PREDICTIVE ANALYTICS
STRATEGIC COUNCIL

Applying Predictive Analytics in Construction

Whitepaper: Applying Predictive Analytics in Construction

By The Predictive Analytics Strategic Council

April 1, 2020

PASC Operations Workstream Leads:

Andrew Burg, Messer Construction

Timothy Gattie, Smartvid.io

Member Firms:

Aon	Ferrovial	Shawmut Design and Construction
AXA-XL	JE Dunn Construction	Skanska (Sweden)
Bagatelos Architectural Glass Systems	Lithko Contracting	Skanska (USA)
Barton Malow Company	Messer Construction	Suffolk Construction Company
Bouygues Construction	Mortenson Construction	VINCI Construction
DPR Construction	Obayashi Corporation	Webcor Builders

Table of Contents

1.0 Outline & Key Concepts

1.1 Document Goal

1.2 Background

1.3 Why this Document is Needed

1.4 History's Lessons on Technology Adoption Challenges

1.5 Limits of Predictive Analytics

2.0 Technology Rollout Considerations

2.1 Rollout Decision Trees

2.2 Predictive Results vs. Report Values

3.0 Data Gathering & Transparency

3.1 Data Gathering Best Practices

3.2 Data Transparency

3.3 Common Digital Tools

4.0 Executive-Level Decision Making using Predictive Results

4.1 Executive Level Considerations

4.2 Executive Level Tools

5.0 Project-Level Decision Making using Predictive Results

5.1 Project Level Considerations

5.2 Project Level Tools

6.0 Conclusion

Appendices

Appendix A - Data Gathering Technology Implementation Decision Tree

Appendix B - Implementing Predictive Analytics Into Existing Processes Decision Tree

Appendix C - Glossary of Key Terms

Appendix D - Recommended Reading

Appendix E - Acknowledgements & Legal Disclaimer

E.1 - Acknowledgements

E.2 - Legal Disclaimer

Applying Predictive Analytics in Construction

1.0 Outline & Key Concepts

1.1 Document Goal

The goal of this whitepaper is to produce guidelines on how to incorporate predictive analytics outputs within existing operational processes of a construction firm. The intended audiences are company executives, project managers, and field teams, all of whom can be positively affected by having more predictive data on project risk, if presented properly. While this document focuses on the specific example of determining the risk of incurring a safety incident, the concepts can be extended to other applications of predictive analytics.

1.2 Background

This whitepaper is the result of collaboration among the membership of the Predictive Analytics Strategic Council (PASC) and is presented to the industry for comment and feedback. Membership of a company in the PASC does not imply that any specific member company has officially or formally endorsed the views presented in this whitepaper (See Legal Disclaimer in Appendix E). For more on the PASC, visit www.pasc.ai.

1.3 Why this Document is Needed

Predictive analytics is changing many industries. For example, retailers use it in customer sentiment analysis, banks and insurers use it in risk modeling and manufacturers use it in predictive maintenance. You can find links to resources and articles discussing these and other applications of predictive analytics in Appendix D - Recommended Reading.

Despite the use of predictive analytics in other industries, the application of predictive analytics is just beginning in construction. This whitepaper attempts to provide guidance on how new metrics generated by predictive analytics can be incorporated into the executive and project level operational processes currently employed by contractors, both large and small.

1.4 History's Lessons on Technology Adoption Challenges

Any technology is only as useful as the processes and people around it - the system of which it is a part. Ignoring any one of the three will result in suboptimal results. With regard to technology adoption, it is imperative to consider the affected people and modifications to the process (see Digitization and Digital Transformation in Appendix C).

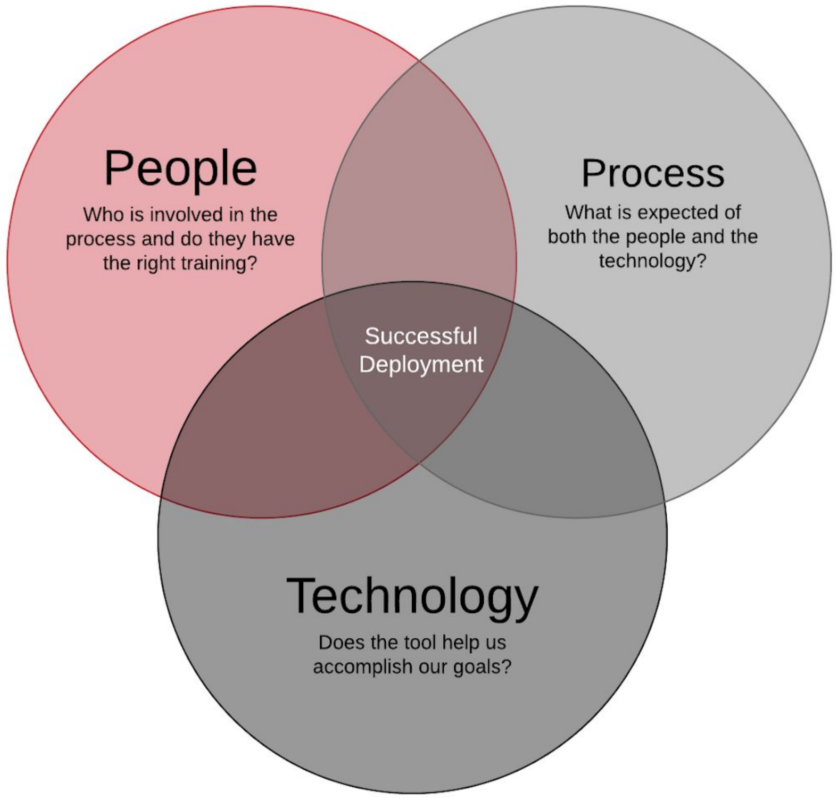


Figure 1 - Defining an approach to building a process, developing the people and defining the goals for technology are all key ingredients to a successful technology deployment. For purposes of this paper, a successful deployment includes consistent capture of a uniform data set across the organization that supports building an accurate predictive model.

A new technology must also prove itself. By way of illustration, it is absurd to think of double checking the results that come from a calculator using long hand calculations now, but when calculators first emerged as a business tool, that process was commonplace. This hardly resulted in productivity gains, that is until the calculator was proven to be a trustworthy tool.

Unlike a calculator however, employing a process to double check the results of a predictive model is not straight forward. The results from predictive models are produced by patterns established in the data we feed them and do not come from a simple executable formula. Put another way, they are only as good as the data we build them from and the continued performance of those patterns of data going forward. From the Artificial Intelligence Index Report 2019: Only 19% of large companies surveyed say their organizations are taking steps to mitigate risks associated with explainability of their algorithms…" With this understanding, a vast majority of models currently deployed lack explainability but this trade-off is accepted in light of the value the models bring.

The good news is that predictive model accuracy can be measured and described. This allows us as users of predictive outputs to apply judgement based on the level of confidence reported by the model.

Food for Thought

We are investing time, money and resources into building these complex systems so they can provide insights that we wouldn't otherwise have. If the systems were simple enough for us to easily understand the results, they would be too simple to be valuable.

History also tells us that three things are required for the successful rollout of a new technology:

Executive Commitment:

There must be support from the top of the organization in the form of a vision of what the future state using the technology will look like. Like most technologies, the trend will be to immediately try to weigh the ROI and intrinsic value of the tools against effort to implement them. The challenge will be putting a value on not encountering something that you might have otherwise encountered.

A good example of this is BIM modeling and clash detection. The cost of the investment to build the BIM model could be estimated. The value of finding and resolving all of the clashes inherent in the design was more difficult to enumerate. After all, if the cost of the clashes were known, they could have been priced into the bid to begin with!

Likewise, adoption of a predictive model requires a similar comparison. The cost to deploy the data gathering technology and to build the model can be estimated. With model accuracy being highly dependent on data quality, the value must be based on the belief that the data can be collected accurately and that the patterns within the data exist to generate valuable predictions. Fortunately, we know that there are patterns within safety data that enable predictions as proven by deployments outside of the PASC. This means we are left with ensuring the data collection process is accurate and repeatable across an organization.

Factors that should be considered in an ROI calculation for the model subject of this whitepaper include:

a) The cost of deploying the data collection system
b) The cost of building the predictive model
c) The value generated by reducing the number incidents including direct savings associated with not having the incident and indirect savings such as resulting fewer meetings, fewer work stoppages, less reporting, improved productivity, ect.
d) There is some anecdotal evidence indicating that predictive models result in lower severity claims when incidents do occur.
e) Tertiary effects of improving safety performance such as higher morale, attracting and retaining higher quality personnel, improving the firm's data systems and the ability to win more work as a result of these effects.

Internal Public Relations:

There must be a solid internal public relations campaign to communicate the following:

The rollout plan for technology adoption (there may be multiple technologies adopted).

Successes achieved along the way.

Pre-established and clear goals that each user must meet.

The benefits users will experience as a result of adoption.

Incentive:

There must be an incentive structure in place to drive adoption. Incentives will vary with firm culture but might include all or some of the following:

Connecting a portion of the user’s performance bonus to use of the new tool/process

Awards and/or recognition for those who meet or exceed the pre-established and clearly communicated goals.

1.5 Limits of Predictive Analytics

It’s important to set the correct expectations on the power of and the challenges associated with predictive outputs. The outputs of predictive models are indeed different from traditional reports as predictive results are predicting a specific outcome vs. showing a trend derived from actual results. Since descriptive analytics are showing trends, they are by nature deterministic, while predictive analytics are by nature probabilistic. This difference represents a paradigm shift in how managers will interpret and interact with these different types of analytics.

Additionally, the “confidence” with which the predictive engine generates alerts can be adjusted to match the user’s tolerance for false alerts (either positive or negative, depending on the model). For example, the model contemplated in this document produces predictions on which projects are most likely to have a safety incident in a defined period of time. The users must balance their desire for accuracy (e.g. only reporting the incidents with a very high confidence) with their desire to know about incidents even if there is only a slightly elevated confidence that one will occur (e.g. report on all potential incidents, or low confidence). In this case, since having an incident can be a severe event, one may choose to err on the side of low confidence.

2.0 Technology Rollout Considerations

2.1 Rollout Decision Trees

The successful rollout of technology is highly dependent on the details associated with that rollout:

a) Is there a demonstrated need for change and support for that change at high levels of the organization?
b) Is an existing process being supplanted or does the technology represent an entirely new process that must be adopted?
c) How many people in the organization does the rollout directly affect?
d) What level(s) of the organization is/are most impacted?
e) What level of effort will be required to train users and support the roll out?
f) What data is required to support the desired outcome?

The case of deploying predictive results is somewhat simplified compared to data gathering since the technology primarily impacts the interpretation of results based on data that is already gathered in organizational processes. We are all experts at reading reports. Where we need assistance is in the deployment and interpretation of predictive results.

The challenges associated with changing the processes by which data is gathered on the other hand, are far more complicated however. Refer to the Data Gathering Technology Implementation Decision Tree in Appendix A.

For the case of deploying and implementing predictive analytics that are aimed at predicting the risk of safety incidents on construction projects into existing processes, refer to the Implementing Predictive Analytics Into Existing Processes Decision Tree in Appendix B.

> ### Food for Thought
>
> *Don't treat predictive models as if they are rational human opinions. Also, don't treat them as if they are simple summed columns in a spreadsheet that can be easily verified. Rather, treat them as insights to be used to enhance decision making but not make the decision.*

2.2 Predictive Results vs. Report Values

Predictive values are outputs from predictive models. Report values are values generated in reports - these are the values we are used to looking at on a daily, weekly, monthly and annual basis. Predictive values differ from report values in several ways. Below are some of the differences:

<u>Report Values:</u>

The values we are used to seeing in reports are based on actual outcomes and measures. While we may question whether they accurately reflect reality, we are often able to describe why they may appear inaccurate (measured precisely, coded correctly, entered by the cut-off, accrued properly, etc.).

Examples of report values include man hours for the past week, cost incurred this month, yearly revenue, etc.

<u>Predictive Results:</u>

Predictive results are different from the values we are accustomed to seeing in reports because they use patterns from training data to extrapolate a result. Predictive results have an associated level of confidence (low, medium or high). For this reason, predictive results should not be treated the same as typical values found in reports. It is also recommended that, if predictive results are mixed with report values, they would be differentiated or noted as such using either italic text or an asterisk.

Unlike report values, the inaccuracy of predictive results cannot usually be simply explained away by a breakdown in process. One must look first at the model input data to ensure it is correct. If it is, then an anomalous result can only be explained in one of three ways:

a) The patterns contained in the most recent model inputs are unlike any the model was trained on resulting in unpredictable model outputs;
b) The model is in need of maintenance or retraining that includes more recent data; or
c) The predictive result is actually representative of what is about to happen.

Any AI output should be filtered through a human's understanding. If there is a predictive result showing a project that has no people working on it will have an incident then a human must intervene and apply understanding that this is a probable false positive result. Error analysis can be applied to outcomes like this after the fact to see how such results can be remedied.

Predictive results can come from one of several types of models. As the models grow in difficulty to build, the outputs become more valuable. The most basic of these models are Descriptive Analytics merely reflecting the values collected and thus, providing hindsight and leaving the user to both deduce insight and decide how to respond. Prescriptive models, while more difficult to build, provide insight but it is left to the user to determine what actions are necessary, if any. Finally, prescriptive models, the most complex, provide the user with both a prediction and the actions necessary to obtain the desired outcome (foresight).

Figure 2 shows the three types of models.

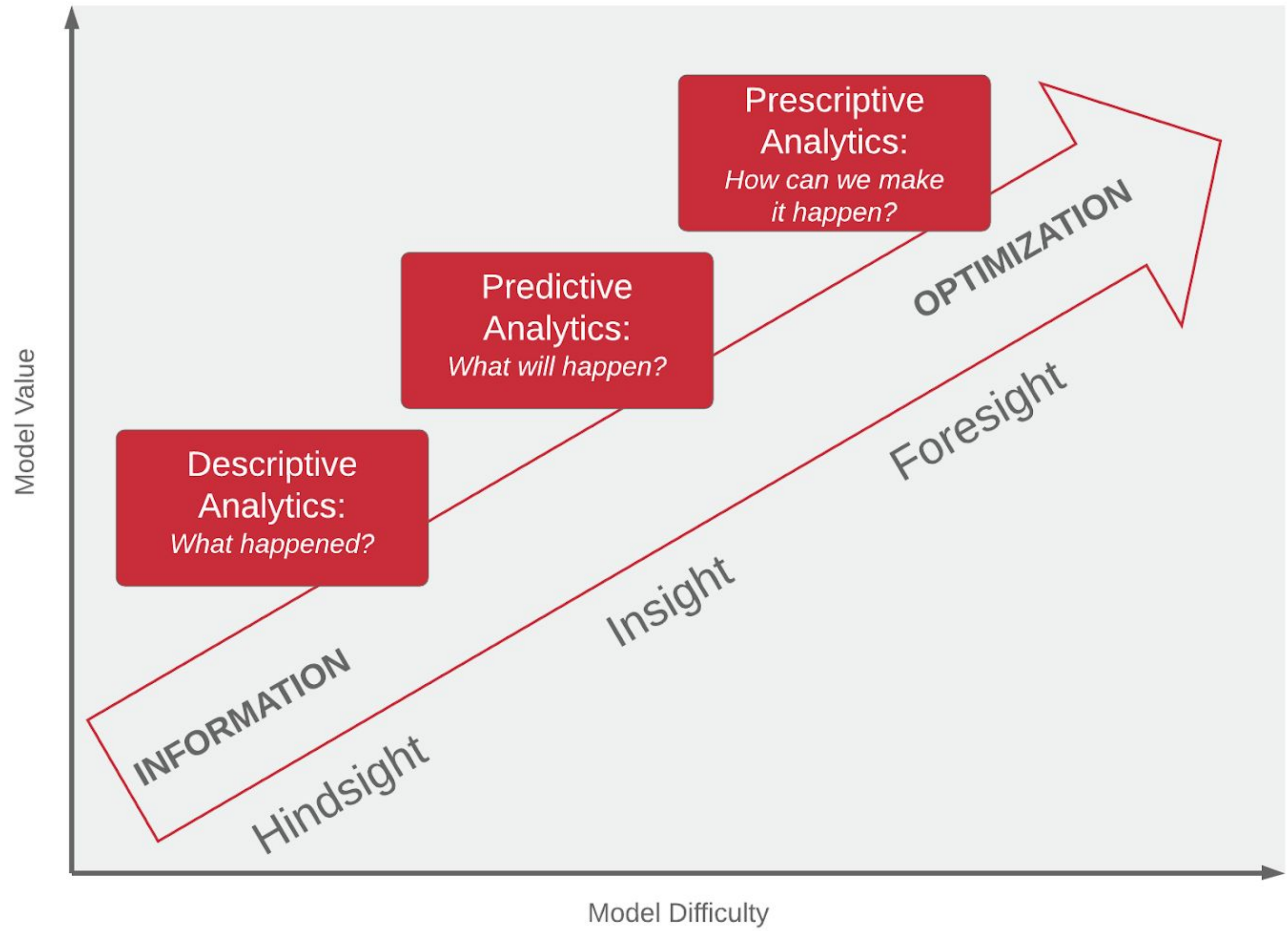


Figure 2 - Graph describing the different types of models and their value. Adapted from Gartner Research.

3.0 Data Gathering & Transparency

3.1 Data Gathering Best Practices

Gathering data is like baking a cake, you're not assured good results until it's done baking. Fortunately, there is a recipe to follow that can help us achieve the results we are looking for. That recipe comes in the form of best practices.

These best practices are intentionally meant to be broad and apply to data and systems outside the scope of this white paper.

When gathering data, first evaluate these questions:

- **What problem is the data trying to solve?**

  It's important to define why the data is being collected and what solution that it is expected to contribute to.

- **How should the data be structured (for structured data)?**

  The second question to ask after determining what problem the data is trying to solve is how the data should be organized. Often, the data needed for predictive models involves more than a single source. That being the case, it is imperative to understand the schema (a diagram how the data must be structured and related). For example, project data from the project management software might be related to a table of safety observations data collected on various projects. In this case, some form of unique Project ID must reside in both tables. These tables can then be joined in a one-to-many relationship on Project ID.

- **How is the data being collected?**

  Manually entered data (e.g. through a form field input) should be kept to an absolute minimum. Collecting manually entered data simply to have it can lead to overburdening the users (people who collect the data) and poor quality results (pencil whipping, etc.).

  Even if data can be gathered automatically (e.g. gathering weather information using GPS and datetime data collected when completing a form), having volumes of data without purpose is cumbersome to manage.

- **How does the data being entered serve the person entering it?**

  When data is collected manually (e.g. through a form field input), it should serve the user in some way (e.g. time saved entering data elsewhere or reducing incidents which eliminates the need to complete incident reports).

- **How should the data be entered?**

  For manual data collection (e.g. through a form field input), careful consideration should be given to the type of field used to gather the data. Drop-down boxes, lists and checkboxes are far more likely to gather consistent, accurate data through their inherent validation.

  Using a free-form text field should be a last resort. For example, using a free text box to collect the name of a trade partner could result in numerous different iterations of each trade partner's name (e.g. Al Brown Construction, Inc. could become ABCI, AB Construction, ABC, Al B Con, etc.)

- **How do we train for good data gathering?**

  When a tool for gathering data is deployed, it is imperative that a minimum level of training be provided. The minimum level of training includes:

  a) A description of why the tool is being used (what the end goal is).
  b) Who is responsible for collecting the data and who has access to it. (people)
  c) What data is to be gathered with the tool, and what data is not. For example, a system that is intended to gather data for incident claims should not be used to gather data for near misses / near hits. (process)
  d) A demonstration of how to complete data entry and what the user can expect to get out of collecting the data (technology)
  e) Training on any behaviors associated with using the tool. For example, if the tool gathers sensitive or private information, there are certain behaviors associated with keeping the data private and secure. (people and process)
  f) When the tool is to be used or how many entries each user is required to complete, if applicable. (process)

Depending on the level of change that use of the tool represents to the organization, this training can be accomplished using a handout/email or it may be complex enough to require online/self-guided or in-person training.

- **Are there certain behaviors associated with good data gathering?**

  By way of illustration, assume an organization collects safety observations that are rated by the level of risk (1 - Serious, 2 - Moderate & 3 - Minor). Whenever a user enters an observation using the risk rank of 1, the project team receives a call from senior management inquiring about the observation and questioning how safe the project is. How many entries using the risk rank 1 do you think the observation system will collect? Not many.

  The process and behaviors surrounding the data must be crafted in a way to promote "truth in the data". If there is a barrier to, or punishment for, collecting objective data, the users will always change their behavior to avoid the pain and/or punishment. It is important to treat data as information and not react to it negatively. Ignoring this fact will result in poor data quality.

3.2 Data Transparency

Once the decision is made to collect data, the next requisite decision that must be made is who gets to see it. Because there are many schools of thought on this topic and because the way in which data is handled is highly dependent on the content of the data itself, this section is narrowly focused on the data and outputs used to predict safety incidents.

There are other data, metrics and models that require discretion. Those reside outside of the recommendations below.

- **Who should see the data and outputs?**

  Unless there is a particularly persuasive reason to restrict access, the source data, metrics and predictive results from the model subject of this whitepaper should be shared as widely as possible within the organization. It is especially important for anyone required to gather data used in the prediction of incidents to have access to the outputs. Preventing incidents is an "all hands on deck" activity. To that end, sharing the predictive results helps get everyone on the same page working together. It also demonstrates the value of collecting the data from an end user perspective provided that it is used correctly and does not cause the user to manipulate the data entry to intentionally affect the model output.

This means that everyone may have access to the safety observations and safety incident information (without private personal or HIPAA restricted information) and metrics. It also means that everyone may have access to the relative project risk rankings.

This requires a certain “organizational maturity” and necessitates that those who have access to the data and results be trained.

- **Now that we can see more, how do we react to all of the new information we now have?**

  By way of illustration, the construction industry is akin to living by candlelight. Since most contractors view their data via reports and graphs and have not yet connected all of their sources of data together, it is like living by candlelight - not able to see the whole picture. Once data is organized in a way that business intelligence tools can be applied to it, the experience is similar to seeing the night lit up by stadium lights for the first time - it can be overwhelming, and enlightening.

  It is a natural response to restrict access to such a powerful tool or use it to identify and punish underperforming areas of the business. Rather than reacting in this way, consider challenging the organization to rise to the occasion. It is likely that the culture will grow along with performance.

Food for Thought

“With great power comes great responsibility”

- Uncle Ben from Spiderman

- **Are there certain behaviors associated with being able to access the data?**

  If you're reading this, you most likely have a driver's license. You know that driving is a privilege and not a right. You are aware that there are certain rules you must follow in order to keep your driving privileges.

  Accessing data and outputs is just like driving. It is up to the organization to define proper and improper uses of the data and outputs as well as the behaviors around using them but rest assured, access can be restricted if they are used improperly.

- **Is there training needed for viewers of the data and outputs?**

  Yes. Viewers should be training in the following at a minimum:

  a) Where the data comes from
  b) How the outputs are generated and their limitations (calculated metrics or predictive results)
  c) Expected uses of the data
  d) Behaviors associated with using the data and outputs
  e) Penalties for misusing the data

3.3 Common Digital Tools

There are countless tools used for the collection, connection, analysis and display of project data. The tools listed below are specific to the scope of this white paper - predicting safety incidents. The lists are not exhaustive, up to date or meant to be construed as recommendations. It is not the intent of this paper to be a reference for these products. There are a variety of sources for finding out more about construction technology. (See Appendix D - Recommended Reading)

- **Tools for collecting data**

  Observations:

  Products: Procore Observations, Autodesk BIM360, Smartvid.io Safety Observations, Viewpoint Fieldview, iAuditor

  What to look for: Observations are different from quality inspections. While checklists work for quality inspections, they are not purpose built for collecting safety observations. Many products have attempted to kill two birds with one stone in this regard. Look for a product that is flexible, easy to use and gathers risk weighted observations. It's also important that the product makes it easy for you to access your raw data for data warehousing and modeling purposes.

Safety Incidents:

Products: Origami, iScout, Intelex, Aclaimant

What to look for: Incidents can occur unexpectedly and can be chaotic. It's important that an incident data collection tool be widely available, be easy to use and guide the user through the process of collecting good data. Look for an application that makes it easy to access your raw data for warehousing and modeling purposes. It's also nice if the incident data collection app can feed your claims management system.

Safety Inspections/Checklists:

Products: Viewpoint Fieldview, Autodesk BIM360, Predictive Solutions

What to look for: Checklists are a good tool for checking repetitive processes and verifying that program elements are in place but are not always a good tool to gather data for predictive analytics. This is due to the binary nature of checklist responses and because all of the context is locked inside the question. Having shades of gray or scale within the data in order to identify outliers and significantly correlated events is more valuable to the predictive model than all black and white (binary) values. With this understanding, checklists can serve an important process verification role if deployed correctly. A good checklist tool should be flexible and provide access to the data and context for analysis.

Visual Data:

Products: Struction Site, Reconstruct, Egnyte, OpenSpace, Smartvid.io Photo Monitoring

What to look for: In this case, the tool may be less important than the process employed by the organization. It is important for the organization to adopt a consistent and uniform method for collecting visual data over the entire project site on a regular (at least 2 times per week) basis.

Project Management Data:

Procore, Autodesk BIM360, Oracle Aconex, CMiC, Viewpoint, Prolog

What to look for: Most software in this category allows access to the raw data for analytics - this is the primary requirement for this tool aside from being a system that fits the needs of the organization.

- **Tools for connecting disparate sources of data**

  Products: Informatica, Dell Boomi, Workato, Jitterbit, SnapLogic

  What to look for: A product that connects to the organization's existing data sources and is easy to use.

- **Dashboards to display both gathered data and predictive outputs**

  Products: Microsoft Power BI, Tableau, Qlik, ThoughtSpot

  What to look for: A tool that connects to the sources of data that your organization uses is of paramount concern. After this, one that allows for access by the necessary users and provides visualizations that convey the data in a meaningful way.

4.0 Executive-Level Decision Making using Predictive Results

4.1 Executive Level Considerations

The accuracy of predictive results will vary depending on the quality of the data they are trained on. Some common questions that are asked when firms are evaluating the use of predictive results from models aimed at predicting safety incidents are listed below along with guidance on how to address them.

- **On which projects should we focus our attention?**

  Predictive results come in the form of a level of confidence. To avoid conflating the predictive results with the actual probability of an incident, the raw predictive results will not be displayed. Rather, the results are usually displayed by ranking the list of projects from most to least likely or by providing a qualitative result such as high, medium or low.

  The modeling that has been done in this area to date shows that the projects ranked higher or listed as high probability correspond to the projects that usually experience incidents. The level of accuracy of each model will vary but some models have been able to identify that the top 15% of projects in a ranked list have 80% of the incidents. Said another way, of a list of 40 projects, the top 6 projects listed have 80% of the incidents. These results are not representative of all models however.

  In summary, it may take some time for the organization to get comfortable with the level of accuracy of the deployed model but once a level of trust is achieved, the highest ranked projects will receive the attention.

- **What specifically should we focus our attention on with regard to the highest ranked projects?**

  Prescriptive models are able to provide feedback on what features are driving the result and actions that can be taken to prevent an incident from taking place. These action items are derived from the feature list that comprises the model. The available feature list is driven by the granularity of the data set provided.

  An example of an output from a prescriptive model is provided below in Figure 3.

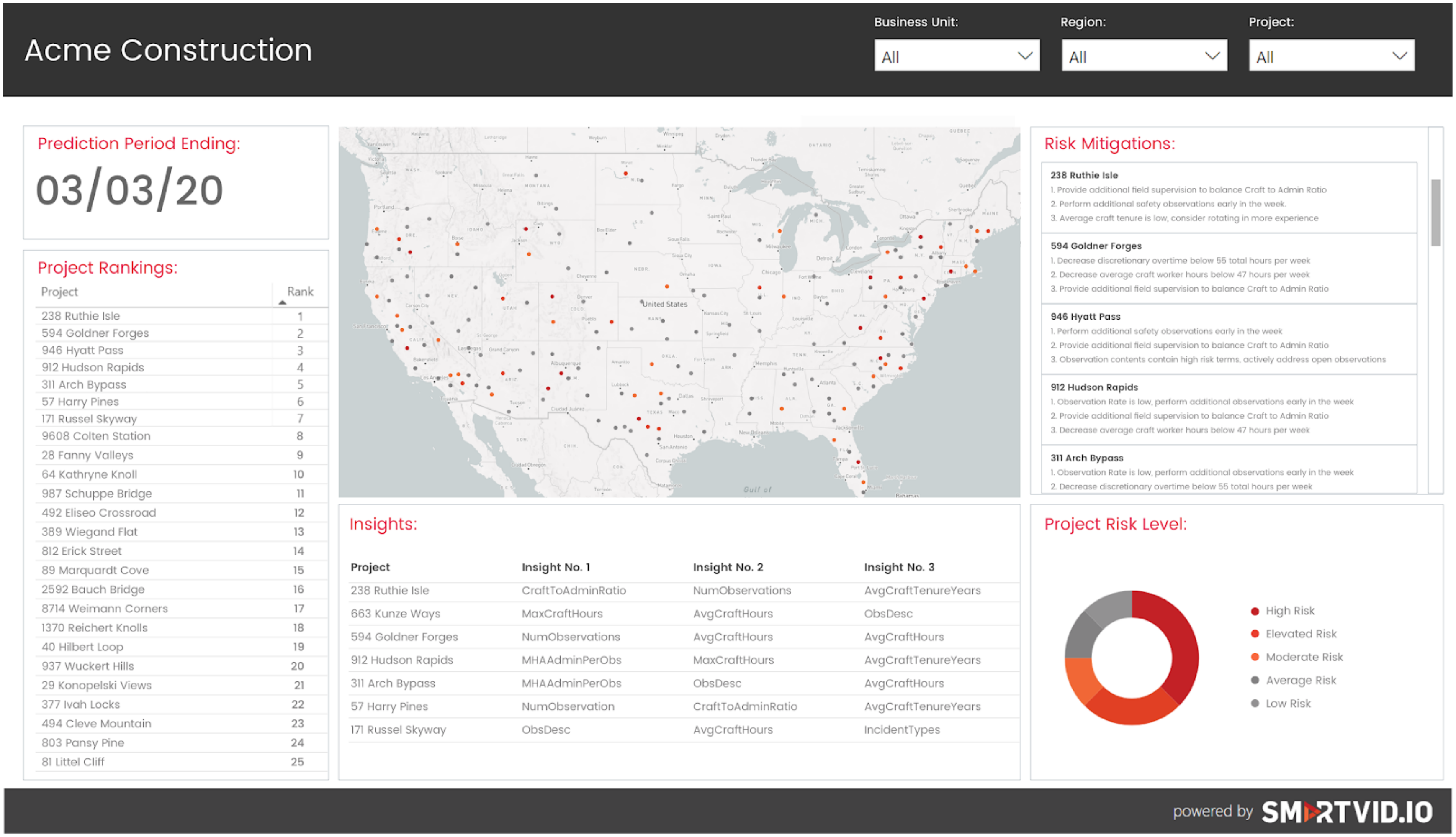


Figure 3 - Prescriptive model output includes a ranking or projects by risk level and actions that can be taken to prevent a potential incident.

The ability to build a prescriptive model is highly dependent upon the quality of the data on which the model is trained. Not all data sets will support prescriptive models. They may however support the development of a predictive model, a descriptive model or metric all of which may provide some idea of the level of risk.

- **Can we know what incident is going to happen?**

  While some data sets are able to be broken down to a level that will support the determination of an incident type, more often than not, the data just isn't there. This level of model granularity requires more form inputs when entering safety observations and incident reports. There is a balance to be struck between data gathering and the ability to build models that will predict events with high levels of granularity. More research is also required in this area.

- **Can we know what we are doing well, where we need to improve and where training is required?**

  While this is not a direct output from a model, there is much that can be learned from the model outputs. Trending the most common action items generated from a prescriptive model would provide insights into what areas the organization is underperforming. Conversely, the actions that are least often presented represent the areas where the organization is performing well.

  When the model features and metrics are compared on a project by project basis, one may also be able to identify areas of opportunity for improvement.

- **What known limitations should we keep in mind?**

  In addition to the limitations associated with interpreting predictive results described in Section 2.2 and the limitation inherent in the level of accuracy of the model, one must understand that these models are based on data usually collected by humans. With varying levels of quality on the input side, the outputs may experience a high degree of variability. This is why a well-defined process and user training are important to the use of predictive models.

- **How do we know if it's working? (e.g. what are the measures of success?)**

  With regard to the case of predicting which projects are most likely to have an incident, it is important to have a set of metrics based on "pre-predictive model deployment" performance. Some metrics might include:

  a) Number of incidents by type reported per week or per month
  b) Average duration between incidents of various types
  c) Maximum duration between recordable incidents
  d) A rating of the level of severity of the incidents that have taken place (if possible)
  e) OSHA metrics like TRIR, DART, LTIR, etc.
  f) Direct comparisons between projects implementing predictive and those that don't during a phased roll-out.
  g) Implementing predictive in the back-end without implementing the process on some projects and comparing results.

  With these metrics in place for the pre-predictive model deployment time period, the organization can continue to measure the same metrics and watch for change. For example, the pre-predictive model deployment measure of the duration between recordable incidents at one firm was 13.9 calendar days and 30.1 days after.

4.2 Executive Level Tools

This section provides recommendations on executive level tools to use and levers to pull when deploying predictive models. They are just that, recommendations. There are many ways to effectively distribute the results and respond to the predictive results. The best way to do this in your organization should be evaluated as part of your deployment.

<u>Predictive Result Visualization</u>

With data visualization and business intelligence tools becoming ubiquitous, it is recommended that the organization adopt an approach of deploying the predictive results via a dashboard. The top tools for this are listed in Section 3.3. It is recommended that the dashboard display the following information:

a) A ranking of projects in order of increasing risk of having an incident in the coming week or rating of projects by level of confidence (low, medium, high).

b) For each project, a list of the metrics and/or model features that contribute to the ranking/rating provided above (if the data supports this).
c) If the model is prescriptive, a list of the actions each project can take to prevent an incident from taking place.

Discretionary Resources

Depending on the size and sophistication of the organization, the executive team may have discretionary support resources at its disposal that can be deployed to the highest risk projects to help the project team execute on the predictive results. The people that make up these discretionary resources should be trained on the interpretation of predictive results, the specific model limitations as well as on the techniques that can be employed to reduce risk. These discretionary resources may become capacity constrained and this tool would focus their attention on the area of greatest opportunity to make best use of their time.

5.0 Project-Level Decision Making using Predictive Results

5.1 Project Level Considerations

Once the model is deployed and the results are available to project level personnel, there are several outcomes to consider:

- **Has our job been flagged as being at risk for an incident by the predictive model?**

  Each week, the project team should review the updated predictive results. This should happen prior to any safety meetings at the start of the week. The predictive results should be used to guide some of the talking points in the agenda - regardless of where the project is ranked or how it is rated.

  > **Food for Thought**
  >
  > *The project needs to be armed with the same information that high level, back office folks have. They are the closest to the work and quickest to take action toward reducing the risk.*

- **What can we do about it?**

  For high and medium risk projects, the team should review the metrics and action items with the aim of assigning the development of action plans to address each action item. If discretionary resources are needed (and available), the project team should request the type of support needed to execute their action plan(s). The fact that the project has been flagged as high or medium risk for the week should be shared with the crews to elevate their level of engagement with respect to safety.

  For low risk projects, the team should share metrics and measures where they performed well in the prior week - stressing that the whole project team is contributing to reducing risks on their project. Any action items provided in the predictive results should have an action plan and be addressed as if the project were rated at medium or high risk.

- **Who is responsible for doing something about it?**

  The project team should establish a process to track their action items, the action plans to address them and mark them completed once the action plan is executed. The Project Manager and site safety representative should assign the action items to their teams and verify completion. If discretionary resources are provided, they should report to the Project Manager and provide any assistance or training needed.

  The Corporate Safety Director should ensure that project teams cross-share ideas, best practices, tactics and plans that were successful in mitigating their risk issues.

- **How do we know if it's working? (e.g. what are the measures of success?)**

  These measures should be consistent with those established at the organizational level in order to promote cross-project comparison. Should the project team have the resources to track additional metrics, they might choose some of the follow (in addition to the organizational metrics):

  a) Number of incidents by type reported per week or per month
  b) Average duration between incidents of various types
  c) Maximum duration between recordable incidents
  d) A rating of the level of severity of the incidents that have taken place (if possible)
  e) OSHA metrics like TRIR, DART, LTIR, etc.

5.2 Project Level Tools

This section provides recommendations on project level tools to use and levers to pull when interpreting predictive results. They are just that, recommendations. There are many ways to effectively respond.

<u>Predictive Results</u>

Timely and responsive review and interpretation of the predictive results is important. The results should be shared, action plans assigned and a summary of metrics distributed to project supervision prior to the start of week safety meeting so that each of these can be reviewed with the project team. There is a danger for project teams with a low ranking getting complacent and taking their focus off actively preventing incidents despite the model providing actionable output. Executives, Safety and Operations Management should continuously encourage these teams to stay active.

Field Tools

Safety Observations - If the project team has access to a safety observation tool, they might perform additional observations during weeks of high and medium risk ratings. These observations should be focused on engaging crews to take responsibility for their safety and ensure that their work area is clean and plans are complete and up to date.

Tool Box Talks - Using the metrics and action items provided in the predictive results as well as queues from ongoing work activities, the project team can participate in additional tool box talks on topics aimed at engaging the crews and addressing trouble metrics.

Safety Training - the project team may use its own or discretionary resources to provide training to specific crews (or the entire project) on topics related to the ongoing work activities.

Safety stand down - The project team may stop all work and conduct a Safety huddle for all folks on the project to address the indicated risk.

Gamification - Promote and/or incentivize the concept of “proving the model wrong”.

While all of these efforts take time and effort to complete, it is less time and effort than would be required should an incident occur and cause the team to complete an investigation and incident paperwork.

6.0 Conclusion

It is our hope that this document achieved the goal of producing meaningful guidelines on how to incorporate predictive analytics outputs within the existing operational processes of a construction firm. By applying the concepts in this white paper, we believe that construction can join other industries in using this powerful analytical approach to reduce risk and improve decision making. Like the manufacturing sector and its implementation of predictive maintenance or the way banking and insurance changed the way they model risk, the construction industry has the potential to evolve. What that evolution looks like is limited only by the data we choose to collect.

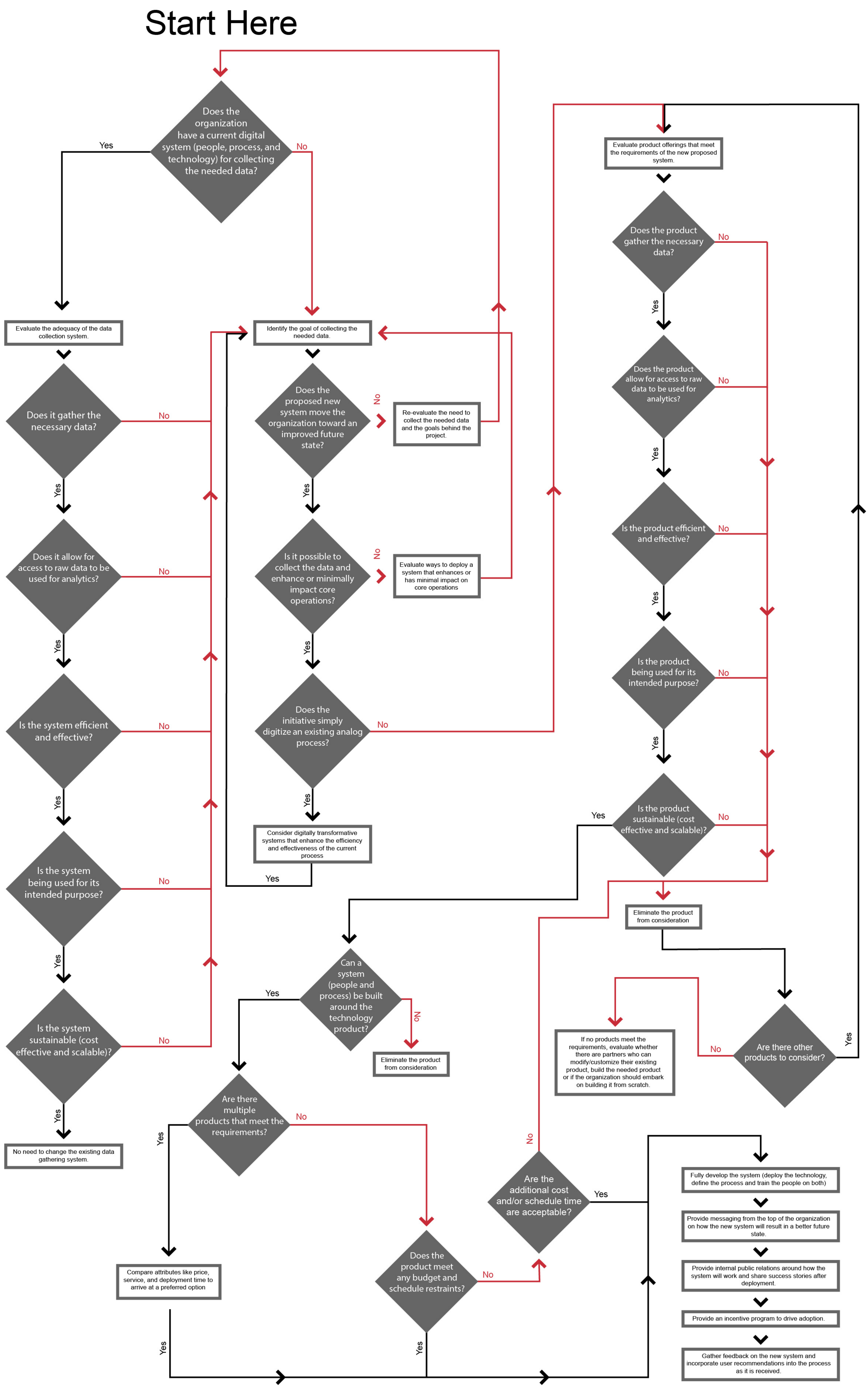
We'd love to hear what you think! To provide feedback on this whitepaper, learn more about predictive analytics or to contact the PASC, please use the form at https://pasc.ai/contact-us/

Thank you for your interest in predictive analytics!

Appendix A - Data Gathering Technology Implementation Decision Tree

(Remainder of page intentionally left blank.)

Data Gathering Technology Implementation

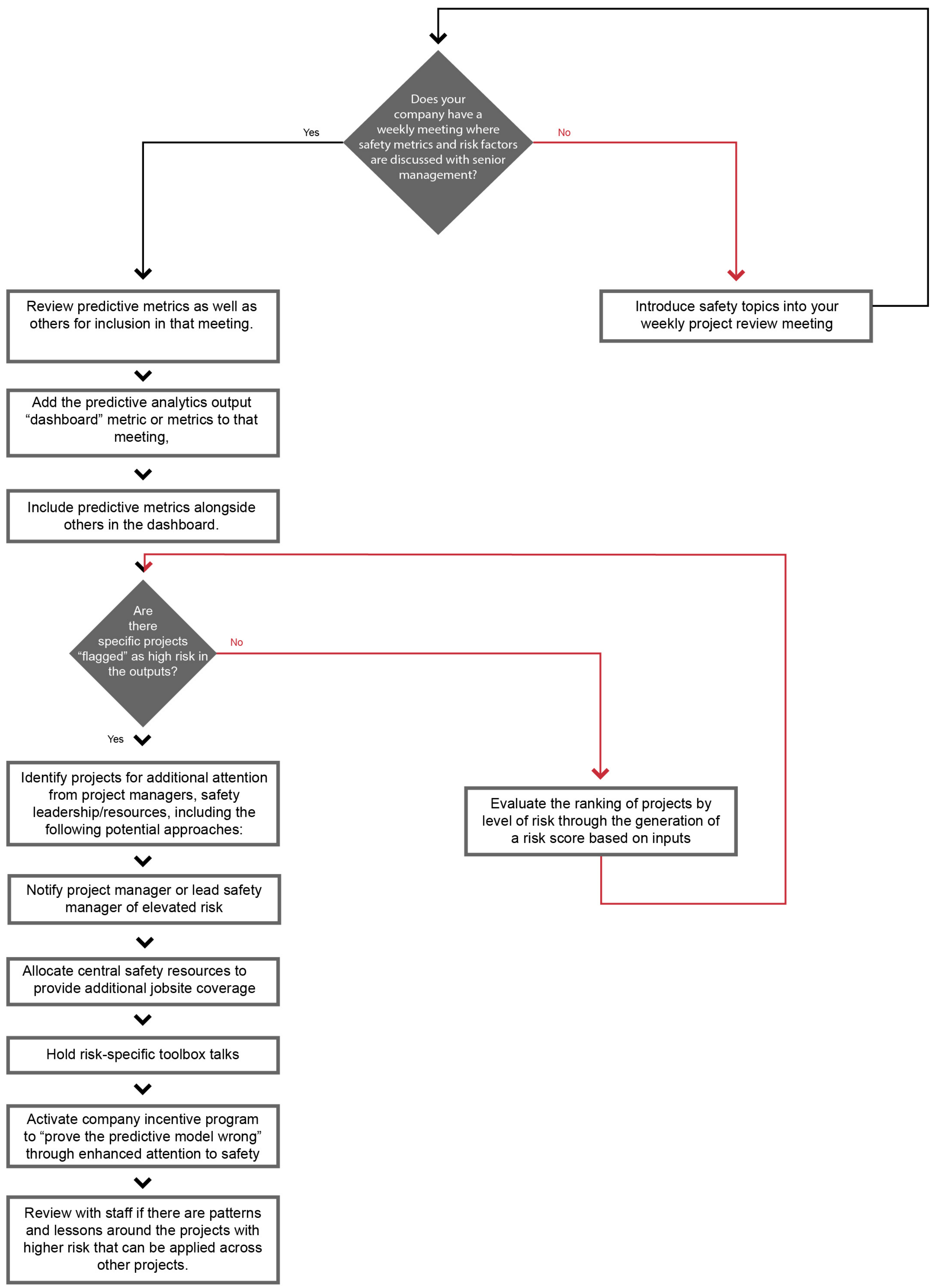


Appendix B - Implementing Predictive Analytics Into Existing Processes Decision Tree

(Remainder of page intentionally left blank.)

Implementing Predictive Analytics Into Current-Day Processes

Start Here



Appendix C - Glossary of Key Terms

Predictive Results - The output or recommendations made by descriptive, predictive or prescriptive models.

Report Values - Also called descriptive analytics, the values in reports based on data and information collected over a fixed period of time.

Structured Data - Data arranged in tables consisting of columns of identically formatted values and rows of unique records.

Unstructured Data - Data in the form of text, sound or imagery.

Predictive Analytics - Outputs from a model providing insight and answering the question “what will happen?”.

Prescriptive Analytics - Output from a model providing foresight and answering the question “how can we make it happen?”.

Descriptive Analytics - Also called report values, the values in reports based on data and information collected over a fixed period of time. Descriptive analytics provide hindsight and answer the question “what happened?”.

Digitization - Conversion of an existing, manual process to an electronic form with identical workflow. (e.g. taking a contract approval process using a folder and signature form that is passed from approver to approver and converting to an electronic contract that is transmitted to each approver electronically)

Digital Transformation - The creation of new processes and workflows using technology to eliminate a manual process or processes. (e.g. email communication allows us to communicate across distances without the need for the postal service)

Appendix D - Recommended Reading

For more on the Predictive Analytics Strategic Council, see www.pasc.ai

Examples of predictive analytics can be found in these resources:

- HBR: What the Companies That Predict the Future Do Differently [Article]
- HBR: Putting Predictive Analytics to Work [Webinar]
- Stanford University AI Index 2019 Report [PDF Download]
- Autodesk: Demystifying AI for Construction - Seven Big Ideas [PDF Download]
- McKinsey: An Executive's Guide to AI [Web Page]
- McKinsey: Various articles on AI [Articles]
- WSJ: CIO Journal [Articles]

For a survey on the current state of Construction Technology:

- JB Knowledge: 2019 Construction Technology Survey & Report

Appendix E - Acknowledgements & Legal Disclaimer

E.1 - Acknowledgements

Thank you to the membership of the Predictive Analytics Strategic Council for your help in preparing, reviewing and finalizing this white paper.

E.2 - Legal Disclaimer

This white paper was created as a voluntary collaboration amongst representatives of the PASC membership. It is presented to the industry and the general public for use without restriction, except that any copying or distribution of its contents is required to include a standard reference to the white paper so that the work can be recognized.

This document is presented for discussion purposes only. It reflects only the views of its contributors and does not represent the formal advice or opinions, or have the approval, of any of the member companies.