



Data Article

Data on the Human Versus artificial intelligence process management experiment



Nicolas F. Soria Zurita^{a,b}, Joshua T. Gyory^c, Corey Balon^d,
Jay Martin^d, Kenneth Kotovsky^e, Jonathan Cagan^{c,*},
Christopher McComb^{c,*}

^a School of Engineering Design, Technology, and Professional Programs, The Pennsylvania State University, University Park, PA, United States

^b Colegio de Ciencias e Ingeniería, Universidad San Francisco de Quito, Diego de Robles y Vía Interoceánica, Quito, Pichincha, Ecuador

^c Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA, United States

^d Applied Research Laboratory, The Pennsylvania State University, University Park, PA, United States

^e Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, United States

ARTICLE INFO

Article history:

Received 22 January 2022

Accepted 1 February 2022

Available online 8 February 2022

Keywords:

Artificial intelligence

Collaborative design

Design teams

Engineering design

Human-computer interaction

Process management

Complex engineering systems

ABSTRACT

Human subject experiments are performed to evaluate the influence of artificial intelligence (AI) process management on human design teams solving a complex engineering problem and compare that to the influence of human process management. Participants are grouped into teams of five individuals and asked to generate a drone fleet and plan routes to deliver parcels to a given customer market. The teams are placed under the guidance of either a human or an AI external process manager. Halfway through the experiment, the customer market is changed unexpectedly, requiring teams to adjust their strategy. During the experiment, participants can create, evaluate, share their drone designs and delivery routes, and communicate with their team through a text chat tool using a collaborative research platform called HyForm. The research platform collects step-by-step logs of the actions made by and communication amongst participants in both the design team's roles and the process managers. This article presents the data sets collected for 171 participants

* Corresponding authors.

E-mail addresses: cagan@cmu.edu (J. Cagan), ccm@cmu.edu (C. McComb).

Social media: [@josh_gyory](https://twitter.com/josh_gyory) (J.T. Gyory), [@joncagan](https://twitter.com/joncagan) (J. Cagan), [@ccmccomb](https://twitter.com/ccmccomb) (C. McComb)

assigned to 31 design teams, 15 teams under the guidance of an AI agent (5 participants), and 16 teams under the guidance of a human manager (6 participants). These data sets can be used for data-driven design, behavioral analyses, sequence-based analyses, and natural language processing.

© 2022 The Author(s). Published by Elsevier Inc.
This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Specifications Table

Subject	Engineering
Specific subject area	Design teams, process management, artificial intelligence
Type of data	Table
How data were acquired	HyForm - a collaborative, online research platform [1] (www.github.com/hyform)
Data format	Raw data
Parameters for data collection	The experiment includes two conditions: (1) Human teams operate under the influence of an external human process manager; (2) Human teams operate under the influence of an external AI process manager.
Description of data collection	In the team experiment, teams of five participants designed and operated a fleet of drones to deliver packages to different locations in a specified market. During the experiment, HyForm, a collaborative research platform, collects step-by-step logs of the actions made by each participant and communications amongst the team.
Data source location	Institution: Carnegie Mellon University Region: Pittsburgh, Pennsylvania, 15,213 Country: United States of America
Data accessibility	With the article as supplemental material Hosted in the public data repository Repository name: Mendeley Data Data identification number: DOI: 10.17632/x7z48dvtbp.1 Direct URL to data: https://data.mendeley.com/datasets/x7z48dvtbp/1
Related research article	J.T. Gyory Mem ASME, N.F. Soria Zurita Mem ASME, J. Martin, C. Balon, C. McComb Mem ASME, K. Kotovsky, J. Cagan, Human Versus Artificial Intelligence: A Data-Driven Approach to Real-Time Process Management During Complex Engineering Design, Journal of Mechanical Design. 144 (2022). 10.1115/1.4052488 .

Value of the Data

- The data provide a step-by-step history of actions and communication of participants collaborating in a team trying to solve a complex engineering design problem. This data is important to continue research efforts in areas such as engineering design, team problem solving, computer science, psychology, and human-computer interaction.
- Researchers in cognitive science, artificial intelligence, and management may be interested in these data sets since they provide detailed process information on the design and operation of an engineered system by both teams and process managers.
- The data sets can be used for behavioral analyses, sequence-based analyses, and natural language processing to study the collaboration strategies process between human designers and to explore the effects of AI management on the design process and human problem-solving behaviors.

1. Data Description

The data sets collected from the experiment are included in the supplementary material as a ZIP file. In the Zip file the “Experiment Data” folder contains six CSV files named “Experiment Log”, “Experiment Chat”, “Pre-experiment Questionnaire”, “Post-experiment Questionnaire Human Process Managers”, “Post-experiment Questionnaire Teams_Human_Process_Manager_condition”, and “Post-experiment Questionnaire Teams_AI_Process_Manager_condition”, respectively. Descriptions of the column headings for the experiment Log CSV files are provided in [Tables 1](#) and [2](#). Descriptions of the column

Table 1
Description of columns in CSV file entitled “Experiment Log”.

Column header	Description
Record ID	This integer indicates the unique identification of the record.
Time	This string indicates the real-time data entry where the data log was recorded using the format [YYYY-MM-DD hh:mm:ss.s].
User ID	This string specifies the participant role identification in the team. The 3 letters of the identification point to the name of the database to which the team submits their solutions. The 2 numbers of the identification point to the participant team role.
Action	This string stores the composition of the action. Including the name of the action, the variables and their corresponding values associated with the action, and the user interface in which the action takes place.
Session	This integer indicates the session of the experiment, {1, 2}.
Condition	This string indicates whether the team has AI agents as Process Manager, {'AI', 'nonAI'}.
Team ID	This string indicates the unique team's identification.
Team Info	This string indicates the team information.
Role	This string indicates the role of the participant who takes the action or sends the chat, {'Problem Manager', 'Planner 1', 'Planner 2', 'Designer 1', 'Designer 2', 'Process Manager', 'Experimenter'}.
Action Name	This is string indicates the type and name of the action taken by the participant.
Drone String	This string indicates the drone design configuration. Please see Section 2.3 for more detailed information.
Position	This float indicates customer location along the X-axis and Y-axis in the mile unit.
Range	This float indicates the drone flying range in the mile unit.
Payload	This integer indicates the drone payload capacity in the pound unit.
Cost	This float indicates the drone cost in US dollars.
Velocity	This float indicates the drone's flying velocity in miles per hour.
Index	This float indicates the customer ID in the market
Selected	This string indicates whether the customer is selected, {'True', 'False'}.
Profit	This integer indicates the profit made through this plan in US dollars.
Operation-Cost	This float indicates the cost of the drone fleet operation in US dollars, calculated by summing up the fuel cost and distance traveled of all vehicles used in the delivery plan.
Start-Up Cost	This float indicates the start-up cost of the delivery plan in US dollars, which is determined by summing up the costs of every drone used in the plan.
Number of Deliveries	This integer records the number of parcels (both food and packages) delivered through this plan.
Weight Delivered	This integer registers the weight of parcels (both food and packages) in the unit of pounds delivered through the plan.
Package	This integer indicates the weight of the package in the unit of pounds delivered through this plan.
Food	This integer indicates the weight of food in the unit of pound delivered through this plan.
Value	This integer indicates the profit in US dollars that can be made if the parcel is delivered to the customer.

Table 2
Description of columns in CSV file entitled “Experiment Chat”.

Column header	Description
Record ID	This integer indicates the unique identification of the record.
Time	This string indicates the real-time data entry where the data log was recorded using the format [YYYY-MM-DD hh:mm:ss.s].
User ID	This string specifies the participant role identification in the team. The 3 letters of the identification point to the name of the database to which the team submits their solutions. The 2 numbers of the identification point to the participant team role.
Chat	This string contains the content of the chat entry.
Type	This string indicates if the information of the chat is an intervention from the Process Manager or a chat between team participants.
Session	This integer indicates the session of the experiment, {1, 2}.
Condition	This string indicates whether the team has AI agents as Process Manager, {'AI', 'nonAI'}.
Team ID	This string indicates the team's unique identification.
Team Info	This string indicates the team information.
Role	This string indicates the role of the participant who takes the action or sends the chat, {'Problem Manager', 'Planner 1', 'Planner 2', 'Designer 1', 'Designer 2', 'Process Manager', 'Experimenter'}.
Channel	This string indicates the identification of the chat channel.

headings for the questionnaires' CSV files are provided in [Tables 3, 4, and 5](#). The columns list the corresponding experiment questionnaire questions and possible answer ranges.

2. Experimental Design, Materials and Methods

In the experiment, participants using a collaborative research platform called HyForm¹ can create, assemble, evaluate, share their designs, and communicate with their team. More detailed information regarding HyForm can be found in previous work [\[1,2\]](#).

2.1. Experiment overview

Participants anonymously connect online through the HyForm research platform. Teams of five participants work together to build and implement a fleet of drones specializing in package and food delivery. To be successful, teams need to maximize the profit of deliveries in a specific target market. Each participant is assigned to one role in the team. The roles include one *Problem Manager*, two *Design Specialists*, and two *Operations Specialists*. Additionally, an external (does not directly work on solving the problem) *Process Manager* observes each team and can intervene to influence the team's design process. [Fig. 1](#) presents the team, and consequently communication, structures used in the experiment.

- The *Problem Manager* is responsible for managing the design team and is the communication link between the *Operations Specialists* and the *Design Specialists*. Additionally, this role is responsible for handling the team's budget, choosing the customers in the market, and approving or rejecting the final plans generated by the *Operation Specialist*.
- Two *Design Specialists* are responsible for designing drones and submitting completed drone designs to the *Operations Specialists*.
- Two *Operations Specialists* are responsible for developing delivery routes and drone fleets within the target market with the constructed drones to deliver parcels.

¹ <https://github.com/hyform/drone-testbed-server/releases/tag/2021-March-v2>.

Table 3
Description of columns in CSV file entitled “Post-experiment Questionnaire Human Process Managers”.

Column header	Description
Start Date	This integer indicates the start time of the questionnaire in the format of [MM-DD-YYYY hh:mm:ss.s].
End Date	This integer indicates the end time of the questionnaire in the format of [MM-DD-YYYY hh:mm:ss.s].
Finished	This string indicates if the participant completed the questionnaire, {'True', 'False'}.
What is your User ID:	This string specifies the participant role identification in the team. The 3 letters of the identification point to the name of the database to which the team submits their solutions. The 2 numbers of the identification point to the participant team role.
What is your role in the experiment?	This string indicates the role of the participant who takes the action or sends the chat, {'Problem Manager', 'Planner 1', 'Planner 2', 'Designer 1', 'Designer 2', 'Process Manager', 'Experimenter'}.
How would you rate the quality of the performance of the team you oversaw? - Quality (0–100)	An integer indicating the quality of the team performance, {'0'(Extremely Low) – '100' (Extremely High)}
How would you rate the cohesion (i.e., how the team worked together) of the team you oversaw? - Cohesion (0–100)	An integer indicating the cohesion of the team, {'0'(Not At All Cohesive) – '100' (Extremely Cohesive)}
Why did you rate the quality and cohesion of the team the way you did?	Open question
Did you feel that the interventions you used were effective (i.e., did they elicit the intended effects in your team?)	The effectiveness of the interventions {' Very effective', 'Slightly effective', 'Unsure', 'Slightly ineffective', 'Very ineffective'}
Did you feel constrained in the items you could intervene with?	{ ' Yes', 'Maybe', 'No'}
What else would you have liked to intervene with that you were not able to? Why?	Open question

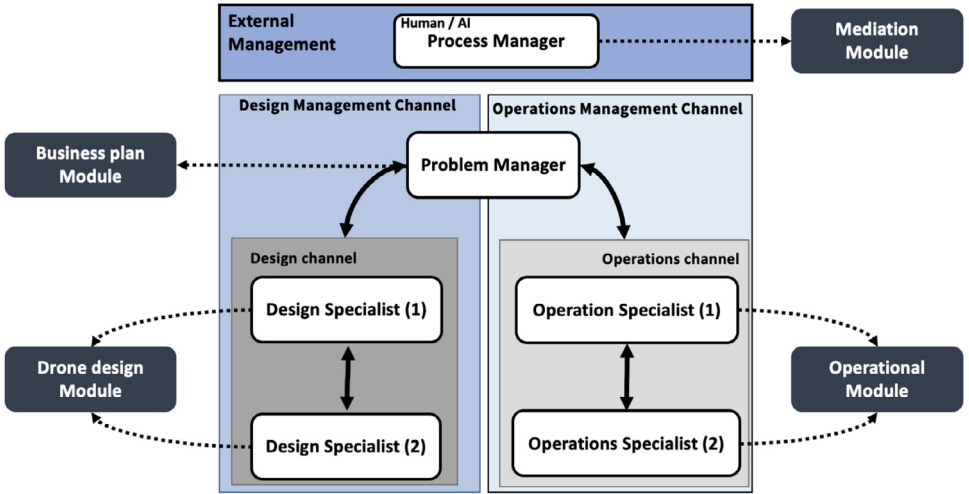


Fig. 1. Team structure. Only particular members can communicate with one another through explicit channels. The solid arrows represent the communication channels between team members, and the dashed arrows show the interaction between role and the respective module in HyForm. The experimental conditions place teams under the guidance of either an AI or a human process manager. Modified from [3].

Table 4
Description of columns in CSV file entitled “Post-experiment Questionnaire Teams_Human_Process_Manager_condition”.

Column header	Description
Start Date	An integer indicating the start time of the questionnaire in the format of [MM-DD-YYYY hh:mm:ss.s].
End Date	An integer indicating the end time of the questionnaire in the format of [MM-DD-YYYY hh:mm:ss.s].
Finished	A string indicating if the participant completed the questionnaire, {'True', 'False'}.
What is your User ID:	This string specifies the participant role identification in the team. The 3 letters of the identification point to the name of the database to which the team submits their solutions. The 2 numbers of the identification point to the participant team role.
What is your role in the experiment?	This string indicates the role of the participant who takes the action or sends the chat, {'Problem Manager', 'Planner 1', 'Planner 2', 'Designer 1', 'Designer 2', 'Process Manager', 'Experimenter'}.
How would you rate the quality of the performance of the team you oversaw? - Quality (0–100)	This integer indicates the quality of the team performance, {'0'(Extremely Low) – '100' (Extremely High)}
How would you rate the cohesion (i.e., how the team worked together) of the team you oversaw? - Cohesion (0–100)	This integer indicates the cohesion of the team, {'0'(Not At All Cohesive) – '100' (Extremely Cohesive)}
Why did you rate the quality and cohesion of the team the way you did?	Open question
Did you receive interventions from the process manager? (Yes - No)	(Yes - No)
Did you follow the interventions?	('Always', 'Most of the time', 'About half the time', 'Sometimes', 'Never')
Did you feel that the interventions were relevant? - Relevancy (0–100)	This integer indicates the relevancy of the interventions, {'0'(Not Relevant) – '100' (Extremely Relevant)}
Did you feel that the interventions were helpful? - Helpfulness (0–100)	This integer indicates the helpfulness of the interventions, {'0'(Not Helpful) – '100' (Extremely Helpful)}
Was the process manager sensitive to the needs of the team? - Sensitivity (0–100)	This integer indicates the sensitivity of the interventions, {'0'(Not Sensitive) – '100' (Extremely Sensitive)}
Do you think that your team could have benefited from any additional interventions from the process manager? If so, in what ways?	Open question
Did you feel that your team could have benefited from an intervention? If so, in what ways? If not, why?	Open question
Do you believe that the interventions you received were from a human process manager or an artificial intelligent (AI) process manager?	('Human', 'Artificial Intelligent (AI) system')
What characteristics of the interventions made you feel that way?	Open question

- The *Process Manager* is responsible for examining the team’s actions in real-time and can provide recommendations at specific times during the experiment. The Process Manager is not part of the team and cannot communicate with the participants in the team. Instead, the process manager guides the team with a set of prescribed, process-related interventions from a predefined list, thereby controlling the types of interventions for consistency across managers. In this experiment, the process manager is either a human or an AI agent. More

Table 5

Description of columns in CSV file entitled "Post-experiment Questionnaire Teams_AI_Process_Manager_condition".

Column header	Description
Start Date	This integer indicates the start time of the questionnaire in the format of [MM-DD-YYYY hh:mm:ss.s].
End Date	This integer indicates the end time of the questionnaire in the format of [MM-DD-YYYY hh:mm:ss.s].
Finished	This integer indicates if the participant completed the questionnaire, ('True', 'False').
What is your User ID:	This string specifies the participant role identification in the team. The 3 letters of the identification point to the name of the database to which the team submits their solutions. The 2 numbers of the identification point to the participant team role.
What is your role in the experiment?	This string indicates the role of the participant who takes the action or sends the chat, ('Problem Manager', 'Planner 1', 'Planner 2', 'Designer 1', 'Designer 2', 'Process Manager', 'Experimenter').
How would you rate the quality of the performance of the team you oversaw? - Quality (0–100)	This integer indicates the quality of the team performance, {'0'(Extremely Low) – '100' (Extremely High)}
How would you rate the cohesion (i.e., how the team worked together) of the team you oversaw? - Cohesion (0–100)	This integer indicates the cohesion of the team, {'0'(Not At All Cohesive) – '100' (Extremely Cohesive)}
Why did you rate the quality and cohesion of the team the way you did?	Open question
Did you receive interventions from the process manager? (Yes - No)	(Yes - No)
Did you follow the interventions?	('Always', 'Most of the time', 'About half the time', 'Sometimes', 'Never')
Did you feel that the interventions were relevant? - Relevancy (0–100)	This integer indicates the relevancy of the interventions, {'0'(Not Relevant) – '100' (Extremely Relevant)}
Did you feel that the interventions were helpful? - Helpfulness (0–100)	This integer indicates the helpfulness of the interventions, {'0'(Not Helpful) – '100' (Extremely Helpful)}
Was the process manager sensitive to the needs of the team? - Sensitivity (0–100)	This integer indicates the sensitivity of the interventions, {'0'(Not Sensitive) – '100' (Extremely Sensitive)}
Do you think that your team could have benefited from any additional interventions from the process manager? If so, in what ways?	Open question
Did you feel that your team could have benefited from an intervention? If so, in what ways? If not, why?	Open question
Do you believe that the interventions you received were from a human process manager or an artificial intelligent (AI) process manager?	('Human', 'Artificial Intelligent (AI) system')
What characteristics of the interventions made you feel that way?	Open question

information regarding the computational framework for the AI agent can be found in related work [3].

Each role has access to one of the following interface modules within HyForm. More detailed information about each module can be found in prior work [2–5].

1. The *business plan module* enables the Problem Manager to create market scenarios and evaluate business plans. The market scenario includes the locations of the customers, the number

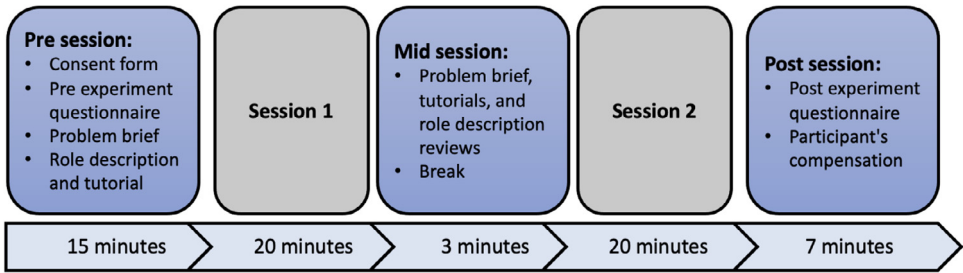


Fig. 2. Time allocation of the team experiment. Modified from [3–5].

- of customers for whom service can/should be provided, and the type of incentive each customer requires.
2. The *drone design module* enables Design Specialists to create and evaluate drones. Using this module, participants can verify through simulations the drones' feasibility, performance, and cost.
 3. The *operational module* enables the Operation Specialists to create and evaluate drones' delivery routes. The module provides through simulations estimates the cost of operation, total profit, time-to-deliver, the number of vehicles required, and time on/off station.
 4. The *mediation module* allows the Process Manager to observe, in real-time, the team discourse occurring through all communication channels and all the actions being performed by the team. Based on their observations, the Process Manager can intervene at specific times across the experiment session to transform the team's problem-solving strategy. The Process Manager can select an intervention to use from a specific prescribed set of interventions.

Participants recruited for the experiment are engineering undergraduate and graduate students at Pennsylvania State University (USA), with 171 participants in total. Participants are randomly assigned to 31 design teams across two experiment conditions, 15 teams under the guidance of an AI Process Manager (5 participants), and 16 teams under the guidance of a human Process Manager (6 participants, including the process manager). The experiment takes approximately 65 min to complete (Fig. 2), where participants have two 20 min design sessions to develop a drone fleet to deliver parcels in a market scenario. For each experiment session, each team has an initial budget of \$15,000 to build and operate a drone fleet and select which customers on the market scenario they want to acquire to obtain higher profit. Halfway through the experiment, during Session 2, the customer market is changed unexpectedly, requiring teams to adapt their strategies. In this scenario, the customers with low weight medical deliveries are included to the market, and the drone cost is reduced by 30%. During the market change, participants maintain the assigned role in the experiment. Participants complete a questionnaire at the beginning and at the end of the experiment. Participants with the Process Manager role complete a different post-study questionnaire tailored to the role.

2.2. Overview of process manager

During the two sessions of the experiment, an external Process Manager can intervene to influence the design teams' performance. The Process Manager can observe, in real-time, features of the teams' design process and can provide prescribed recommendations at specific times during the experiment. During the experiment, the team members cannot communicate with the Process Manager. Table 6 shows the list of interventions predefined for this experiment. The Process Managers can intervene at 2.5 min intervals, with the first available intervention starting at 5 min into each session. The intervention is transferred to the team through a specific communication channel, presented in the right-hand column of Table 6. For this experiment, the two

Table 6
List of predefined interventions available to the Process Manager.

Intervention	Communication Channel
Ops. planners, it would be good to continue working on and refining your plans a bit more.	Operations
Hey operations team, I suggest that you try evaluating and submitting your plan and starting fresh.	Operations
Hey operations team, try running the path-planning agent to help.	Operations
Drone designers, it would be helpful if you can continue working on and refining your drone designs a bit more.	Design
Hey drone design team, I would recommend evaluating and submitting your current design and starting fresh.	Design
Hey drone design team, check out the suggestions from the drone design agent.	Design
No intervention	NA
Team, I think you should try focusing more on adjusting the design parameters to meet the goals of the problem, and share this with each other (cost, capacity, speed, budget, weight, etc.)	Design, Operations, and Problem Manager
Team, try focusing more on your strategy. Try optimizing and increasing/decreasing size of components and share this with each other.	Design, Operations, and Problem Manager
Hi team, try sharing your goals with each other a bit more and make sure they're aligned.	Design, Operations, and Problem Manager
Ops team, please try to communicate with each other more.	Operations
Drone designers, please try to communicate with each other more.	Design
Hi problem manager, please try to communicate with your team more.	Problem Manager

experimental conditions differ on whether the Process Manager is a human or an AI agent. Both Process Managers can choose from the prescribed interventions using the specific time intervals. More detailed information about the interventions and the AI Process Manager can be found in the prior work [3].

2.3. Grammar string format of drone configuration in log files

Each drone design configuration is recorded using a grammar node string in the experiment logfiles. Within the grammar node string, each node records information about the component identification, position, and size. In addition, the connection between two nodes creates edges with specific character identifiers. The grammar allows any drone design configuration to be recorded as a variable number of nodes and connections. More detailed information regarding the grammar string for the drone design can be found in previous work [2,4,5].

Ethics Statement

The authors declare that informed consent was obtained for experimentation with human subjects. These human subject experiments were approved by the Carnegie Mellon University Institutional Review Board (IRB) IRBSTUDY2015_00000042.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article.

CRediT Author Statement

Nicolas F. Soria Zurita: Investigation, Writing – original draft; **Joshua T. Gyory:** Investigation, Data curation, Writing – original draft; **Corey Balon:** Resources; **Jay Martin:** Resources; **Kenneth Kotovsky:** Supervision, Writing – review & editing; **Jonathan Cagan:** Supervision, Writing – review & editing; **Christopher McComb:** Supervision, Writing – review & editing.

Acknowledgments

This material is based upon work supported by the [Defense Advanced Research Projects Agency](#) through cooperative agreement N66001–17–1–4064. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

Supplementary Material

Supplementary material associated with this article can be found in the online version at <https://data.mendeley.com/datasets/x7z48dvtbp/1>.

References

- [1] GitHub – hyform/drone-testbed-server: backend server tools for the drone delivery problem, (n.d.). <https://github.com/hyform/drone-testbed-server>. Accessed January 10, 2022.
- [2] B. Song, N.F. Soria Zurita, G. Zhang, G. Stump, C. Balon, S.W. Miller, M. Yukish, J. Cagan, C. McComb, Toward hybrid teams: a platform to understand human-computer collaboration during the design of complex engineered systems, in: Proceedings of the Design Society: DESIGN Conference, 1, 2020, pp. 1551–1560, doi:[10.1017/dsd.2020.68](https://doi.org/10.1017/dsd.2020.68).
- [3] Mem ASMEMem ASMEMem ASME J.T. Gyory, N.F. Soria Zurita, J. Martin, C. Balon, C. McComb, K. Kotovsky, J. Cagan, Human Versus artificial intelligence: a data-driven approach to real-time process management during complex engineering design, J. Mech. Des. 144 (2022), doi:[10.1115/1.4052488](https://doi.org/10.1115/1.4052488).
- [4] B. Song, J.T. Gyory, N.F. Soria Zurita, G.M. Stump, J.D. Martin, S.W. Miller, C.M. Balon, M.A. Yukish, C. McComb, J. Cagan, Decoding the agility of human-artificial intelligence hybrid teams in complex problem solving, In Submission. (2021), doi:[10.1016/j.destud.2022.101094](https://doi.org/10.1016/j.destud.2022.101094).
- [5] G. Zhang, N.F. Soria Zurita, G. Stump, B. Song, J. Cagan, C. McComb, Data on the design and operation of drones by both individuals and teams, Data Brief 36 (2021) 107008, doi:[10.1016/j.dib.2021.107008](https://doi.org/10.1016/j.dib.2021.107008).