

FORRESTER®

The Total Economic Impact™ Of Amazon Intelligent Document Processing

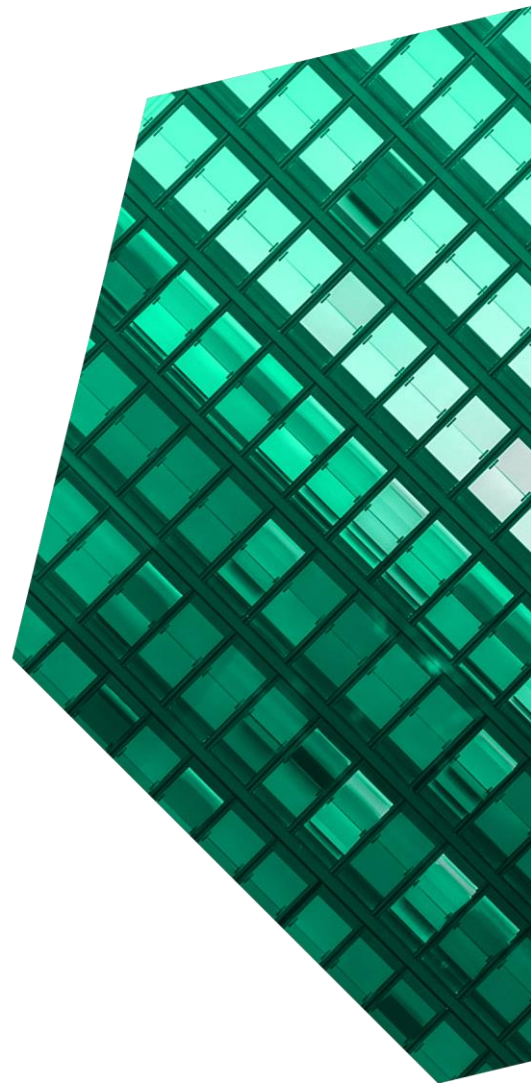
Cost Savings And Business Benefits
Enabled By Amazon Intelligent Document Processing

OCTOBER 2021

Table Of Contents

Consulting Team: Jennifer Adams
Isabel Carey

Executive Summary	1
The Amazon Intelligent Document Processing Customer Journey	6
Key Challenges.....	6
Composite Organization.....	7
Analysis Of Benefits	8
Text And Data Extraction Productivity Improvement	8
Text And Data Content Comprehension Productivity Improvement	10
Human Review Productivity Improvement	12
Machine Learning Model Development Productivity Improvement	14
Decommissioned Legacy OCR System.....	16
Unquantified Benefits.....	17
Flexibility.....	18
Analysis Of Costs	19
Amazon IDP Tools Cost	19
Amazon IDP Services Cost.....	22
Amazon IDP Implementation And Maintenance Cost.....	23
Financial Summary	25
Appendix A: Total Economic Impact	26
Appendix B: Endnotes	27



ABOUT FORRESTER CONSULTING

Forrester Consulting provides independent and objective research-based consulting to help leaders succeed in their organizations. For more information, visit forrester.com/consulting.

© Forrester Research, Inc. All rights reserved. Unauthorized reproduction is strictly prohibited. Information is based on the best available resources. Opinions reflect judgment at the time and are subject to change. Forrester®, Technographics®, Forrester Wave, RoleView, TechRadar, and Total Economic Impact are trademarks of Forrester Research, Inc. All other trademarks are the property of their respective companies.

Executive Summary

Amazon's Intelligent Document Processing (IDP) solution leverages machine learning to automate data processing and extract key structured and unstructured data from various document formats. Companies can harness Amazon's IDP solutions to improve document processing efficiency, save costs, gain new insights from their data, and bring new products and services to market faster.

With Amazon IDP, organizations can extract text from millions of documents, understand the sentiment between these documents, and manually validate machine learning (ML) results for higher accuracy and compliance. Amazon's IDP solution offers both pretrained IDP ML models and services to create custom IDP ML models.

The interviewees' companies used three Amazon IDP pretrained models: Amazon Textract, Amazon Comprehend, and Amazon Augmented AI (A2I). Amazon Textract is a ML service that automatically extracts text, handwriting, and data from scanned documents, going beyond simple optical character recognition (OCR) to identify, understand, and extract data from forms and tables. Amazon Comprehend is a natural-language processing (NLP) service that uses ML to uncover information in unstructured data. Amazon A2I is a ML service that makes it easy to build the workflows required for human review.

Machine learning model developers at the interviewees' companies used Amazon SageMaker to develop proprietary ML models.

Amazon commissioned Forrester Consulting to conduct a Total Economic Impact™ (TEI) study and examine the potential return on investment (ROI) enterprises may realize by deploying the [Amazon Intelligent Document Processing](#) solution.¹ The purpose of this study is to provide readers with a framework to evaluate the potential financial impact of Amazon IDP on their organizations.

KEY STATISTICS



Return on investment (ROI)
73%



Net present value (NPV)
\$4.60M

To better understand the benefits, costs, and risks associated with this investment, Forrester interviewed four decision-makers with experience using Amazon IDP. For the purposes of this study, Forrester aggregated the interviewees' experiences and combined the results into a single [composite organization](#).

Prior to using Amazon IDP, these interviewees noted how their organizations typically relied on large document processing teams to manually update documents that had been digitized via a legacy OCR system. The process was time-consuming and often introduced errors. It was difficult to digitize unstructured information and review document content to classify and identify key information within the documents. ML model development was inefficient.

After the investment in Amazon IDP, the interviewees noted their organizations could more efficiently extract information from millions of documents and use the information and insights to better manage

The really big benefit is inspiration. AWS capabilities, services, and technologies inspired us to architect and come up with new solutions and products. It makes us rethink what is the art of the possible.

— Solutions architect AI/machine learning, media

their businesses and serve their customers. Key benefits included:

- **Faster and more accurate document processing.** With Amazon Textract, less manual intervention was required to process documents, which reduced errors. It was easier for the interviewees' companies to scale their document processing as needed.
- **Enhanced document classification and entity recognition processes.** Amazon Comprehend allowed the interviewees to classify documents and recognize entities more efficiently. Less manual effort was required to review and classify the documents.
- **Easier and faster human review of machine learning models.** Amazon A2I allowed interviewees' organizations to include human review more easily into the ML process. Developers could more easily check if their ML models were performing as expected.
- **Time spent more efficiently for machine learning experts.** Interviewees noted that their

organizations used Amazon's pretrained tools, such as Amazon Textract, Amazon Comprehend, Amazon A2I, to automate the more basic document processing tasks. ML model experts used Amazon SageMaker to build their own proprietary IDP ML models. Amazon SageMaker helped the machine learning model experts build models faster and more efficiently.

KEY FINDINGS

Quantified benefits. Risk-adjusted present value (PV) quantified benefits include:

- **Text and data extraction productivity improvements of more than \$7.3 million.** With Amazon Textract, the composite organization digitizes information more easily, including unstructured data from a variety of document formats, such as PDF and scanned PDF files. Less manual effort is required to digitize unstructured and unconventional data.
- **Text and data content comprehension productivity improvements of more than**

\$730,000. The composite organization uses Amazon Comprehend's NLP abilities to uncover and understand information in documents' unstructured data. Use cases include document classification and named entity recognition. Amazon Comprehend automates the process, making it faster and more cost effective and reducing required human document review.

- **Human review productivity improvements of nearly \$1 million.** The composite organization uses Amazon A2I to incorporate human review more efficiently in ML workflows.
- **Machine learning model development productivity improvements of \$1.2 million.** With Amazon SageMaker, ML model developers build ML models faster and more efficiently.
- **Decommissioned legacy OCR, saving nearly \$760,000.** The composite organization decommissions its legacy OCR system after deploying Amazon IDP tools.

Unquantified benefits. Benefits are not quantified for this study include:

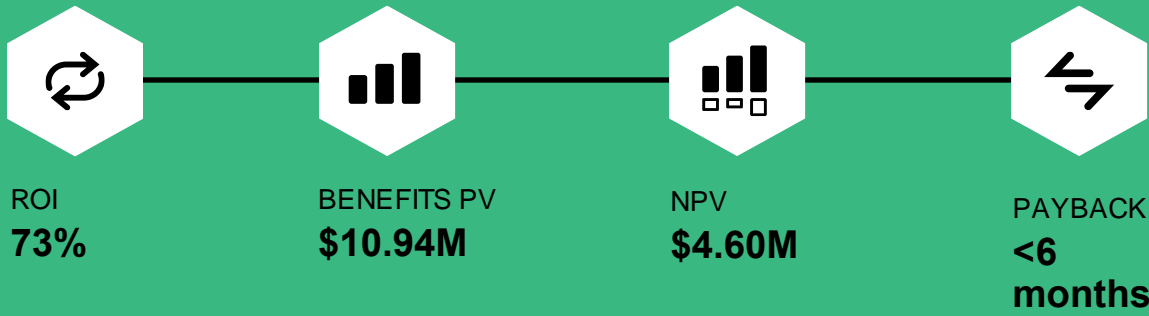
- **Enabled new products and incremental revenue.** Amazon's IDP products and services allowed several of the interviewees' companies to create new product offerings that would not have been possible previously without Amazon. This led to new incremental revenue sources
- **Faster go-to-market.** Amazon IDP helped several of the interviewees' companies go to market faster. The interviewees concentrated on their core competencies and leveraged Amazon's ML expertise to launch new products more quickly.
- **Improved data accuracy.** The interviewees found that Amazon IDP tools eliminated their organizations' manual processes and improved the accuracy of digitized data.

- **Enhanced data security.** Security was of the utmost importance to the interviewees. Their organizations needed to ensure their customers' sensitive information was safe. The overall integrity and security of the Amazon cloud infrastructure gave them confidence that any data stored or analyzed in the cloud remained secure.
- **Scalability.** Amazon's cloud-based IDP solutions allowed the interviewees' companies to quickly scale and grow their document processing as business needs required.

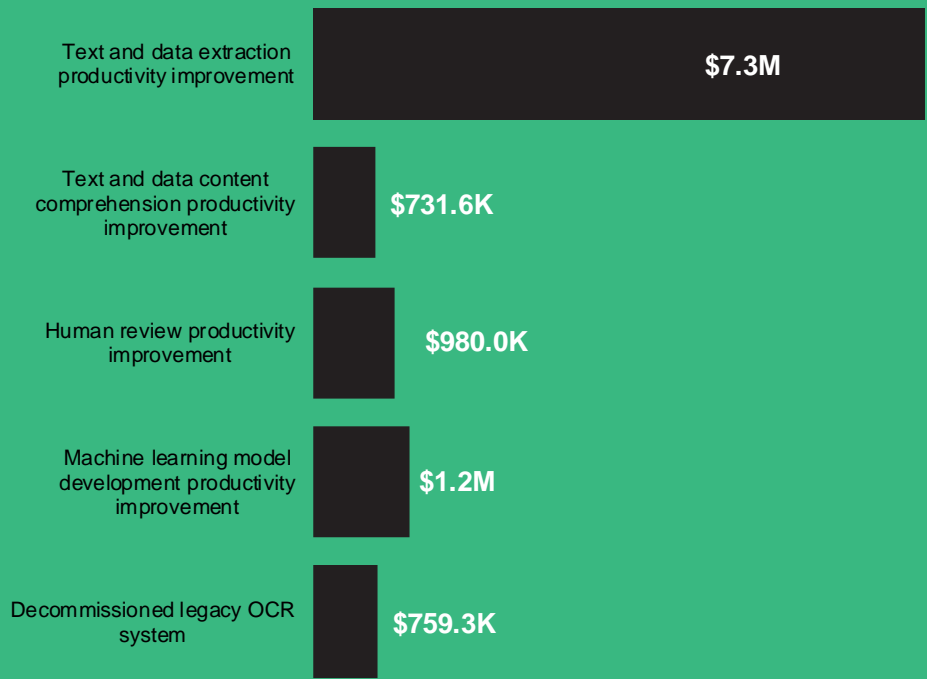
Costs. Risk-adjusted PV costs include:

- **Amazon IDP tools costs of less than \$5.0 million.** The composite organization pays Amazon fees based on the volume of documents and information processed using IDP tools, which include Amazon Textract, Amazon Comprehend, and Amazon A2I.
- **Amazon IDP services costs of just over \$266,000.** The composite organization pays Amazon a fee based on the number of hours the ML model developers use SageMaker.
- **Amazon IDP implementation and maintenance costs of just over \$1.1 million.** An internal team manages the composite organization's Amazon IDP implementation effort. A team of seven FTEs is dedicated to support the initial rollout over 15 weeks. The composite continues to phase in and expand use of Amazon IDP services over the first year. A team of 2.5 FTEs provides any required support going forward after the initial deployment effort.

The decision-maker interviews and financial analysis found that a composite organization experiences benefits of \$10.94M over three years versus costs of \$6.34M, adding up to a net present value (NPV) of \$4.60M and an ROI of 73%.



Benefits (Three-Year)



TEI FRAMEWORK AND METHODOLOGY

From the information provided in the interviews, Forrester constructed a Total Economic Impact™ framework for those organizations considering an investment in Amazon Intelligent Document Processing.

The objective of the framework is to identify the cost, benefit, flexibility, and risk factors that affect the investment decision. Forrester took a multistep approach to evaluate the impact that Amazon Intelligent Document Processing can have on an organization.

DISCLOSURES

Readers should be aware of the following:

This study is commissioned by Amazon and delivered by Forrester Consulting. It is not meant to be used as a competitive analysis.

Forrester makes no assumptions as to the potential ROI that other organizations will receive. Forrester strongly advises that readers use their own estimates within the framework provided in the study to determine the appropriateness of an investment in Amazon Intelligent Document Processing.

Amazon reviewed and provided feedback to Forrester, but Forrester maintains editorial control over the study and its findings and does not accept changes to the study that contradict Forrester's findings or obscure the meaning of the study.

Amazon provided the customer names for the interviews but did not participate in the interviews.



DUE DILIGENCE

Interviewed Amazon stakeholders and Forrester analysts to gather data relative to Amazon Intelligent Document Processing.



DECISION-MAKER INTERVIEWS

Interviewed four decision-makers at organizations using Amazon Intelligent Document Processing to obtain data with respect to costs, benefits, and risks.



COMPOSITE ORGANIZATION

Designed a composite organization based on characteristics of the interviewees' organizations.



FINANCIAL MODEL FRAMEWORK

Constructed a financial model representative of the interviews using the TEI methodology and risk-adjusted the financial model based on issues and concerns of the decision-makers.



CASE STUDY

Employed four fundamental elements of TEI in modeling the investment impact: benefits, costs, flexibility, and risks. Given the increasing sophistication of ROI analyses related to IT investments, Forrester's TEI methodology provides a complete picture of the total economic impact of purchase decisions. Please see Appendix A for additional information on the TEI methodology.

The Amazon Intelligent Document Processing Customer Journey

Drivers leading to the Intelligent Document Processing investment

Interviewed Decision-Makers

Interviewee	Industry	Region	Revenue (USD)
VP, infrastructure technology	Financial services	US	\$75 billion
Solutions architect, AI/machine learning	Media	Canada	\$6 billion
VP, cloud architecture	Financial services	US	\$6 billion
CTO	Software	Canada	\$ 2 million

KEY CHALLENGES

Before implementing the Amazon IDP solution, the interviewees' companies typically employed large document processing teams to manually update and correct documents that had been digitized via a legacy OCR system.

The interviewees noted how their organizations struggled with common challenges, including:

- **The legacy OCR system did not provide adequate accuracy and flexibility.** The companies needed to digitize and extract information from a variety of formats including PDF, Word, Excel, XML, XHTML, HTML, and JCEL files. It was difficult to extract information from PDF files and to process unstructured data, such as handwritten signatures, with the legacy OCR system.
- **The document processing process was inefficient and required significant manual intervention.** The companies employed large data processing teams to add information manually to digital files and fix errors. It was difficult for them to scale and grow document processing as the volume of documents increased. It was time-consuming for the document processing teams to classify documents and identify entities.

- **Machine learning resources were limited.** The companies' ML teams could not build all the required IDP models internally. The companies wanted off-the-shelf ML models for the more basic, commoditized document processing tasks. The companies' ML team could then focus on custom proprietary models.

“We don’t have the enormous resources that it takes to build a robust OCR like Amazon Textract. Our customers are never going to pick us because of how great our OCR is; they’re going to pick us because of how great our technology is. Amazon Textract and Amazon Comprehend are building blocks for us that would be cost prohibitive for us to create on our own.”

CTO, software

COMPOSITE ORGANIZATION

Based on the interviews, Forrester constructed a TEI framework, a composite company, and a ROI analysis that illustrates the areas financially affected. The composite organization is representative of the four decision-makers that Forrester interviewed and is used to present the aggregate financial analysis in the next section. The composite organization has the following characteristics:

Description of composite. The composite organization is a \$5 billion global company with 20,000 employees. It processes 10 million documents per month. Before deploying the Amazon IDP solution, the composite organization uses a team of 300 document processers to manually update, correct, and classify documents that are digitized via the organization's legacy OCR system. The composite organization has a team of 30 ML model developers who develop custom ML models.

Deployment characteristics. The composite organization uses both pretrained Amazon IDP ML models and Amazon services to create its own proprietary ML models.

The composite organization uses Amazon Textract as its OCR and to digitize document information, including tables and unstructured data such as signatures. The organization uses Amazon Comprehend for documents classification and entity recognition and Amazon A2I to facilitate the human review of ML model results.

The organization's machine learning model developers use Amazon SageMaker to develop custom proprietary ML models.

Key assumptions

- **\$5 billion annual revenue**
- **10 million documents processed per month**
- **300 document processors**
- **30 machine-learning developers**

Analysis Of Benefits

■ Quantified benefit data as applied to the composite

Total Benefits						
Ref.	Benefit	Year 1	Year 2	Year 3	Total	Present Value
Atr	Text and data extraction productivity improvement	\$1,800,000	\$3,600,000	\$3,600,000	\$9,000,000	\$7,316,304
Btr	Text and data content comprehension productivity improvement	\$180,000	\$360,000	\$360,000	\$900,000	\$731,630
Ctr	Human review productivity improvement	\$241,542	\$481,950	\$481,950	\$1,205,442	\$979,986
Dtr	Machine learning model development productivity improvement	\$283,500	\$567,000	\$567,000	\$1,417,500	\$1,152,318
Etr	Decommissioned legacy OCR system	\$118,750	\$356,250	\$475,000	\$950,000	\$759,251
	Total benefits (risk-adjusted)	\$2,623,792	\$5,365,200	\$5,483,950	\$13,472,942	\$10,939,489

TEXT AND DATA EXTRACTION PRODUCTIVITY IMPROVEMENT

Evidence and data. Interviewees noted that their companies needed to digitize and extract information from a variety of formats, including PDFs, Word, Excel, XML, XHTML, HTML, and JCEL files.

- PDF files were a very common document format. It was difficult for the interviewees' legacy OCR systems to extract information from the PDF files, since they are essentially a document markup language like HTML with a noisy and irregular information structure. PDF files that look identical to the eye can be expressed in fundamentally different ways in the PDF markup describing the documents.
- If a PDF document was scanned, it ended up as an image without the original markup language, making it even more difficult to extract information.
- The legacy OCR systems could not easily digitize unstructured data, such as handwritten signatures.

- Before Amazon Textract, the interviewees noted their companies needed to manually add and correct information that the legacy OCR system had not properly digitized. The data operators often made errors when manually entering the information.
- The CTO at a software organization noted that, before Amazon Textract, his company resorted to manual parsing and intensive code to extract information from PDF files. The process did not scale well. With Amazon Textract, the company easily extracted information from PDF files, including scanned documents and any handwritten information.
- The VP, infrastructure technology at a financial services organization explained: "We have a lot of unconventional and unstructured data to process and turn into structured data and we use Amazon Textract. Before, we would rely on manual operators to process the unconventional data. Textract saves time and it converts unstructured data into structured data in a readable format."

Modeling and assumptions. For the analysis Forrester assumes:

- The composite organization processes 10 million documents per month.
- 300 FTEs process and manually correct documents before Amazon Textract is deployed.
- Amazon Textract is phased in throughout Year 1. By Year 2, it is fully deployed.
- With Amazon Textract, the document processors spend 50% less time manually correcting documents.
- The average annual salary, including benefits, for a document processor is \$40,000.
- The document processors convert 75% of the hours saved time into productive time.

“Amazon Textract fit exactly what we were looking for in terms of business needs, scalability, volume, and types of templates.”

VP, cloud architecture, financial services

Risks. The Amazon Textract productivity improvement benefit will vary based on:

- The number of document processing staff manually reviewing and classifying documents.
- Average document processor annual salary including benefits.

Results. To account for these risks, Forrester adjusted this benefit downward by 20%, yielding a three-year, risk-adjusted total PV (discounted at 10%) of over \$7.3 million.

Text And Data Extraction Productivity Improvement

Ref.	Metric	Source	Year 1	Year 2	Year 3
A1	Documents processed per month	Composite	10,000,000	10,000,000	10,000,000
A2	Document processing staff for manual corrections before Amazon Textract (FTEs)	Composite	300	300	300
A3	Amazon Textract phased implementation	Interviews	50%	100%	100%
A4	Time savings with Amazon Textract	Interviews	50%	50%	50%
A5	Time savings with Amazon Textract (FTEs)	A2*A3*A4	75	150	150
A6	Document processing staff fully burdened annual salary	TEI Standard	\$40,000	\$40,000	\$40,000
A7	Amazon Textract productivity improvement	A5*A6	\$3,000,000	\$6,000,000	\$6,000,000
A8	Productivity recapture	TEI Standard	75%	75%	75%
At	Text and data extraction productivity improvement	A7*A8	\$2,250,000	\$4,500,000	\$4,500,000
	Risk adjustment	↓20%			
Atr	Text and data extraction productivity improvement (risk-adjusted)		\$1,800,000	\$3,600,000	\$3,600,000
Three-year total: \$9,000,000			Three-year present value: \$7,316,304		

TEXT AND DATA CONTENT COMPREHENSION PRODUCTIVITY IMPROVEMENT

Evidence and data. The interviewees' companies used Amazon Comprehend's NLP abilities to uncover and understand information in their documents' unstructured data.

- The interviewees' companies used Amazon Comprehend primarily to classify documents and for named entity recognition.
- Before Amazon Comprehend, document classification and entity recognition required significant human review of the documents. Amazon Comprehend automated the process, making it faster and more cost effective.
- The solutions architect, AI/machine learning at a media company used Amazon Comprehend to identify facts and entities in news items and legal content going back to the 1800s. The interviewee noted, "Amazon Comprehend definitely reduced the time we spend in terms of named entity recognition."
- The CTO at a software company used Amazon Comprehend to identify entities across legal documents and contracts. The company first uses Amazon Textract to extract core information from the documents and then uses Amazon Comprehend to review the contracts' terms and to develop inferences. The CTO shared: "The outcome for our customers would not be as accurate — or even possible — without Textract and Comprehend. We would have had to build additional models to get the context we get from Comprehend."

Modeling and assumptions. For the analysis Forrester assumes:

- Thirty document processing team members review and classify documents before the composite organization deploys Amazon Comprehend.

“Amazon Comprehend gives us certain entity recognition and document classification supplementation. Amazon Comprehend gave us a way to automate and shortcut work that we would have to do in a more labor-intensive way if we didn’t have Amazon Comprehend.”

CTO, software

- Amazon Comprehend is phased in throughout Year 1. By Year 2, it is fully deployed.
- With Amazon Comprehend, the document processors spend 50% less time manually reviewing documents.
- The average annual salary, including benefits, for a document processor is \$40,000.
- The document processors convert 75% of the hours saved into productive time

Risks. The Amazon Comprehend productivity improvement benefit will vary based on:

- The number of document processing staff manually reviewing and classifying documents.

Document processing time savings with Amazon Comprehend



50%

- Average document processor annual salary including benefits.

Results. To account for these risks, Forrester adjusted this benefit downward by 20%, yielding a three-year, risk-adjusted total PV of over \$731,000.

Text And Data Content Comprehension Productivity Improvement					
Ref.	Metric	Source	Year 1	Year 2	Year 3
B1	Document processing staff for categorizing, classifying, and searching documents before Amazon Comprehend (FTEs)	Composite	30	30	30
B2	Amazon Comprehend phased implementation	Interviews	50%	100%	100%
B3	Time savings with Amazon Comprehend	Interviews	50%	50%	50%
B4	Time savings with Amazon Comprehend (FTEs)	B1*B2*B3	7.5	15.0	15.0
B5	Document processing staff fully loaded annual salary	TEI Standard	\$40,000	\$40,000	\$40,000
B6	Amazon Comprehend productivity improvement	B4*B5	\$300,000	\$600,000	\$600,000
B7	Productivity recapture	TEI Standard	75%	75%	75%
Bt	Text and data content comprehension productivity improvement	B6*B7	\$225,000	\$450,000	\$450,000
	Risk adjustment	↓20%			
Btr	Text and data content comprehension productivity improvement (risk-adjusted)		\$180,000	\$360,000	\$360,000
Three-year total: \$900,000			Three-year present value: \$731,630		

HUMAN REVIEW PRODUCTIVITY IMPROVEMENT

Evidence and data. The interviewees' companies used Amazon A2I to incorporate human review into ML workflows.

- Before Amazon A2I, it was inefficient and time-consuming to assemble the information required for human review. The companies needed to find and identify relevant data points across millions of documents stored in disparate locations.
- The VP, infrastructure technology at a financial services company used Amazon A2I to quickly search and gather information across millions of documents and data points for human review. Before Amazon A2I, the company used an internal tool to identify all the relevant data. The process was time-consuming, and it could take up to 30 minutes to find a single data point.
- The VP, infrastructure technology at a financial services company noted: "Amazon A2I will collect the millions of data points which need to be quickly and effectively processed. With the help of Amazon A2I, you can extract critical data from the forms, whether it is unstructured or structured data, and have a human review. It can take off all the information in a single shot."

Modeling and assumptions. For the analysis Forrester assumes:

- Ten developers review documents and the related the ML model predictions. They spend 50% of their time of the review, so five FTEs are dedicated to the task.
- Amazon A2I is phased in throughout Year 1. By Year 2, it is fully deployed.
- Amazon A2I saves the developers 85% of the time they were spending on model review.
- The average ML model developer annual salary including benefits is \$189,000.

- The ML model developers convert 75% of the hours saved into productive time.



Risks. The Amazon A2I productivity benefit will vary based on:

- The number of developers reviewing model results.
- The desired level of human review.
- The average annual salary for a ML model developer.

Results. To account for these risks, Forrester adjusted this benefit downward by 20%, yielding a three-year, risk-adjusted total PV of nearly \$980,000.

“We use Amazon A2I a lot. It makes it easy to build the workflows required for human review.”

VP, infrastructure technology, financial services

Human Review Productivity Improvement					
Ref.	Metric	Source	Year 1	Year 2	Year 3
C1	Developers reviewing documents and ML model predictions before Amazon A2I	Composite	10	10	10
C2	Percentage of time on review	Interviews	50%	50%	50%
C3	Developers reviewing documents and ML model predictions before Amazon A2I (FTEs)	C1*C2	5	5	5
C4	Amazon A2I phased implementation	TEI Standard	50%	100%	100%
C5	Time savings with Amazon A2I	Interviews	85%	85%	85%
C6	Time savings with Amazon A2I (FTEs)	C3*C4*C5	2.13	4.25	4.25
C7	Developers fully loaded annual salary	TEI Standard	\$189,000	\$189,000	\$189,000
C8	Amazon A2I productivity improvement	C6*C7	\$402,570	\$803,250	\$803,250
C9	Productivity recapture	TEI Standard	75%	75%	75%
Ct	Human review productivity improvement	C8*C9	\$301,928	\$602,438	\$602,438
	Risk adjustment	↓20%			
Ctr	Human review productivity improvement (risk-adjusted)		\$241,542	\$481,950	\$481,950
Three-year total: \$1,205,442			Three-year present value: \$979,986		

**MACHINE LEARNING MODEL DEVELOPMENT
PRODUCTIVITY IMPROVEMENT**

Evidence and data. The interviewees noted that their companies used Amazon SageMaker to build their own custom ML models.

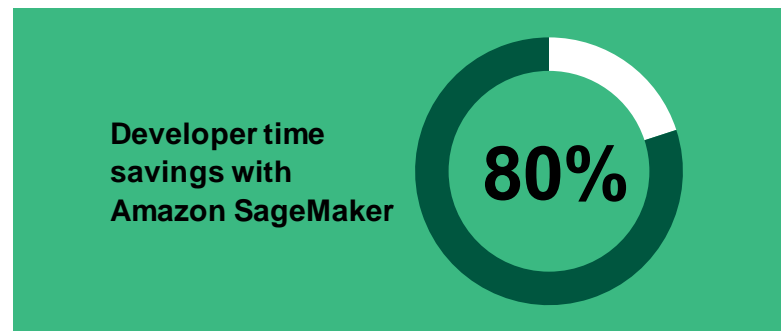
- With Amazon SageMaker, ML model developers at the interviewees’ companies built ML models faster and more efficiently.
- The solutions architect, AI/machine learning at a media company noted that ML model developers previously used on-premises infrastructure to develop and test ML models before SageMaker. It could take six to eight months for a developer to get access to the required server capacity. Now, with Amazon SageMaker, the developers have infrastructure in five to ten minutes.
- The VP, cloud architecture, at a financial services company noted that the organization used Amazon SageMaker to build custom ML models to review invoices for potential fraud and identify appropriate actions. The interviewee noted: “SageMaker is a platform that helps us improve the productivity of machine learning model developers. It gives them some out of the box features and helps them expedite development instead of building from scratch.” The VP estimated that Amazon SageMaker saved developers 20% to 25% of their time vs. building a ML model in a more traditional way.
- The CTO at a software company reported that Amazon SageMaker helped the ML developers operate more efficiently. The company avoided hiring new, additional ML developers to maintain and operate the ML tools. The CTO estimated that SageMaker saved 20% to 40% of the cost of ML model development.

“[With Amazon SageMaker] within five to 10 minutes, we can spin up new infrastructure and endlessly scalable training jobs to support whatever model works for our latest cutting-edge research.”

Solutions architect, AI/machine learning, media

Modeling and assumptions. For the analysis Forrester assumes:

- Thirty ML model developers use Amazon SageMaker. They spend 83% of their time developing new ML models. This equates to 25 FTEs building custom ML models.
- Amazon SageMaker is phased in throughout Year 1. By Year 2, it is fully deployed.
- With Amazon SageMaker, the ML model developers save 20% of the time they previously spend building ML models.
- The average ML model developer annual salary including benefits is \$189,000.



- The ML developers convert 75% of the hours saved into productive time.

Risks. The Amazon SageMaker productivity improvement benefit will vary based on:

- The number of ML model developers.
- The average annual salary for a ML model developer.

Results. To account for these risks, Forrester adjusted this benefit downward by 20%, yielding a three-year, risk-adjusted total PV of over \$1.15 million.

Machine Learning Model Development Productivity Improvement					
Ref.	Metric	Source	Year 1	Year 2	Year 3
D1	Machine learning model developers (FTEs)	Composite	30	30	30
D2	Percentage of time building machine learning models	Interviews	83%	83%	83%
D3	Machine learning model developers (FTEs)	D1*D2	25	25	25
D4	Amazon SageMaker phased implementation	Interviews	50%	100%	100%
D5	Time savings with Amazon SageMaker	Interviews	20%	20%	20%
D6	Time savings with Amazon SageMaker (FTEs)	D3*D4*D5	2.5	5.0	5.0
D7	Machine learning model developers fully loaded annual salary	TEI Standard	\$189,000	\$189,000	\$189,000
D8	Amazon SageMaker productivity improvement	D6*D7	\$472,500	\$945,000	\$945,000
D9	Productivity recapture	TEI Standard	75%	75%	75%
Dt	Machine learning model development productivity improvement	D8*D9	\$354,375	\$708,750	\$708,750
	Risk adjustment	↓20%			
Dtr	Machine learning model development productivity improvement (risk-adjusted)		\$283,500	\$567,000	\$567,000
Three-year total: \$1,417,500			Three-year present value: \$1,152,318		

DECOMMISSIONED LEGACY OCR SYSTEM

Evidence and data. Several of the interviewees; noted their organizations decommissioned their legacy OCR systems after deploying Amazon IDP.

Modeling and assumptions. For the analysis, Forrester assumes:

- The composite organization decommissions its legacy OCR system as it deploys the Amazon IDP solution.
- The composite organization begins to decommission its legacy OCR in Year 1 and the legacy system is fully decommissioned by Year 3.

Risks. The benefit of decommissioning the legacy OCR system will vary based on:

- The cost of the legacy OCR system.
- The speed of decommissioning.

Results. To account for these risks, Forrester adjusted this benefit downward by 5%, yielding a three-year, risk-adjusted total PV of more than \$759,000.

Decommissioned Legacy OCR System					
Ref.	Metric	Source	Year 1	Year 2	Year 3
E1	Legacy OCR system cost	Interviews	\$500,000	\$500,000	\$500,000
E2	Phased decommissioning	Interviews	25%	75%	100%
Et	Decommissioned legacy OCR system	E1*E2	\$125,000	\$375,000	\$500,000
	Risk adjustment	↓5%			
Etr	Decommissioned legacy OCR system (risk-adjusted)		\$118,750	\$356,250	\$475,000
Three-year total: \$950,000			Three-year present value: \$759,251		

UNQUANTIFIED BENEFITS

Additional benefits that interviewees experienced but were quantified for this study include:

- **Enabled new products and incremental revenue.** Amazon's IDP products and services allowed several of the interviewees' companies to create new product offerings that would not have been previously possible without Amazon. This led to new incremental revenue sources.

“We are making more revenue because we can create products which were impossible to create before. [Amazon] SageMaker enables us to do a lot more with our granular data sets using intelligent document processing.”

Solutions architect ,AI/machine learning, media

- **Faster go-to-market.** Amazon IDP helped several of the interviewees' companies go to market faster. The customers could concentrate of their core competencies and leverage Amazon's ML expertise to launch new products more quickly.
- **Improved data accuracy.** Because Amazon IDP tools eliminated manual processes, the interviewees' organizations saw improved accuracy of digitized data. The CTO, software shared, “The accuracy we get now is much better and more complete than it was prior to [Amazon] Textract and Comprehend.”

“[Amazon] Textract and Comprehend are boilerplate capabilities, building blocks, that would be cost prohibitive for us to create on our own. They are accelerants for us to prove all the deeper value that we need without having to get stuck around those commodity capabilities.”

CTO, software

- **Enhanced data security.** Security was of the utmost importance to the interviewees and their companies. They needed to ensure their customers' sensitive information was safe. The overall integrity and security of Amazon's cloud infrastructure gave them confidence.

“Amazon is responsible for protecting the global infrastructure that all the AWS services run on. Therefore, we do not worry about our data being leaked or used by others. You will not see any data breaches which is highly important for us.”

VP, infrastructure technology, financial services

FLEXIBILITY

The value of flexibility is unique to each customer. There are multiple scenarios in which a customer might implement Amazon IDP and later realize additional uses and business opportunities, including:

- **Scalability.** Amazon's cloud-based IDP solutions, allowed the interviewees' companies to quickly scale and grow their document processing as business needs required.

“Scalability is never a question because we’re talking about the cloud environment and one of the best advantages of using cloud is the quick scalability. We don’t have to worry about scalability or on premises infrastructure; these headaches go away because of the cloud environment.”

VP, cloud architecture, financial services

Flexibility would also be quantified when evaluated as part of a specific project (described in more detail in [Appendix A](#)).

Analysis Of Costs

■ Quantified cost data as applied to the composite

Total Costs							
Ref.	Cost	Initial	Year 1	Year 2	Year 3	Total	Present Value
Ftr	Amazon IDP tools cost	\$0	\$1,220,175	\$2,440,350	\$2,440,350	\$6,100,875	\$4,959,539
Gtr	Amazon IDP services cost	\$0	\$65,520	\$131,040	\$131,040	\$327,600	\$266,313
Htr	Amazon IDP implementation and maintenance cost	\$140,387	\$438,008	\$365,006	\$365,006	\$1,308,407	\$1,114,468
	Total costs (risk-adjusted)	\$140,387	\$1,723,703	\$2,936,396	\$2,936,396	\$7,736,882	\$6,340,320

AMAZON IDP TOOLS COST

Evidence and data. The interviewees reported that their companies paid Amazon fees based on the volume of documents and information processed. These fees covered the use of the following IDP tools: Amazon Textract, Amazon Comprehend, and Amazon A2I.

- All four interviewees reported that their companies used Amazon Textract. Two companies used Amazon Comprehend and one company used Amazon A2I.
- The interviewees' organizations typically began using the Amazon IDP solutions on small batches of documents and rolled out increased usage over time.

Modeling and assumptions. For the analysis, Forrester assumes:

- The composite organization deploys the Amazon IDP tools in Year 1 and an average of 50% of the organization's documents are processed with them. This increases to 100% of the organization's documents in Years 2 and 3.
- The composite organization processes 10 million documents per month. On average, each document has four pages. The composite

organization processes 40 million pages per month in total.

- The composite organization uses Amazon Textract Direct Document Text API to process all 40 million document pages per month. The cost for the first million pages per month is \$0.0015 per page; the cost for additional pages is \$0.0006 per page.
- Five percent of the document pages contain tables and forms. The composite organization processes these documents with Amazon Textract for pages with tables and forms. The cost for the first million pages per month is \$0.065 per page; the cost for additional pages is \$0.050 per page.
- The composite organization uses Amazon Comprehend to process 10% of its documents. The cost of Amazon Comprehend is based on 100-character units. On average each document contains 500 characters or five units. The Amazon Comprehend cost is \$0.0005 per unit.
- The composite organization uses Amazon A2I to enable human review for 3% of the documents. The cost for the first 100,000 pages per month is \$0.02 per page; the cost for additional pages is \$0.03 per page.

Risks. The cost for Amazon's IDP tools will vary based on:

- Which Amazon tools and functionality are required.
- The volume of documents.

Results. To account for these risks, Forrester adjusted this cost upward by 10%, yielding a three-year, risk-adjusted total PV (discounted at 10%) of under \$5.0 million.

Amazon IDP Tools Cost						
Ref.	Metric	Source	Initial	Year 1	Year 2	Year 3
F1	Documents processed per month	Composite (A1)		10,000,000	10,000,000	10,000,000
F2	Average pages per document	Composite		4	4	4
F3	Pages processed per month	F1*F2		40,000,000	40,000,000	40,000,000
F4	Amazon Textract pricing for Detect Document Text API (OCR) (first million, cost per page)	List price		\$0.0015	\$0.0015	\$0.0015
F5	Amazon Textract pricing for Detect Document Text API (OCR) (over 1 million, cost per page)	List price		\$0.0006	\$0.0006	\$0.0006
F6	% of pages with tables and forms	Composite		5%	5%	5%
F7	Amazon Textract pricing for pages with tables and forms (first million, cost per page)	List price		\$0.0650	\$0.0650	\$0.0650
F8	Amazon Textract pricing for pages with tables and forms (over 1 million pages, cost per page)	List price		\$0.0500	\$0.0500	\$0.0500
F9	Amazon Textract phased implementation	Phased implementation (A3)		50%	100%	100%
F10	Amazon Textract cost	((1M pages *F4)+((F3-1M pages)*F5)+(1M pages *F7)+((F3*F6)-1M pages)*F8))*12 months*F9		\$839,400	\$1,678,800	\$1,678,800
F11	% of documents processed with Amazon Comprehend	Interviews		10%	10%	10%
F12	Average units per page	10 units per page		5	5	5
F13	Units processed per month	Composite (A1)*F11*F12		20,000,000	20,000,000	20,000,000
F14	Amazon Comprehend pricing for NLP (price per unit, custom)	List price		\$0.0005	\$0.0005	\$0.0005
F15	Amazon Comprehend cost	Phased implementation (B2)		50%	100%	100%
F16	Amazon Comprehend cost	(F13-50K)*F14*F15*12 months		\$59,850	\$119,700	\$119,700
F17	% of pages with human review	Interviews		3%	3%	3%
F18	Pages with human review	(F3)*F17		1,200,000	1,200,000	1,200,000
F19	Amazon A2I cost with Textract (first 100K pages per month, cost per page)	List price		\$0.0200	\$0.0200	\$0.0200
F20	Amazon A2I cost with Textract (subsequent pages per month, cost per page)	List price		\$0.0300	\$0.0300	\$0.0300
F21	Amazon A2I phased implementation	Phased implementation (C4)		50%	100%	100%
F22	Amazon A2I cost	((100K pages *F19)+((F18- 100K pages)*F20))*12 months*F21		\$210,000	\$420,000	\$420,000
Ft	Amazon IDP tools cost (Amazon Textract, Amazon Comprehend, Amazon A2I)	F10+F16+F22	\$0	\$1,109,250	\$2,218,500	\$2,218,500
	Risk adjustment	↑10%				
Ftr	Amazon IDP tools cost (Amazon Textract, Amazon Comprehend, Amazon A2I) (risk-adjusted)		\$0	\$1,220,175	\$2,440,350	\$2,440,350
Three-year total: \$6,100,875			Three-year present value: \$4,959,539			

AMAZON IDP SERVICES COST

Evidence and data. The interviewees reported that their companies paid Amazon a fee based on the number of hours the ML model developers used Amazon SageMaker, an Amazon IDP service.

- Three of the four interviewees used Amazon SageMaker.
- The number of ML model developers using Amazon SageMaker for IDP varied based on the size of the interviewees’ company and the extent to which the company utilized custom ML models. The number of ML developers using SageMaker ranged from four to 40.
- The ML modelers at the interviewees’ companies increased their use of Amazon SageMaker over time.

Modeling and assumptions. For the analysis Forrester assumes:

- Twenty-five ML model developers use Amazon SageMaker.

- Amazon SageMaker is implemented in stages in Year 1. Each developer averages seven Amazon SageMaker jobs per week in Year 1, increasing to 14 per week in Years 2 and 3.
- Each job runs on Amazon SageMaker for 6 hours.
- Amazon SageMaker costs \$1 per hour.

Risks. The cost of Amazon SageMaker will vary based on:

- The number of ML model developers using Amazon SageMaker.
- The complexity of the ML models.
- The number of hours the developers use Amazon SageMaker.

Results. To account for these risks, Forrester adjusted this cost upward by 20%, yielding a three-year, risk-adjusted total PV of just over \$266,000.

Amazon IDP Services Cost						
Ref.	Metric	Source	Initial	Year 1	Year 2	Year 3
G1	Machine learning model developers (FTEs)	D3		25	25	25
G2	Amazon SageMaker jobs per week per developer	Interviews		7	14	14
G3	Total Amazon SageMaker jobs per week	G1*G2		175	350	350
G4	Amazon SageMaker hours per job	Interviews		6	6	6
G5	Total Amazon SageMaker hours per week	G3*G4		1,050	2,100	2,100
G6	Average cost per hour for Amazon SageMaker	List price		\$1	\$1	\$1
Gt	Amazon IDP Services cost (SageMaker)	G5*G6*52 weeks	\$0	\$54,600	\$109,200	\$109,200
	Risk adjustment	↑20%				
Gtr	Amazon IDP Services cost (SageMaker) (risk-adjusted)		\$0	\$65,520	\$131,040	\$131,040
Three-year total: \$327,600			Three-year present value: \$266,313			

AMAZON IDP IMPLEMENTATION AND MAINTENANCE COST

Evidence and data. Internal teams at the interviewees' companies supported the Amazon IDP deployment.

- The interviewees found Amazon's IDP solutions easy to install and maintain. Their organizations usually rolled out the Amazon IDP solution in phases.
- The interviewees reported that their companies typically stored the documents and data in Amazon's S3 cloud databases. Before moving to Amazon's IDP solutions, several of the companies stored information in on-premises databases. The cost of the Amazon cloud database and the prior on-premises databases were roughly similar, so there was no incremental storage cost.
- The VP, cloud architecture at a financial services organization noted, "Implementation is pretty easy because it's completely cloud-based." The financial services company began processing documents in small batches four to six weeks after signing the contract with Amazon. The company dedicated a core team of seven people to the Amazon IDP installation for the initial two months, while a broader team of 30 people also provided input. The complete global implementation took six months.
- The CTO at a software company said that the software organization spent three months experimenting and testing the Amazon IDP tools. Once the decision-maker decided to move ahead

Implementation time:

15 weeks



with the solution, it took about 10 weeks to get up and running. Three FTE engineers were dedicated to the deployment effort.

Modeling and assumptions. For the analysis Forrester assumes:

- Five FTEs work on the installation over a 15-week period.
- Once the Amazon IDP tools and services are installed, 2.5 FTEs update and maintain the solution.
- The average IT team member salary, including benefits, is \$139,050.

Risks. The Amazon IDP implementation and maintenance cost will vary based on:

- The complexity of the installation.
- The internal IT team's skill set and experience

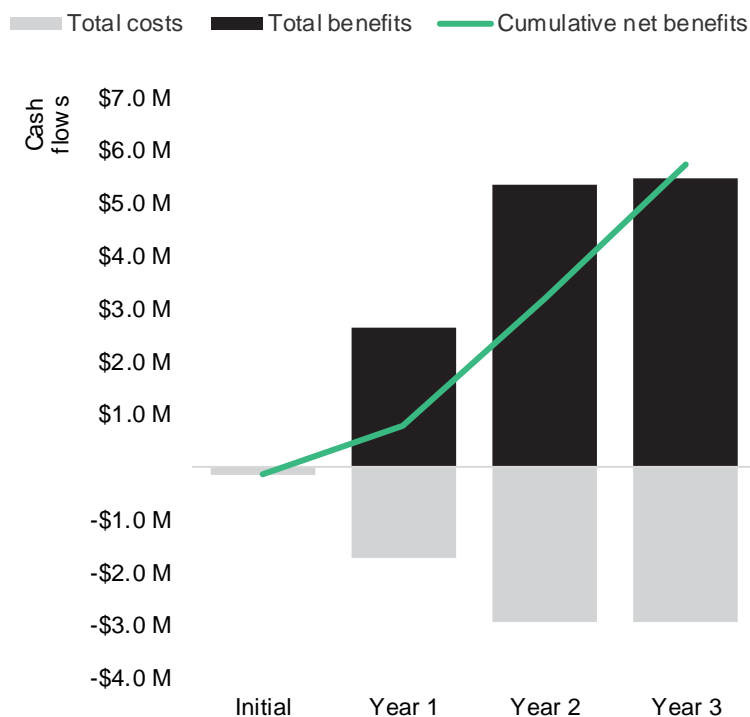
Results. To account for these risks, Forrester adjusted this cost upward by 5%, yielding a three-year, risk-adjusted total PV of just over \$1.1 million.

Amazon IDP Implementation And Maintenance Cost						
Ref.	Metric	Source	Initial	Year 1	Year 2	Year 3
H1	Implementation team (FTEs)	Interviews	5	5		
H2	Implementation time	15 weeks	19%	10%		
H3	Maintenance team (FTEs)	Interviews		2.5	2.5	2.5
H4	Implementation and maintenance team annual salary fully burdened	TEI Standard	\$139,050	\$139,050	\$139,050	\$139,050
Ht	Amazon IDP implementation and maintenance cost	$\frac{((H1 * H2) + H3) * H4}{4}$	\$133,702	\$417,150	\$347,625	\$347,625
	Risk adjustment	↑5%				
Htr	Amazon IDP implementation and maintenance cost (risk-adjusted)		\$140,387	\$438,008	\$365,006	\$365,006
Three-year total: \$1,308,407			Three-year present value: \$1,114,468			

Financial Summary

CONSOLIDATED THREE-YEAR RISK-ADJUSTED METRICS

Cash Flow Chart (Risk-Adjusted)



The financial results calculated in the Benefits and Costs sections can be used to determine the ROI, NPV, and payback period for the composite organization's investment. Forrester assumes a yearly discount rate of 10% for this analysis.

These risk-adjusted ROI, NPV, and payback period values are determined by applying risk-adjustment factors to the unadjusted results in each Benefit and Cost section.

Cash Flow Analysis (Risk-Adjusted Estimates)

	Initial	Year 1	Year 2	Year 3	Total	Present Value
Total costs	(\$140,387)	(\$1,723,703)	(\$2,936,396)	(\$2,936,396)	(\$7,736,882)	(\$6,340,320)
Total benefits	\$0	\$2,623,792	\$5,365,200	\$5,483,950	\$13,472,942	\$10,939,489
Net benefits	(\$140,387)	\$900,090	\$2,428,804	\$2,547,554	\$5,736,060	\$4,599,169
ROI						73%
Payback period (months)						<6

Appendix A: Total Economic Impact

Total Economic Impact is a methodology developed by Forrester Research that enhances a company's technology decision-making processes and assists vendors in communicating the value proposition of their products and services to clients. The TEI methodology helps companies demonstrate, justify, and realize the tangible value of IT initiatives to both senior management and other key business stakeholders.

TOTAL ECONOMIC IMPACT APPROACH

Benefits represent the value delivered to the business by the product. The TEI methodology places equal weight on the measure of benefits and the measure of costs, allowing for a full examination of the effect of the technology on the entire organization.

Costs consider all expenses necessary to deliver the proposed value, or benefits, of the product. The cost category within TEI captures incremental costs over the existing environment for ongoing costs associated with the solution.

Flexibility represents the strategic value that can be obtained for some future additional investment building on top of the initial investment already made. Having the ability to capture that benefit has a PV that can be estimated.

Risks measure the uncertainty of benefit and cost estimates given: 1) the likelihood that estimates will meet original projections and 2) the likelihood that estimates will be tracked over time. TEI risk factors are based on "triangular distribution."

The initial investment column contains costs incurred at "time 0" or at the beginning of Year 1 that are not discounted. All other cash flows are discounted using the discount rate at the end of the year. PV calculations are calculated for each total cost and benefit estimate. NPV calculations in the summary tables are the sum of the initial investment and the discounted cash flows in each year. Sums and present value calculations of the Total Benefits, Total Costs, and Cash Flow tables may not exactly add up, as some rounding may occur.



PRESENT VALUE (PV)

The present or current value of (discounted) cost and benefit estimates given at an interest rate (the discount rate). The PV of costs and benefits feed into the total NPV of cash flows.



NET PRESENT VALUE (NPV)

The present or current value of (discounted) future net cash flows given an interest rate (the discount rate). A positive project NPV normally indicates that the investment should be made, unless other projects have higher NPVs.



RETURN ON INVESTMENT (ROI)

A project's expected return in percentage terms. ROI is calculated by dividing net benefits (benefits less costs) by costs.



DISCOUNT RATE

The interest rate used in cash flow analysis to take into account the time value of money. Organizations typically use discount rates between 8% and 16%.



PAYBACK PERIOD

The breakeven point for an investment. This is the point in time at which net benefits (benefits minus costs) equal initial investment or cost.

Appendix B: Endnotes

¹ Total Economic Impact is a methodology developed by Forrester Research that enhances a company's technology decision-making processes and assists vendors in communicating the value proposition of their products and services to clients. The TEI methodology helps companies demonstrate, justify, and realize the tangible value of IT initiatives to both senior management and other key business stakeholders

FORRESTER®