TEMPLATE ● Ability to bring past data (NASA or other dataset) into the new structure ● Effective filtering of risk data (only risks, all risks, high-level risks) ● Output of risk data from the new template	White Paper Description	15
	Code and code outputs	10
BUILDING A RISK LIST FROM HISTORICAL DATA ● Ability to pull all risks from past data ● Ability to categorize risk ● Ability to identify previously unknown risks	White Paper Description	10
	Data structure and code	5
EXPLANATION OR DEMONSTRATION OF FUTURE RISK ANALYSIS Bringing all the pieces together to build risk analysis for projects in the future	White Paper Description	25
	Code and code outputs	10

RULES

NASA and employees of Challenge partners reserve the right to update or modify this challenge at any time. Updated material will be made available to all participants.

Eligibility Requirements

By participating in a Freelancer.com contest, each participant (whether an individual, group of individuals, or entity) must agree to and abide by the following: Freelancer Eligibility Policies, Freelancer User Agreement, and the Freelancer Copyright Infringement Policy. Each participant must complete and comply with the Freelancer KYC process. All team members must meet Freelancer eligibility policies.

All team members must be at least 18 years old. The team lead must be eligible to receive payment under the laws of the United States; U.S. federal sanctions prohibit participation from certain countries. (see: https://www.treasury.gov/resource-center/sanctions/Programs/Pages/Programs.aspx).

NASA and employees of Challenge partners are not eligible to participate. Federal employees acting within the scope of their employment are not eligible to participate. Any individual or entity associated with the development or administration of this challenge is ineligible to compete. Government contractors working on the same or similar projects are ineligible to participate in the Challenge.

Funds from U.S. or foreign government organizations should not be used to directly fund the development of a solution to this Challenge. Solutions that were previously developed with Government/Federal funds, or where Government/Federal funds, including but not limited to, employee time, materials, and reviews, were utilized to prepare the submission or solutions are prohibited.

Submission Requirements

Solutions must originate from either the U.S. or a designated country (see definition of designated country at https://www.acquisition.gov/far/part-25#FAR_25_003), OR have been substantially transformed in the US or designated country prior to prototype delivery pursuant to FAR 25.403(c).

Submission Forms must be uploaded to the Challenge page and data solutions must be available on GitHub before the deadline, February 7, 2022, at 5:00 pm ET. No late submissions will be accepted.

All submissions must be submitted using the PDF Submission Form. Submissions must be in English, and machine-readable (not handwritten).

Payment

For any prize award, the winner will receive the full amount awarded; any Freelancer seller fee will be refunded to the winner.

If participating as a team, all payments will be made to the Team Leader who is solely responsible for the distribution of funds among team members.

Intellectual Property

Each Winning Team (whether an individual or group of individuals) selected to win a prize under this Challenge must apply a permissive open source license, such as Apache 2.0, to their solution(s) and retain the software and accompanying documentation in a publicly available repository until June 3, 2025.

Ownership of all intellectual property rights, if any, in the idea or concept demonstrated by the proposed solution will remain with the solver.

NASA reserves the right to negotiate licenses for additional submitted solutions.

If you are unable to comply with the IP agreement for all or part of your solution, contact the challenge team at nasa-gcd@freelancer.com

Please submit questions to the challenge Public Clarification Board or to nasa-gcd@freelancer.com.