Reporting bias occurs when the frequency of events, properties, and/or outcomes captured in a data set does not accurately reflect their real-world frequency. This bias can arise because people tend to focus on documenting circumstances that are unusual or especially memorable	**EXAMPLE A**: A sentiment-analysis model is trained to predict whether book reviews are positive or negative based on a corpus of user submissions to a popular website. The majority of reviews in the training data set reflect extreme opinions (reviewers who either loved or hated a book), because people were less likely to submit a review of a book if they did not respond to it strongly. As a result, the model is less able to correctly predict sentiment of reviews that use more subtle language to describe a book.
Automation bias is a tendency to favor results generated by automated systems over those generated by non-automated systems, irrespective of the error rates of each.	**EXAMPLE I**: Software engineers working for a sprocket manufacturer were eager to deploy the new "groundbreaking" model they trained to identify tooth defects, until the factory supervisor pointed out that the model's precision and recall rates were both 15% lower than those of human inspectors.
Selection bias occurs if a data set's examples are chosen in a way that is not reflective of their real-world distribution. One type of selection bias -- **coverage bias** -- occurs when data is not selected in a representative fashion.	**EXAMPLE C**: A model is trained to predict future sales of a new product based on phone surveys conducted with a sample of consumers who bought the product. Consumers who instead opted to buy a competing product were not surveyed, and as a result, this group of people was not represented in the training data.

Selection bias occurs if a data set's examples are chosen in a way that is not reflective of their real-world distribution. One type of selection bias -- **non-response bias**, occurs when data are unrepresentative due to participation gaps in the data collection process.	**EXAMPLE F**: A model is trained to predict future sales of a new product based on phone surveys conducted with a sample of consumers who bought the product and with a sample of consumers who bought a competing product. Consumers who bought the competing product were 80% more likely to refuse to complete the survey, and their data was underrepresented in the sample.
Selection bias occurs if a data set's examples are chosen in a way that is not reflective of their real-world distribution. One type of selection bias -- **sampling bias**, occurs when proper randomization is not used during data collection.	**EXAMPLE D**: A model is trained to predict future sales of a new product based on phone surveys conducted with a sample of consumers who bought the product and with a sample of consumers who bought a competing product. Instead of randomly targeting consumers, the surveyor chose the first 200 consumers that responded to an email, who might have been more enthusiastic about the product than average purchasers.
Group attribution bias is a tendency to generalize what is true of individuals to an entire group to which they belong. One manifestation of this is **in-group bias:** A preference for members of a group to which *you also belong*, or for characteristics that you also share.	**EXAMPLE H**: Two engineers training a resume-screening model for software developers are predisposed to believe that applicants who attended the same computer-science academy as they both did are more qualified for the role.

Group attribution bias is a tendency to generalize what is true of individuals to an entire group to which they belong. One manifestation of this is **out-group homogeneity bias**: A tendency to stereotype individual members of a group to which *you do not belong*, or to see their characteristics as more uniform.	**EXAMPLE E**: Two engineers training a resume-screening model for software developers are predisposed to believe that all applicants who did not attend a computer-science academy do not have sufficient expertise for the role.
Implicit bias occurs when assumptions are made based on one's own mental models and personal experiences that do not necessarily apply more generally.	**EXAMPLE G**: An engineer training a gesture-recognition model uses a head shake as a feature to indicate a person is communicating the word "no." However, in some regions of the world, a head shake actually signifies "yes."
A common form of implicit bias is **confirmation bias**, where model builders unconsciously process data in ways that affirm preexisting beliefs and hypotheses. In some cases, a model builder may actually keep training a model until it produces a result that aligns with their original hypothesis; this is called **experimenter's bias**.	**EXAMPLE B**: An engineer is building a model that predicts aggressiveness in dogs based on a variety of features (height, weight, breed, environment). The engineer had an unpleasant encounter with a hyperactive toy poodle as a child, and ever since has associated the breed with aggression. When the trained model predicted most toy poodles to be relatively docile, the engineer retrained the model several more times until it produced a result showing smaller poodles to be more violent.