peek IQ | 2020 White Paper |

REVEALING IMMEDIATE AND ACTIONABLE INSIGHTS: THE PEEK IQ WAY

What to do when your useful sources of data are unstructured, scattered, in multiple formats and disconnected, and when you need answers now, without the resource and time implications of a traditional data project.

Publication Date: August 2020

Founder's Overview

———

At PeekIQ we are on a passionate journey to equip you and your organisation with the ability to interrogate all sources of available data quickly and easily, unstructured or otherwise. To rapidly look for data that supports or negates ideas; to search for common trends and hidden narratives; to deliver insights that fuel action and rapid decision making in near real-time.

Our approach is non-traditional and has been used across insurance, banking, retail, health, law enforcement and many other sectors.

In October 2019 we decided to invest in the next generation of our solution set. One danger of the investment that we have made in areas such as automation and ease of use is that PeekIQ may at first appear simplistic compared to other approaches. Frankly, it is! We have hidden underlying complexity to ensure customers do not need specialist technical skills and can, with very little exposure, start to immediately produce meaningful results. We have made our platform open and extensible therefore PeekIQ can easily integrate into your existing process flows and tools.

Thank you for choosing to read this white paper. I invite you to put the power of PeekIQ to the test and to start a conversation with PeekIQ at any time.

Table of Contents

Abstract

According to recent research we create over 2.5 quintillion bytes of data every day. Much of it is unstructured, perhaps in the form of emails, or in social media feeds, or in a multitude of other diverse formats and platforms. All of it has potential to deliver critical and actionable insight that accelerates accurate decision making and actions.

It is also estimated that 90% of all data in existence today has been created in the last two years and that data production is increasing at over 4000% annually. This represents an unparalleled, and largely untapped source for intelligent decision making and action.

To understand where the value in your data lies, or what customer narratives are hidden behind the curtain of multiple interactions and channels, or how you can use your data in all its forms to quickly test hypotheses requires a break from tradition. Traditional approaches suggest creating consolidated and normalised data models, engaging in large data cleansing exercises and applying sophisticated and specialist intelligence tools and advanced decision-making technologies such as machine learning. These approaches are valid and valuable in the correct context. The downside however, is that they tend to require the support of large, specialist project teams, large budgets and long timescales. What do you do if you require critical, time-sensitive answers and insights now? How can you arm your decision making with the ability to access and interrogate all sources of data, both structured and unstructured? How can you interrogate it? How can you test hypothesis, immediately gain insight and take better informed decisions and actions?

In this paper, using use cases, you will learn how PeekIQ has addressed these requirements; how the PeekIQ solution can be deployed across multiple industry scenarios and how you can access your data sources and gain insights without data movement or manipulation, without large project teams, without specialist skills, without large investments and in short timescales (sometimes hours, frequently days). You will also learn how this can all be automated and how it will inform the scope of large, traditional data analysis projects.

The problem with existing approaches

The approach of the current traditional data landscape tends to put the cart before the horse so to speak. Teams of data scientists come in and begin the tedious process of normalization, data cleansing, data mapping, etc… to get the data to a state where it is all consolidated (aka replicated) and in a form that can be queried using traditional languages like sql and R.

The problem this presents, is that all of this work is done upfront to answer what are often times basic questions. Who has product A vs product B? (where product A and B rest in different silos of data) or where have people taken a new drug and complained to their doctors about symptoms x, y, and z? Typically answering these questions requires traditional data manipulation and project processes since the data is scattered across multiple locations, formats and channels in unstructured form.

The analogy of a person who is attempting to get fit is helpful in conceptualizing the difference between a traditional approach and the PeekIQ approach. Typically, when an individual decides to get fit they may choose running and gym work to do so. They invest heavily in inexpensive running shoes, new running clothes, a new iWatch to track progress, better electronic scales, they may subscribe to a food service, hire a personal trainer, they may also join a gym. Unfortunately, three weeks into this demanding routine an underlying and painful joint condition emerges that prevents any further exercise! Wouldn't it have been helpful and less expensive to have taken a different, simpler, less expensive approach that revealed the underlying problem earlier? In a very simplistic way this analogy represents the difference between traditional projects and the PeekIQ approach which delivers fast insight.

Traditional approaches first need to identify all the possible data sources. Then a common tagging/naming/format template needs to be developed which is then enforced. Structured data sources are then modified to build a mapping between all of them and the new common data conventions. Unstructured documents such as pdf's, word documents, and emails are processed either manually or by utilizing other proprietary natural language parsers to strip out what is "believed" to be important and then force compliance with the aforementioned naming and tagging scheme. Once this work is complete the data sources are replicated and stored in a single data lake to create a unified data view. Once this massive task is completed, typically over months and occasionally years, data scientists and data base admins begin the process of communicating with the rest of the team to attempt to understand what queries to write in order to generate reports. These reports then need to be visualized and interpreted by subject matter experts to validate the results and provide actionable insights.

Often, when an organization finally reaches this last step, they realize they didn't have the data they needed all along to answer the question they wanted to ask. Other times it is the case that these values generate counts that asker a lot of the "what" questions but few of the "why" because now the data is not as it originally was and context is lost. Without context you do not know why the result is what it is.

PeekIQ starts at this last step, and allows you to simply ask questions of your companies data, where it already rests, in its original format, and allows for insights to be investigated via an intuitive UI that puts subject matter experts immediately and directly in touch with the original data. They can then query this data, view the results along with surrounding snippets of the source material, and export results from memory, without ever disturbing the original sources.

3 Use Cases that demanded a different approach

We have already discussed the limitations of the traditional pathways to deriving data insights. The following examples represent time-sensitive, business critical issues and opportunities that could not easily and quickly be addressed by traditional methods. We will explain later in this white paper how we addressed these use cases and the resultant impact.

Use case 1 : Law Enforcement/Counter Terrorism

This sector relies heavily on being able to quickly decipher millions of police reports, suspicious activity reports, evidence records, social media feeds and more, in order to make intelligence decisions. These data types are the hardest for traditional scripting approaches to handle due to the volumes, possible combinations that exist and the importance that context plays when making decisions. Policing becomes more difficult due to the substantially large range of subjects and specialist areas for fighting crime. These specialists are usually consumed in their own investigations and not often available to the rest of those policing the frontlines. Here is a couple example use cases that provide insight into to the methodology and technology behind our solution:

1) In today's society Terrorism presents itself in many forms which makes it very hard to combat. As a terrorist plot unfolds, the importance of immediate access to the vast quantities of intelligence streaming across multiple channels internally and externally is extremely hard for response teams to consume and filter. Furthermore, depending on the attack it is hard to organise and update the relevant experts on the situation in order to make the correct decisions necessary to minimise damage on existing attack while trying to discover any other pre-empted attacks. PeekIQ could provide officers, the ability to capture their knowledge through the creation of rules/queries based on their domain expertise and quickly share across the intelligence community. For example, bomb making experts in California could create rules/queries about signs of bomb making materials being purchased and then send them to the relevant task force to upload and apply them against their data instantaneously. To expand this functionality, these rules/queries can now be saved for re-use by anyone within the intelligence community.

2) Officers and response teams when receiving imminent threats from counter terrorism teams are responsible for analysing vast amounts of suspicious activity reports, which are often typed by an officer in pdf format, and try to identify if any of threats that are potentially live. PeekIQ provides the opportunity to save these threats in the form of rules/ queries into the system, and arm the officers with real time alerts the moment a report comes through containing any or all threats previously identified. True intelligence directly placed in the hands of those protecting our streets when time is of the essence.

Use case 2: Insurance

Insurance subrogation is the process of recovering funds paid by an insurer on behalf of their customer after identifying their customer was not at fault and the liability fell on another parties' insurance provider. This process requires a costly and time-consuming legal exercise of chasing funds from another Insurer or individual. This process is usually initiated by the insurer if it hits a minimum value that is "worth" pursuing without spending more on recovering than the value gained by its recovery. In order to reach that decision analysts spend extra time manually reading through all case reports to make a decision as to whether or not to proceed. The cost and challenge of this exercise, compounded by limited amount human resource availability, is both inefficient and costly. Analysts are then required to identify cases with a high probability success rate in order to deliver value to the organisation.

PeekIQ would provide these subject matter experts with the ability to teach the system what they look for manually, and use the platform to help identify all claims that qualify as subrogation claims in real time, as well as sorting them by highest to least chance of success in terms of pursuing these claims.

Case 3: Healthcare

Erm systems contain most of a patient's main relevant medical data. A lot of this data is very structured, and thus, is easy to query and retrieve. However, there is a hidden wellspring of information that goes overlooked by traditional methods. This includes the doctor and nurse notes made in free form text. These notes show things such as side effects of medications, symptoms present, anomalies noted, and many more. What is required is the ability to analyse past and present notes while corelating these results with the other structured information, providing the full picture of the situation for both researchers and medical personnel alike. This ability to provide a 360 degree view of the patient journey provides the insight needed to determine patterns along the patient journey broken down by demographics, ailments, symptoms, and anything else contained in the source material.

A better way

In many cases, an ability to search an indexed version of data, in its natural state, and see the results within the documents themselves, would solve the problem without the need for a prolonged, expensive and tedious process. In cases where a traditional process is necessary for other reasons, an initial consolidated search and reporting mechanism allows for an up-front audit to know exactly what needs mapping, what is important, and most importantly, what is missing. Furthermore, organizations tend to already have subject matter experts that know what they need to find, and many times how it looks in the organization's data, they just simply don't have the means to look for it themselves without going through IT and the above steps.

PeekIQ is a new technology implementing a methodology that allows for the creation of a virtualized unified view of all connected data sources (which are read at rest in their current environment) and allows that virtualized unified view to be searched by subject matter experts using an intuitive and easy to train UI. This allows an organization to have fast actionable insights, puts data exploration at the fingertips of the organization's existing personnel, and exports their results in a variety of formats all while never disturbing the source data and without penalty of number of queries.

This isn't to say that the more traditional methods of data manipulation do not have their place. We have simply created an offering that allows for more immediate results to be found either before, or in tandem with the more traditional approach. For many mid-tier organizations, the budget just isn't there for a large big data implementation of Hadoop, IBM, or Palentere, yet it's equally as important for them to have a view of all their organizations data to better be able to make key decisions that nurture growth and expansion within the organization and their customer base. The ability to use their data the way it already is, saves time and money on storage, consolidation efforts, and IT resources to design and implement the queries, the unified view removes the need for replication of the data and time and cost associated with said efforts, and finally, the ability to use your organization's existing personnel and subject matter expertise empowers those who already know what to find to get in quickly, find what they need, and move on with their task.

6 Important Freedoms

The PeekIQ approach liberates you from many time-consuming and costly tasks that require specialist knowledge.

1. **Freedom from Preparation**
   Time and effort is not required to prepare data for analysis

2. **Freedom from Normalisation**
   No need for common data models to be developed or for normalisation

3. **Freedom from Transformation**
   No manipulation of data is required.

4. **Freedom pre-cleaning**
   No need to cleanse data, eliminate duplicates and a multitude of other pre-analysis data cleansing tasks

5. **Freedom from Data Migration**
   Data remains in its native state in its current location. Data is not moved.

6. **Freedom from a Centralised Repository**
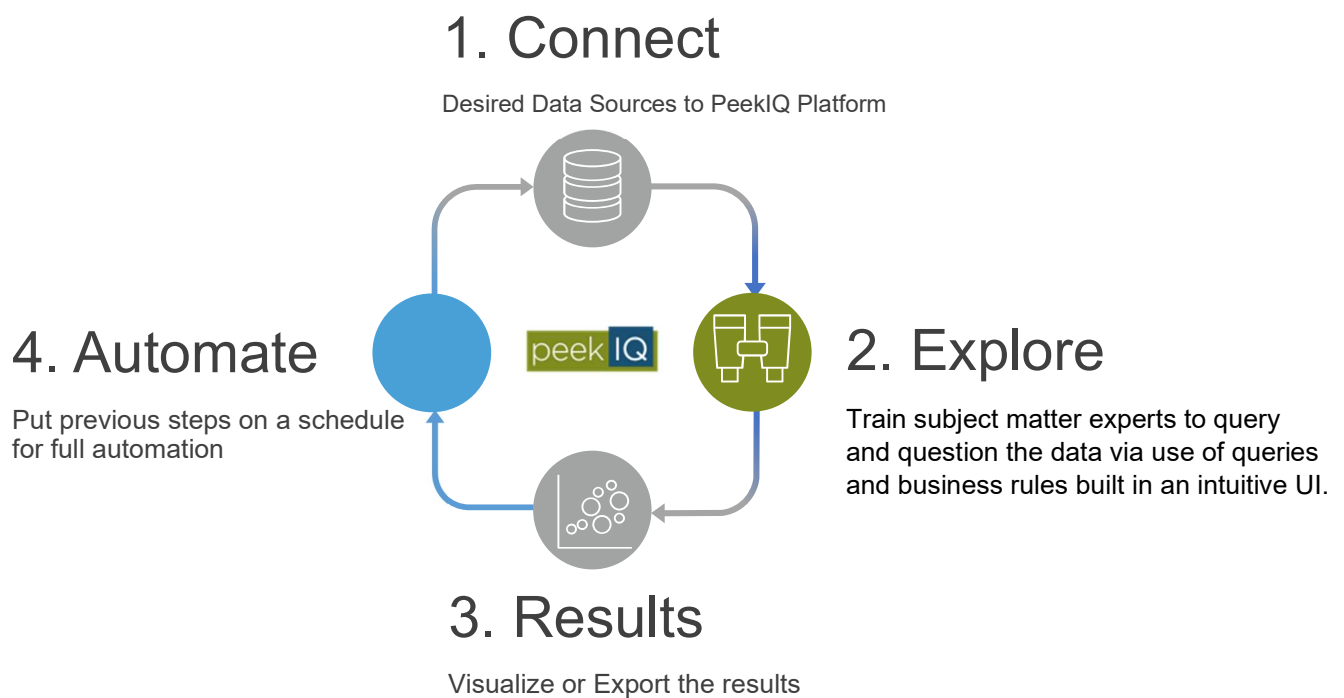   No large, complicated centralised repository of data is required.

9 Impactful Promises

1. **The Unique Promise of fast insight**
   In some instances we have used PeekIQ to deliver insights during a one-hour pitch meeting. In most instances it takes a few days

2. **The Promise of multi-file type data access**
   PeekIQ supports all major textual based file formats (such as pdf, word, txt, log, xml, json, emails, etc…) alongside all your major database technologies.

3. **The Promise of maintaining data integrity**
   PeekIQ does not alter, append, move, delete or manipulate source data in any way.

4. **The Promise of automated indexing of data**
   Peek IQ automatically creates a virtualized index of all of your data

5. **The Promise of ease of use**
   PeekIQ was designed to be used by Subject Matter experts who are not necessarily IT professionals. Our modern and intuitive GUI allows for queries to build using common UI elements and are translated automatically by the platform into the query language needed to do the heavy lifting.

6. **The Promise of automation**
   Everything done in the PeekIQ platform from data indexing, result generating, exporting, alerts and rule execution are all designed to be put on a schedule causing the platform to be more and more automated the more it is used.

7. **The Promise of re-use**
   Business rules, concepts, and queries input into the system can be exported and shared with anyone else using the PeekIQ platform, allowing for both intra and extra organizational collaboration to happen seamlessly without sharing data pools.

8. **The Promise of security**
   PeekIQ utilizes the powerful and proven Okta technology to guarantee data security.

9. **The Promise of openness and extensibility**
   PeekIQ's functionalities are web service capable making them easily extendable to fit growing customer requirements, and permit integration with existing systems.

3 Use Cases Revisited

At the end of the previous section we introduced three use cases that were difficult and costly to address using traditional pathways. By applying the approaches described in this section each of these cases were addressed with substantial major benefits. We invite you to discuss these in detail directly with PeekIQ.

How we do what we do

1. Connect

Desired Data Sources to PeekIQ Platform

4. Automate

Put previous steps on a schedule
for full automation

2. Explore

Train subject matter experts to query
and question the data via use of queries
and business rules built in an intuitive UI.

3. Results

Visualize or Export the results

Blueprint of a typical PeekIQ project

PeekIQ accomplishes results in two major steps. First, the platform is given read only access by the organization to all sources of data and information that they want to explore. These can be SQL databases, mainframes, network drives of documents (like pdf's, word documents, power points, emails, .txt files, and log files) as well as uploading from your local machine. Once these sources have been read by the system, the platform treats all resulting metadata (normally around 14% the size of the data that was read) as if it came from a single data source while simultaneously keeping track of where each piece came from initially. This virtual unified view is then able to be sliced and diced by the system through a proprietary query system that has been built to be intuitive and user friendly so that even non-technical subject matter experts can test their hypotheses against the data, see their results, edit their query, and repeat this process until they have the perfect query and the desired correlation/results from their data. These results can then be exported in a variety of formats to be used for their tasks.

All of these functionalities are implemented with security in mind. PeekIQ utilizes Octa for security.

The PeekIQ methodology and platform enables you to assess your organization's data more quickly, and through a single pane of glass.

Key Architectural Components

Component	Decision	Reason	Benefit
Development Platform	.Net Core	Easy expandability. Able to be compiled for both Windows and Linux systems.	This will allow us to be agile in feature additions for our customers and support their varied environments easily.
Back End Tech	Support for both open source and non-open source database back ends such as SQL Server and SQLite	For some customers SQL Server Licenses may be cost prohibitive. So, having an open source free option can help with that.	Provides both our customers and PeekIQ with better options based on the task at hand and better margins.
Security	Octa	Easy integration via web services. Well established in the marketplace. Well documented	An industry recognised solution that fits seamlessly with customer environments.

Three additional use case examples

Law Enforcement:

PeekIQ enables inter-departmental data to be viewed by authorised officers all at once while consolidating their search based on characteristics of the crimes they are investigating leveraging their individual domain expertise and experience. This provides the ability to identify if there have been any past instances, similarities in the crimes, or partial evidence across all connected police data providing quick insight and identification of serial crime, past offenses and imminent threats to police on the streets.

Furthermore, the added capability with the use wild cards to allow officers to use their expertise of gang language or slang to be entered into the system and treated as normal text. For example; the use of '3' for 'e' in words like weed (written as w33d on many sites known for being platforms of drug sales).

Healthcare:

Patient Journey - Hospitals tend to use a variety of EMR, paper doc, CRM, call centre information, and other sources of data all at once while caring for its' patients. Also, many of the high value pieces of insight for their patient's care come in the form of unstructured notes from caretakers explaining complaints of symptoms, reactions to medications, etc.

PeekIQ would provide the platform for these systems to be viewed at once while providing the ability to quickly search the vast amounts unstructured information and report the results. Thus providing much more insight on current patients, as well as an aggregated view and identification of trends based on past patient records.

From the business side of healthcare, PeekIQ can be used to identify what is being said during patient recruitment efforts (call centre data, patient surveys, etc) to identify ways to raise patient recruitment into the organization.

Banking:

Fraud/Anti money laundering/FSAR (Financial Suspicious Activity Reporting) (see above example for insurance fraud as it follows the same flow but ofte times using suspicious activity reports from bank tellers and other feeds of data.)

Know Your Customer - Many organisations are large enough to have multiple business units. In the case of a bank, they will have a mortgage division, credit cards, checking and savings, along with many other services. These services tend to have their own siloed data systems preventing the organization from having targeted marketing of services their customer currently doesn't utilize, or offer help and up sale services that they already do. PeekIQ's ability to unify all these siloed data sources under a single pane of glass will allow for targeted marketing and growth by upselling to their existing customer base since they can easily report on which services a single customer is or isn't using.

Conclusion and getting started

PeekIQ is ready to accelerate insights and support operations in many verticals. It's ability to virtually unify data sources without need for traditional consolidation and normalization efforts accelerates time to insights and puts the focus on generating results. We would be honoured to discuss any needs, wants, or difficulties presented in your organization and find how a PeekIQ solution can enhance your organization.

About PeekIQ

PeekIQ is a pioneering technology company creating solutions and methodologies to better equip organizations and task forces to handle their largely disparate and unstructured data to generate faster actionable insights by bypassing the many of the more cumbersome tradition data steps and empowering the local subject matter experts to test their own hypothesis against their data without penalty.

Headquartered in the USA with a growing presence in the UK, our team has implemented PeekIQ's methodologies in large scale medical institutions, Government agencies, financial and insurance institutions, government agencies and task forces, legal discovery teams, marketing firms, and many more.

Contact us:

www.peekiq.com

info@peekiq.com

Offices:

USA : Jacksonville, Boston, Chicago, Houston

UK: Edinburgh