# Spend Matters®

# What Can the Israeli Military and a Nobel Prize-Winning Economist Teach Us About Better Technology Decision Making?

**by Jason Busch**

## Comparing Spend Matters SolutionMap to Other Industry Analyst Ratings Methodologies

# About the Author

*Jason Busch is the founder of [Azul Partners](#) and [Spend Matters](#), and a contributor to the original SolutionMap methodology.*

# Abstract

*Based on a Nobel Prize-winning thesis, this white paper explores how traditional ratings and decision making approaches based on "expert" opinions typically are no more accurate than guesses. It explains how Nobel Prize-winning economist and behavioral scientist Daniel Kahneman inspired the SolutionMap ratings methodology. The analysis also attempts to objectively explain the differences between Spend Matters' SolutionMap analyst ratings approach compared to other industry analyst ranking models, using the Gartner Magic Quadrant (MQ), Forrester Wave and the IDC MarketScape as a basis for methodological comparison.*

# Reader Knowledge Assumption

*The analysis does not presuppose any knowledge of existing analyst ratings methodologies, and explores the topic at a generalist level of consideration (i.e., it is applicable to a general business audience).*

## Our Thesis: Legacy Ratings Models Are Based on Flawed Methodologies

The types of comparative ratings methodologies that are commonplace today among industry analysts have been proven by Nobel Prize-winning economists and behavioral scientists to lead to flawed ratings models. Yet these approaches:

- Are still used every day by procurement, finance and IT organizations as a key factor in shortlisting technology providers
- Can drive the wrong vendor selection decisions
- Can lead to reduced ROI, technology investment write-offs and added change management costs

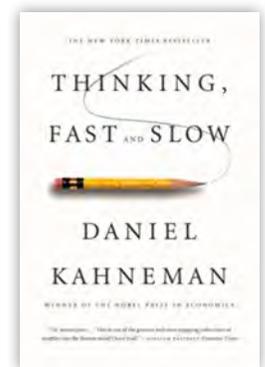How do we know these legacy analyst ratings models do not work?

What if you were told that even expert opinions and hunches were no more accurate than guesses? That they could cost lives (or dollars) in comparison to superior models for making decisions that were more predictive and accurate?

This is precisely the decision that the Israeli military made in pioneering an approach to leadership and candidate selection and placement. It is also a model we can all now make use of in everyday business decisions, such as the buying of the optimal set of technologies to support procurement, finance and other business operations.

## Fixing Poor Decision Making

In Thinking, Fast and Slow, Nobel Prize winner Daniel Kahneman suggests that even "experts" can make serious mistakes when making choices that leverage their past knowledge and experience as the foundation for judgment. In fact, he proved these decisions are no more accurate than guesses!

*"Daniel Kahneman suggests that even 'experts' can make serious mistakes...in fact, he proved these decisions are no more accurate than guesses!"*

Kahneman's laboratory for this finding was the Israeli military's candidate selection and placement process. In the past, expert evaluators would assess whether they believed a military candidate would outperform or underperform others in certain roles prior to field training and combat. They based these decisions on their experience and intuition. Yet these decisions proved no more accurate than random selections when compared with actual candidate success.

This is a pattern that continued to repeat itself, despite evidence that suggested the decision process was flawed. Here, Kahneman notes that:

*"The global evidence of our previous failure should have shaken our confidence in our judgments of the candidates, but it did not. It should also have caused us to moderate our predictions, but it did not. We knew as a general fact that our predictions were little better than random guesses, but we continued to feel and act as if each of our specific predictions was valid. I was reminded of the Müller-Lyer illusion, in which we know the lines are of equal length yet still see them as being different. I was so struck by the analogy that I coined a term for our experience: the illusion of validity."*

Kahneman's genius was in creating a new methodology, using the same exact expert ratings of candidates, to create a predictive approach that would prove accurate. In an essay on the topic of what Kahneman would end up pioneering to overcome the Illusion of Validity, Kevin McLaughlin summarizes that:

*"Before Kahneman, the military evaluated candidates based only on a single interviewer's holistic impression after a 20-minute interview … these holistic impressions had no correlation with a candidate's eventual success in combat. Kahneman, based on Paul E. Meehl's book 'Clinical Versus Statistical Prediction,' believed he could more effectively predict candidates' success by scoring them on independent, specific, objective tests."*

*"To do so he created a simple set of questions to evaluate traits relevant to combat duty like responsibility, sociability, and masculine pride. The interviewers only had to mechanically score the answers to these questions during the interviews. Kahneman then created an algorithm to weight these trait-scores to determine the final overall score for each candidate. As you might have guessed, Kahneman's algorithmic evaluation of the candidates was more correlated with success in combat than the holistic impressions."*

> *"Having interviewers rate specific, narrow criteria rather than make summary observations was more predictive (and hence, accurate) than the legacy approaches."*

In other words, having interviewers rate specific, narrow criteria rather than make summary observations was more predictive (and hence, accurate) than the legacy approaches. It was this Nobel prize-winning observation that the Spend Matters team adopted in scoring technology and solutions as part of the SolutionMap process.

## Spend Matters SolutionMap: Background and Methodological Introduction

How does SolutionMap apply Kahneman's rating methodology compared to legacy models?

Let's take the analogy of rating/ranking the front door of houses as part of an overall real estate appraisal. An "expert" appraiser using a legacy approach might look at a front door of a house and say "that's a good door" based on instinct, giving it a high score, and then move on to the interior or siding.

The SolutionMap ratings methodology, in contrast, would ask the rater to assign specific values to different elements of the door to determine an overall score. These might include:

- Type of composite or wood substance (e.g., higher scores for durable composites and hardwood vs. softwood)
- Condition of wood (based on a specific grading scale)
- Type of paint
- Number of layers of paint
- Paint sheen (with specific criteria for rating gloss / non-gloss)
- Quality of hardware (based on third-party ratings of the hardware which is in use)
- Working condition of hardware

Then, after rating all these individual aspects of the door to arrive at an overall score, this appraiser would move on to all other elements of the house itself, following a similar set of rules for rating the entire house.

This is similar to the SolutionMap ratings methodology. Under this model, the ultimate scoring of a technology module or suite reflects the ratings of a much more granular set of sub-criteria. It uses a binary ratings approach that attempts to remove bias and subjectivity that are inherent to broader, overarching questions.

Including technology and platform-level scoring criteria and functional/modular criteria, typical SolutionMaps comprise 400+ individual scored (rated) analyst fields in addition to customer-rated scoring (which we will not cover in this white paper). Procurement suites may have 1,000 or more rated fields.

Let us consider one of the scoring examples that SolutionMap considers under language support. Below is one of the questions that analysts are asked to rate each provider on, giving a rating based on a highly defined set of criteria that an expert looking at the technology can assess.
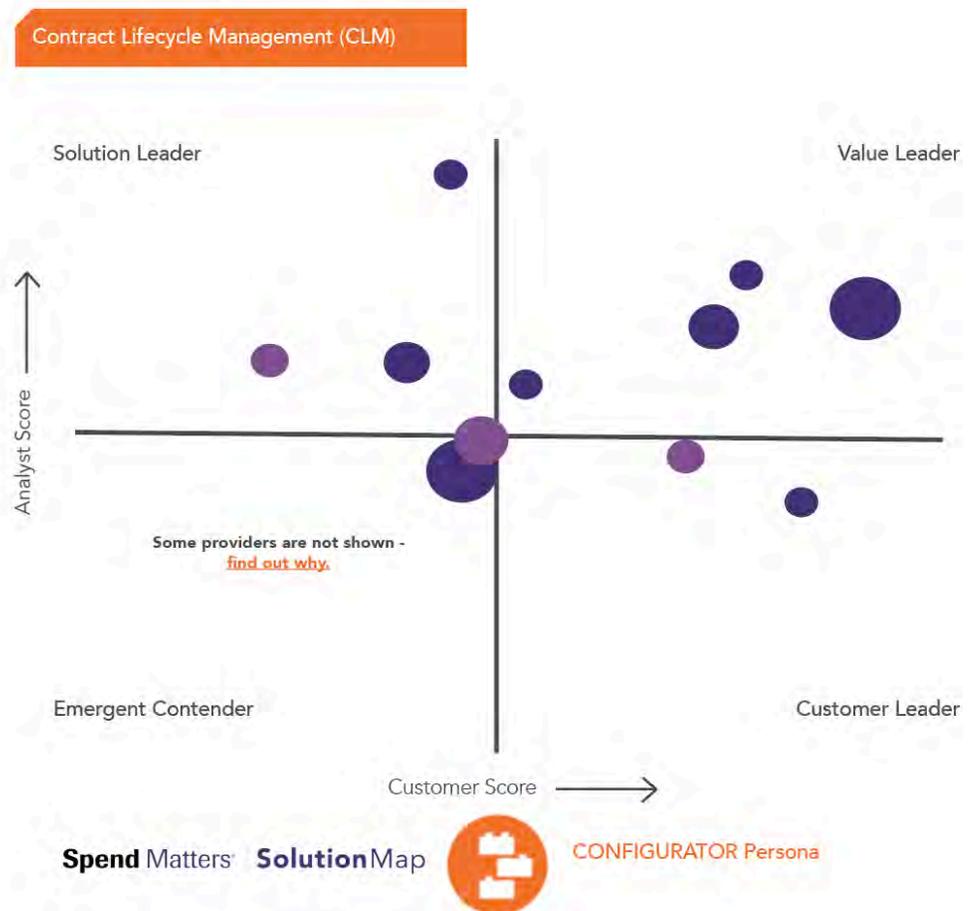
*Sample SolutionMap scoring*

| Multi-Lingual | How extensive are the multi-lingual support capabilities? | 0 - not currently supported / not applicable<br>1. flat file menu mappings for a small sest of languages;<br>2. dynamic mappings of menu options and help text based on standard translations and regional linguistic variances and support for over a dozen languages;<br>3. override features that allow a buyer to override mappings on documents / menu options being shared with a supplier and over twenty languages supported;<br>4. includes capability beyond which is previously addressed and beyond peers |
|---|---|---|

*Source: SolutionMap RFI, Q4, 2019*

To apply a rating to this question (and all others), the scorer (i.e., analyst) must validate the vendor-provided RFI scoring based on highly defined, narrow criteria gathered via interview sessions, administrative screens and demonstrations. And then, they are only providing a single score (not applying qualitative judgment).

Kahneman's model, as applied to SolutionMap ratings, takes out the subjective overarching questions/impressions and instead reduces them to a set of scores that are not prone to interpretation. After the scoring of this (and every other) individual question, the sum of these scores, as compared to the overall average across vendors and then weighted by specific need/requirement within a persona (i.e., a unique view of the data), will determine the overall rating of the analyst component of the scoring.

*Sample SolutionMap quadrant*



Contract Lifecycle Management (CLM)

Solution Leader — Value Leader

Analyst Score

Some providers are not shown - find out why.

Emergent Contender — Customer Leader

Customer Score

Spend Matters SolutionMap — CONFIGURATOR Persona

## SolutionMap in Context

The SolutionMap analyst ratings methodology follows Daniel Kahneman's recommended model to eliminate bias and create more accurate, reliable ratings.

SolutionMap analysts are never asked to judge or rate a vendor at a business level (or even an aggregate platform, product or solution level). SolutionMap does not consider any business-level criteria in creating comparative ratings (this type of subjective analysis can be found in Spend Matters qualitative research services: PRO and Nexus).

Specifically:

- Analysts rate only highly specific technical or functional criteria, against a defined scale, requiring proof to assign a specific score.
- Analyst opinion is designed to never factor into scores, ratings and the comparative view of rankings.
- Analyst scoring (described above) is separated from customer scoring, which is represented on its own axes (without consideration of any analyst input or rating).

What impact is SolutionMap having in the market today? Spend Matters published its first SolutionMap in 2017, and SolutionMap has been used by thousands of organizations to make shortlisting and buying decisions.

Since Q4 2017, Spend Matters has published 100+ SolutionMap ratings charts and reports. From a publishing and methodological basis, SolutionMap:

- Covers procurement and select finance technology only (to date).
- Does not discriminate on potential candidates for inclusion based on screening criteria such as revenue, size, number of customers, etc. (although a minimum of 3 customer references are required for consideration).
- Is focused entirely on creating an underlying dataset of rated fields that are divided into individual datasets scored by analysts and customers. At minimum, each underlying SolutionMap dataset includes 400+ analyst-rated fields and 26 customer-rated fields.

- Does not follow a "one-size-fits-all" methodology for representing data. Products and suites are represented in multiple charts based on common "buying personas" that weight underlying analyst and customer scores differently based on persona requirements (i.e, "Nimble," "Deep," "Turn-Key," "CIO-Friendly," "Configurator," etc.)
- Allows organizations selecting procurement technology to create custom views of the underlying data mapped to their own specific business requirements and technology needs through custom programs.



Nimble     Deep     Configurator     Turn-Key     CIO Friendly

## Contrasting SolutionMap With Other Analyst Methodologies: Gartner

Gartner has become ubiquitous in technology circles because of its ratings methodology, the "Magic Quadrant," or MQ, for short. It suggests that "research activities" for each MQ include:

- "Attending vendor briefings and product demonstrations"
- "Conducting surveys"
- "Interviewing reference customers identified by the vendors"
- "Interacting with industry contacts"
- "Discussing pertinent topics with clients and nonclients"
- "Consulting public sources, such as U.S. Securities and Exchange Commission filings, articles, speeches and published papers"
- "Seeking input from other Gartner analysts"
- "Assessing social media contributions, reviews and interactions, including Gartner Peer Insights"

These activities then form the basis of the two axes on which Gartner then rates providers based on the following criteria*:

- "Products/services: Core goods and services offered by the vendor that compete in and serve the market. This category includes product and service capabilities, quality, feature sets and skills (offered natively or through original equipment manufacturers), as defined in the market definition and possibly further detailed by other criteria."

- "Overall viability: Includes an assessment of the vendor's overall financial health, the financial and practical success of the relevant business unit, and the likelihood of that business unit continuing to invest in and offer the product within the vendor's product portfolio."
- "Sales execution/pricing: The vendor's capabilities in presales and sales activities and the structure that supports them in this market. This criterion also includes deal management, pricing and negotiation, presales support and the overall effectiveness of the sales channel."
- "Market responsiveness and track record: The vendor's ability to respond, change direction, be flexible, and achieve competitive success as opportunities develop, competitors act, customer needs evolve, and market dynamics change. This criterion also considers how responsive the vendor has been over time."
- "Marketing execution: The clarity, quality, creativity and efficacy of the execution of marketing programs designed to deliver the vendor's message to influence the market, promote its brand and business, increase awareness of its products and services, and establish a positive identification with the product, brand or vendor with buyers. These programs may include, among other elements, a combination of advertising, promotions, thought leadership, word of mouth and sales activities."
- "Customer experience: Relationships, products, and services, and programs that enable clients to succeed with the products being evaluated. This criterion includes the ways in which customers receive technical support or account support for the product being evaluated. It can also include ancillary tools, customer support programs (and their quality), availability of user groups and service-level agreements."
- "Operations: The vendor's ability to meet its goals and commitments. This includes the quality of the organizational structure, such as skills, experiences, programs, systems and other vehicles that enable the vendor to operate effectively and efficiently."
- "Market understanding: The ability of a vendor to understand buyers' needs and translate these needs into products and services. A vendor that shows the highest degree of vision listens to, and understands, what buyers want and need and can use that information to shape or enhance the relationship."
- "Marketing strategy: A clear, differentiated set of messages consistently communicated throughout the organization and publicized through online presence, advertising, customer programs, events and positioning statements."
- "Sales strategy: A strategy for selling products or services that uses the appropriate network of direct and indirect sales, marketing, service and communication affiliates to extend the scope and depth of a vendor's market reach, skills, expertise, technologies, services and customer base."

- "Offering (product) strategy: A vendor's approach to product development and service delivery that emphasizes differentiation, functions, methodology and feature set in relation to current and future requirements."
- "Business model: The validity and logic of a vendor's underlying business proposition in this market."
- "Vertical/industry strategy: A vendor's strategy to direct resources, skills and offerings to meet the needs of individual market segments, including vertical industries.
- "Innovation: Marshaling of resources, expertise or capital for competitive advantage, investment, consolidation or defense against acquisition."
- "Geographic strategy: A vendor's strategy to direct resources, skills and offerings to meet the needs of regions beyond its 'home' or native area — directly or through partners, channels and subsidiaries — as appropriate for that region and market."

*Source: Gartner*

Gartner MQ Summary:

- Technology and functional capability factor into Gartner scoring, but they are two of many considerations.
- The level of granularity in technology and functionality capability scoring is low compared with SolutionMap.
- While it is clear that Gartner dedicates significant rigor to its research for an MQ — including a strong peer review process — it is still asking analysts to pass macro-level judgments on providers, rather than only rating specific technical or functional capabilities that can be measured against a clearly delineated, objective scale.
- Hence, each Gartner MQ scoring is primarily comprised of business-level criteria ratings — of which basic function and technology capabilities comprise limited inputs — which require an analyst to form an opinion on a variety of topics to assign a single score.
- In addition, the Gartner MQ charting approach, based on two axes that individually consider "Completeness of Vision" and "Ability to Execute," co-mingles customer-rated criteria with analyst-rated criteria, making it impossible to separate the two considerations for those who may wish to consider either separately.

*"Hence, each Gartner MQ scoring is primarily composed of business-level criteria ratings -- of which basic function and technology capabilities comprise limited inputs."*

## Contrasting SolutionMap With Other Analyst Methodologies: Forrester

Forrester publishes a comparative analysis matrix, the "Wave," for a number of enterprise technology areas. According to Forrester, the process for criteria selection, inclusion and ratings in the Wave include:

- Identifying "differentiative rather than exhaustive" criteria for comparison.
- Considering "the roles of professionals that the analyst serves."
- Using "three [primary] inputs: a questionnaire, an executive strategy and product demo session and customer references." In addition, Forrester will "evaluate incomplete participants using Forrester estimates based on publicly available information."
- Creating "scale explanations for each criterion, defining what best-in-class looks like for each."
- Using "facts gathered during the evaluation to score each vendor against that scale and to weight criteria according to importance" and also "basing" scoring "on the analyst's experience and expertise in the marketplace."

These activities then form the basis of the two axes on which Forreter rates providers: "Current Offering" (vertical axis) and "Market Presence" (horizontal axis). "Stronger Current Offerings" score higher on the vertical access, whereas "Weaker Current Offerings" score lower. "Stronger Strategy" ratings score further to the right compared with "Weaker Strategy," which score to the left "Market Presence."

*Source: Forrester*

Forrester Wave Summary:

- Forrester uses a variety of means to research and rate vendors, including the "analyst's experience and expertise."
- Analysts are asked to score a range of business-level criteria against a "best-in-class" scale.
- Analysts are also asked to score higher-level functional and technology criteria, also against a "best-in-class" scale.
- The level of granularity in technology and functionality capability scoring is very low compared with SolutionMap.
- Opinion can factor heavily into analyst scoring (e.g., awarding SAP Ariba the lowest possible rating for the broad-category of "Technology" in a recent Wave).
- Analyst ratings on topics involving business strategy are given as much emphasis as technology/solutions.
- Customer ratings are co-mingled with analyst ratings.

## Contrasting SolutionMap With Other Analyst Methodologies: IDC

In publishing its ratings methodology for MarketScape reports, IDC suggests these analyses are "designed to provide an overview of the competitive

fitness of ICT (information and communications technology) suppliers in a given market." Specifically, "the research methodology utilizes a rigorous scoring methodology based on both qualitative and quantitative criteria that results in a single graphical illustration of each vendor's position within a given market."

How do IDC analysts score vendors? The firm suggests that, "MarketScape criteria selection, weightings, and vendor scores represent well-researched IDC *judgment* [emphasis added] about the market and specific vendors. IDC analysts tailor the range of standard characteristics by which vendors are measured through structured discussions, surveys, and interviews with market leaders, participants and end users. Market weightings are based on user interviews, buyer surveys and the input of a review board of IDC experts in each market."

Further, "IDC analysts base individual vendor scores, and ultimately vendor positions on the IDC MarketScape, on detailed surveys and interviews with the vendors, publicly available information and end-user experiences in an effort to provide an accurate and consistent assessment of each vendor's characteristics, behavior and capability."

IDC MarketScape summary:

- IDC bases its ratings approach on "research," including discussions, surveys and interviews.
- It also relies on public information and "end-user" experiences.
- IDC follows an "expert" approach that applies "judgments" to score higher level questions.
- IDC uses business-level considerations as the basis of its underlying scoring.
- The firm does not use a comparatively detailed technology/solution RFI as the basis to assess providers.
- IDC does not require detailed demonstrations to prove specific capabilities as part of its methodology.

## Contrasting SolutionMap With Other Analyst Methodologies

In contrast to other industry ratings approaches, Spend Matters SolutionMap attempts to remove the potential for analyst bias based on Kahneman's methodology as a fundamental tenet of its approach.

Other methodologies require analyst opinion to factor into scoring for more generalized, business-level questions. In fact, Gartner's, Forrester's and IDC's

methodologies explicitly state that they consider business-level considerations when rating solution providers, such as the tenability or a particular corporate, sales, marketing or product strategy or their overall impression of a vendor or technology.

Spend Matters SolutionMap, therefore, is likely to form a more accurate point of reference for understanding comparative technology approaches and demonstrated functional capability (along with transparent customer ratings) based on its depth and its approach to removing analyst scoring bias.

In contrast, Gartner, Forrester and IDC are likely to be more valuable for organizations that value an expert business judgment about a particular provider in comparison to others (even if this type of methodology, per Kahneman, can create an illusion of validity).

## The Final Word

When we apply Daniel Kahneman's research to industry analyst rating approaches, it is clear that legacy models leverage fundamentally unuseful information as the basis of comparative rating criteria. By using expert analyst judgments and hunches to justify ratings and vendor placement, these approaches create potential risk for those that rely on them directly or indirectly.

In short: Leveraging legacy analyst ratings models has created a situation that has led many organizations to make incorrect or suboptimal technology buying decisions.

*The views expressed in this paper are those of the author and do not necessarily reflect those of Azul Partners or Spend Matters.*

> *"Spend Matters SolutionMap, therefore, is likely to form a more accurate point of reference for understanding comparative technology approaches and demonstrated functional capability."*

# Spend Matters® | SolutionMap

## About SolutionMap

Find the best-fit procurement technology for your company - fast. Use SolutionMap to:



Track market developments & disruptors

Assess provider capabilities based on your needs

Identify your best-fit provider shortlist

**Every SolutionMap:**
- Ranks a technology within the procurement and supply chain spectrum
- Comprises equal parts customer and analyst input
- Reflects different organizational needs through 'buying personas'
- Gets updated semiannually to show market developments

## Why SolutionMap?

- Each semiannual update results in a series of benchmarks by buying persona as a first step in tech selection

- SolutionMap RFIs "go deep" to flesh out the differences between providers

- Then we can "map" these differences to specific customer requirements

- Resulting in a tailored technology benchmark identifying which providers best meet a customer's needs!

**spendmatters.com/SolutionMap**