

# Unified Resource Descriptor over KAAS Framework

## Refining Cloud Dynamics

Subhajit Bhattacharya

**Abstract** With the advent of information digitization, virtual social networking, and other means of information sharing protocols, today billions of data are available on the World Wide Web from heterogeneous sources. All these data further contribute to the emergence of Big Data gamut. When these data are processed further, we get a glimpse of information which gives some level of understanding on the subject or the matter (person, place, enterprise, etc.). Knowledge is cohesively logically processed related information with the intellect to give us multidimensional information spectrum for decision-making in real time. In today's global environment, data plays crucial role to understand the social, cultural, behavioral, and demographic attributes of a subject. Knowledge-as-a-Service (KAAS) is a pioneering cloud framework inheriting the "Internet of Things" principles that extract data from various sources in a seamless manner and can further decouple–couple logically processed information based on the "matching chromosome" algorithm. Unified Resource Descriptor (URD) is an innovative information modeling technique that operates over KAAS framework to further publish knowledge on the subject on need basis. Based on this concept, every resource or subject is assigned a unique identifier that can perform multilayered search in the KAAS Database to extract relevant knowledge frames. Considering India's context, second most populated country in the world, URD can play an indispensable role to tighten information dynamics holistically and accumulate a broader spectrum of knowledge of the resource to address adverse situations (natural calamity, medication, insurance, etc.), business process solution (Banking, BPOs, KPOs, etc.), and research practices.

**Keywords** Big data • KAAS • Cloud computing • Knowledgebase • BI

---

S. Bhattacharya (✉)  
Accenture, New Delhi, India  
e-mail: Subhajit.bhattacharya07@gmail.com

# 1 Introduction

Today, Information Technology has spread its wings wide and social sites have become the boon for social connectivity, every day the World Wide Web is getting cluttered with billions of data from heterogeneous sources. These structured, semi-structured, and unstructured data hubs form the big data gamut. Today, the biggest challenge is the utilization and proper processing of these data to derive adequate information.

Knowledge-As-A-Service is one of the pioneering initiatives to redefine cloud dynamics which enables multi-tier filtering and processing of data over “matching chromosome” algorithm to form information cuboids that are further filtered through analytical engine to get intelligently sliced, diced, and re-clustered to build information pool for a particular resource/subject. Matching chromosome is an AI-based algorithm to compare and then couple, decouple, and recouple the relevant data about the resource and thus formalize knowledge framework that further gets processed through KAAS engine to form knowledge warehouses. The ultimate idea is to bring “Information Neutrality” across the globe.

Here, the primary objective is to optimize and convert huge abandon data in the form of knowledge that can provide significant level of information for decision making and further knowledge transition.

Unified Resource Descriptor (URD) is an innovative information modeling technique that operates over KAAS framework to further publish knowledge on the subject/resource comparing behavioral, demographic, social, political, economic, and other aspects. URD ID operates as a primary key assigned to every resource/subject for which significant volume of knowledge is presented to the end user. It can be further associated as “Social Resource Planning (SRP)”.

Considering India’s context, URD can play a central role to tighten information dynamics holistically and accumulate a broader spectrum of knowledge of the resources to address adverse situations (war, natural calamity, medication, insurance, etc.), business process solutions (BFSI/FMCG/BPOs/KPOs, etc.), and education/research institutions resulting to cost efficiencies, productivity, and innovation. Most importantly, it can prove one of the significant and indispensable technologies for rural India for education and other vital facilities.

The URD ID is assigned to a subject/resource; the information about that resource will be available to the end user for knowledge and decision purpose. This URD ID works as cohesive meta-knowledge. Under KAAS framework, URD ID is explicitly associated with the resource for unified information representation.

In the KAAS framework, resources are scanned as an image or by data attributes or by videos/audios to get an in-depth insight. Therefore, when a medical firm scans an image of a patient so it can get the patient’s past medical reports saving time and cost, an insurance institution scans through person details to get his past insurance details, bank can assess the credibility of the resource or company to save itself from bad debts, defense personnel can scan suspect to see his past history, a villager

can scan the ground to understand its fertility, a common person can scan a logo or news headlines to get respective details in fraction of seconds, BPOs/KPOs can get benefits by getting the details of intended clients information in a simplified structured manner, education will be more informative and interactive. E-commerce and commercial firms can get wider information about their existing and prospective consumers and to make the right decision for sales promotions and offer positioning.

## 2 Technical Insight

### 2.1 Why KAAS and URD

Due to emergence of new technologies and social media boom, today we are observing global data warming in the huge datacenters across the globe. Global data warming is a gradual increase of unstructured and unproductive data resulting monstrous data space in the World Wide Web with no significant usage.

Cloud technology has certainly brought a number of pioneering initiatives in the IT sector, and mainly in IT-enabled services. Knowledge-As-A-Service has been introduced as another arm of cloud technology to redefine the information dynamics acting as a scavenger to segregate and unite coherent interrelated data from the global databases and form unique information clusters and further process them to generate knowledge warehouses. This will lay foundation for “Information Neutrality”.

Highly processed information so produced can be accessed by required subscription and the knowledge on the resource can be obtained as dynamically as just a glance on it. Unlike search engines (Google, Bing, or Yahoo, etc.), it will give an in-depth knowledge about a resource along with URD ID associated with it.

*The overall concept works on the below modules:*

- Intellibot Crawlers
- Matching chromosome algorithm
- Information integrator
- Test-tube information marts
- Knowledge warehouses.

The world is driven by information. Any technology, innovation, business, government policies, defense strategies, financial, agricultural, and education plans, etc., are dependent on the information that in turn form knowledge hubs to enable optimized decision-making capabilities.

Today, undoubtedly the whole world is facing challenges due to limited amount of relevant information. Until today, Europe could not come out of Euro-Crisis occurred due to bad debts years ago, most of the developed and developing nations are facing security issues, no centralized patients’ record repository, monotonous

non-interactive education; farmers are handicapped due to limited visibility and non-decision-making capabilities to judge the soil and climate conditions.

*To overcome all these constraints, KAAS framework has been introduced working on seven principles:*

- Capturing and indexing the heterogeneous data from big data clusters
- Data so collected are parsed and run through matching chromosome algorithm to get coupled/decoupled on match basis
- Information collector further collects and collaborates information iteratively to form processed information hubs
- Related information hubs are clubbed together and further segregated and coupled together to form knowledge test-tube marts
- The knowledge test-tube marts are channelized and fused together to form knowledgebase
- URD ID is assigned to every individual resource/subject to uniquely describe a resource
- This URD ID basically makes foundation of meta-knowledge.

Strategically KAAS framework formulates technology endeavor that will enable a person or an institution to have in-depth knowledge about other resources just by a glance either by keyed in the details or scanned through device camera, so explicitly the system will hit the KAAS server and fetch the details onto the screen with all relevant information. This can be well used in the process of pre-job background checking of a candidate or credibility checking of an organization.

KAAS framework iterates the information processing so many times under the information collector and test-tube marts that finally it harvests quality knowledgebases. This knowledgebase is continuously updated on real-time basis.

KAAS framework can further be tuned-up to keep continuous scanning on the global satellite maps for real-time information collaboration to combat natural calamities, crimes, and terrorism.

In Fig. 1, it is shown that in KAAS framework, data are collected from the heterogeneous sources and then went through various levels of ETL processes to get stored into various staging databases. Matching chromosome algorithm and information integrator modules are the heart of KAAS that plugs-in and plugs-out data source connections to perform various permutation/combination for generating highly processed information by coupling/decoupling the processed data.

This behavior enables the KAAS to generate the most relevant information for optimal decision-making.

In Fig. 2, it is shown that in the below KAAS framework, we can see there are six major layers. Data are extracted, processed, transformed, and loaded at every layer. At every layer, different manifestation of information is available until it gets purified at the extreme level to generate knowledge for decision-making. At every staging databases, BI tools are integrated for further segregation, purification,

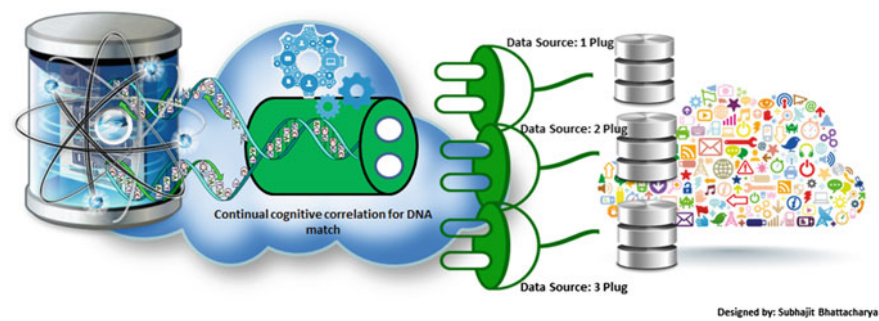


Fig. 1 Data collection mechanism in KAAS

processing, integration, and analytics. Once the knowledge information is collected into the centralized knowledgebase, URD ID is tagged with every resource/subject to provide unified resource description. All these are catered together into global KAAS datacenter to simulate Social Resource Planning for information neutrality and just-in-time decision-making capabilities.

KAAS provides the highest level of abstraction, scalability, and visualization along with security to maintain confidentiality and segregation of knowledge usage.

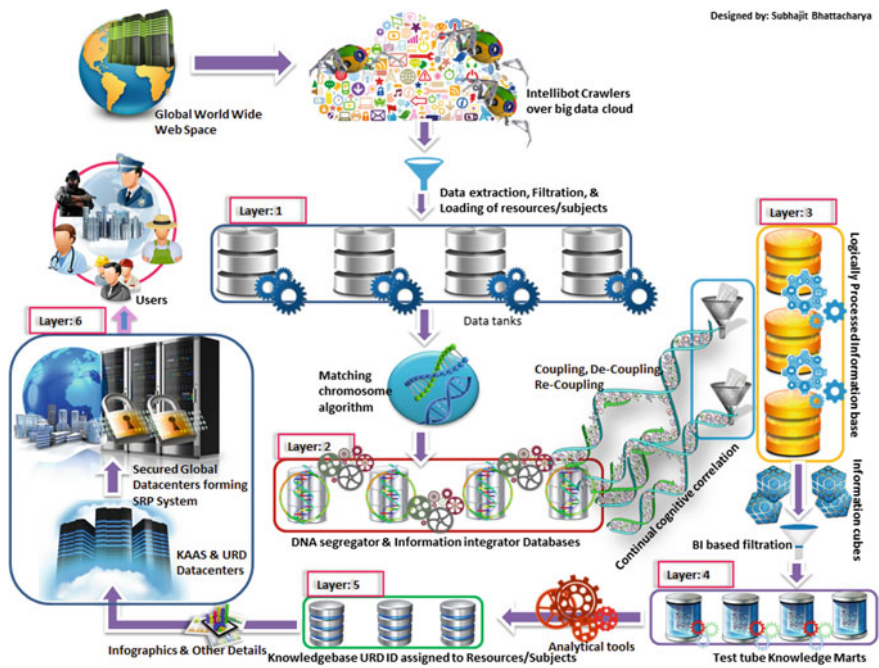


Fig. 2 Working model of KAAS framework

### **3 Case Study**

#### ***3.1 A Well-Known Medical Insurance Company Was Being Cheated by Its Customer for over a Decade—A Case Study***

In 2009, one of India's top consulting firms was in discussion with one of the well-known Indian medical insurance companies for IT solution. As of now, the insurance company was doing decent job and they made a deal with the consultancy firm to device a long-term solution to monitor insurance subscribers' annual medical claims and other background checks. Till the date, the insurance company used to take medical papers and fair background checks for the claims, however, post-solution automation it was realized that few of the subscribers were allegedly cheating the company by showing fake claims and medical reports. The customer actually never had such decreases for which he was claiming the benefits for the past several years.

It was identified not only by customer background reports and other channels of database integration but a critical assessment of past data by the application to conclude some probabilistic reports that were undergone further manual investigations. The application so devised was WHO compliant.

Until today many insurance and other financial institutions claim that their processes are too robust to be cheated, however it was found that around 30–35% of the financial or insurance institutions being cheated and vicé versa despite all possible legitimate checks.

Similar cases are happening with the corporates that perform pre-job background checking and by the time it realizes the fakeness of candidature it is too late.

The bottom line is that despite all hypothetical claims of having holistic and well-protocolled system for information analytics and tracking, till date organizations and end consumers are being cheated in various ways due to lack of relevant information bases that add to knowledge to induce decision-making capabilities. This is because many of the companies were failed miserably, either due to bankruptcy or other means, and on the other hand, consumers and loaners are becoming prey to the fraudulent companies.

### **4 Challenges and Impediments**

#### ***4.1 Key Challenges***

Although most of the organizations claim that they have opted secured and holistic approach to assess their resources and clients but unfortunately it is irrelevant fact. Also when individuals claim that they have significant knowledge about firm or

another person or a place, it is not absolutely correct because at any given instance, he will be having limited information due to limited source of data.

In Indian context, we have often seen that due to limited infrastructure and IT enablement, most of the crucial operations are still being performed manually which is in itself error prone and on top of it, there is no mechanism currently available to set up unified information system working centrally on a distributed cohesive platform providing real-time knowledgebase.

Key challenges to capture relevant and authentic information for knowledge building and decision-making:

- Improper thought process: Every innovation comes through an in-depth thought process and brainstorming. Due to lack of holistic approach, India as a potential country has failed to devise strategic knowledgebase server.
- Inappropriate development framework: Concept alone cannot play a role, but there should always be a development framework that accommodates the concept to model a working solution to address the challenges.
- Lack of infrastructure: Like any complicated long term project, this also needs a promising infrastructure and strong IT process enablement otherwise it will be just like dreaming of castle in the air.
- Lack of data collection and integration mechanisms: Although big data has fantasized IT industries for quite some time, however, due to lack of data exploration, extraction, comparison, integration, and restoration mechanisms, a robust system could never come into existence.
- Inadequate test plan: Test plans and cases to check system readiness are always advisable. Often systems failed due to vulnerabilities and risks areas that could not have been detected proactively.
- All at one go: Planning to develop and onboard the application in single instance without measuring the complexities and challenges can lead to a major mishap.
- Weak project management: There must be a well thought-out project management plan from initiation till closure keeping close eye on every phase otherwise any dodge can turn the table upside down. Lag in proper project plan and flaw in risk mitigation plan can lead the business into disastrous situations. Despite nitty-gritty checks and followed automation principles, improper plan and solution model could not yield successful result.
- Level of information access authority: It has to be made mandatory to segregate information access authority and confidentiality for the company and individuals as per the approval from government body depending case to case basis.
- An integrated collaboration channel should be set up among government, nongovernment organizations, and solution providers to address social resource planning holistically with the help of URD.

## 5 Solution Ahead

### 5.1 Methodology and Process Framework

In the above case study, we have found a number of key challenges and dependencies behind the failure of insurance company to detect the fraudulent practices. If we try to frame the above scenario under KAAS model and perform information scrutiny more holistically, it might have brought some quantifiable results.

Here, we may figure out below five major road blocks:

- (1) Lack of information integration
- (2) Lack of fraudulent check processes
- (3) No standard application in place for information binding
- (4) Lag in resource identification mechanism
- (5) No or little due diligence done on the process quality and test plans.

KAAS framework, on contrary, could have played an important role to deal with the above scenario:

- Intellibot crawlers are the artificial agents based crawlers that crawl through the World Wide Web containing heterogeneous data around and capture all the raw data to store them into data tank
- Matching chromosome algorithm further interprets data sequence and performs coupling, decoupling, and recoupling to form structured information of the resource
- Information integrator is a tool which holistically maps all the relevant information crushed through n-tier filter mechanism to build processed information base built on demographic, behavioral, social, economic, etc. parameters
- These high-end information microbes are further fused together to form test-tube knowledge marts
- Multiple test-tube knowledge marts collaborated and channeled together to form knowledge warehouse
- Although there could be various staging knowledgebases in between before being stored into the knowledge warehouse
- Every resource that has entry in the knowledgebase is assigned a URD ID
- URD plays primary role to identify the resource, based on the scanned image or information attributes and displays infographics onto the screen
- Post go-live, government, and nongovernment users can subscribe to the KAAS to get relevant information of the resource/subject. Knowledgebase keeps getting updated on a regular interval to furnish latest information
- SRP and information neutrality can help various organizations to have a 360° view on the resource/subject
- This hybrid model can further be used for processed information recycling and fix the knowledge gap.



In Fig. 3, it is shown that how KAAS framework can be used by various types of users on subscription basis in real time. Knowledge about a resource can be searched by taking or scanning an image (including live image), videos/audios, and plain search data. Knowledgebase is continuously getting refreshed; therefore likelihood of getting real-time data on just-in-time basis is too high with least latency. Knowledgebase is secured and optimally encapsulated to maintain high degree of confidentiality and at the same time maintains degree of information segregation for business and government benefits.

5.2 Quantified Benefits to Business

- SMART information processing and sharing in terms of knowledge on contrary to the traditional approach
- Information availability and neutrality can bring a major radical leap in industrial development and performance, especially in the case of India and other developing nations

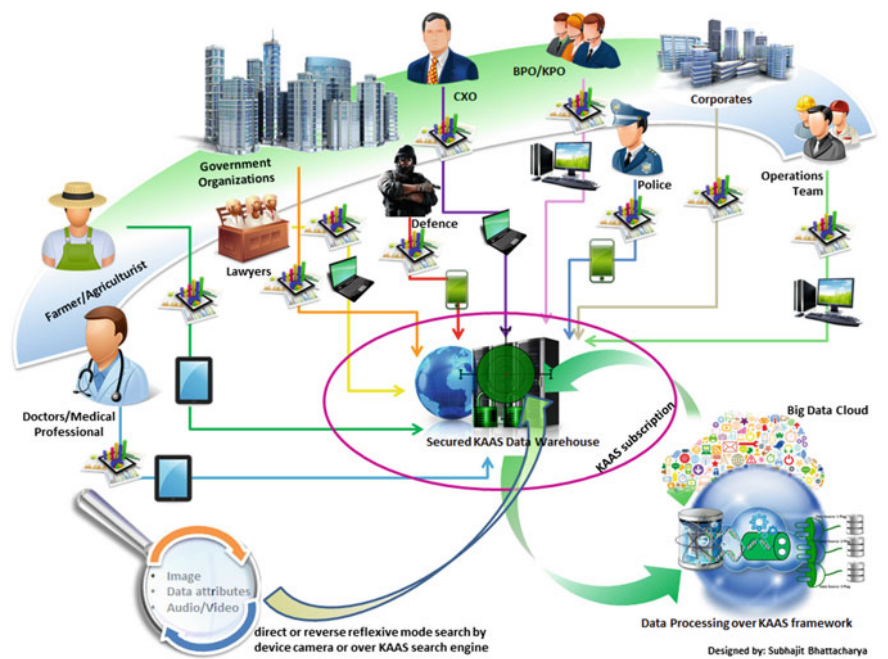


Fig. 3 Operational dynamics of KAAS framework

- Inclusion of KAAS framework can produce better result even though there is no in-house data or information warehouse or repository
- With the strategic alliance along with KAAS, companies can yield higher ROIs and meet challenging KPIs
- Authentic and the most up-to-date information will be available at all time, only a click away
- URD with the capability of direct or reverse reflexive search will change the outlook of data analysis to information analysis
- Infographics will give added advantage for graphical information presentation
- Social Resource Planning (SRP) and URD together may bring a new edge of socio-technology trend giving a new generation to the knowledge harvesting
- Inclusion of information security will make sure of confidentiality and integrity
- Organization can keep focusing on its major line of businesses while the business-related and sensitive information and other information will rest on cloud servers
- Defense, income tax, excise/custom, agriculture commodities, BFSI, hospitality, BPOs/KPOs, health care, etc., organizations will get tremendous benefits out of KAAS
- Forecasting business strategies, risks, budgets, and preparing respective plans will become easier as information is articulated in a highly structured manner topped with analytical capabilities
- As the matching chromosome algorithm not only slices and dices the processed data and coupling/decoupling information, further repositioning the original subject/resource will yield a set of related knowledge sets forming a wider spectrum of cohesive sub-knowledgebase
- As the KAAS framework has been modeled over cloud, it gives an essence of scalability assuring further scale-up of knowledgebase, enabling a platform for virtualization wherein we get a virtual interface to interact with knowledgebase managed at the cloud level.

## 6 Conclusion

Knowledge-As-A-Service is one of the pioneering initiatives to redefine cloud dynamics to process all the heterogeneous data from big data gamut and channelize them through serialized AI processes and forms a holistic knowledgebase for business growth and knowledge awareness across the globe. The idea is to bring information neutrality for all the people while maintaining security and confidentiality at all levels.

Unified Resource Descriptor has been introduced as an information modeling technique that operates over KAAS framework to further publish knowledge on the subject/resource comparing behavioral, demographic, social, political, economic, and financial aspects. URD acts as a unique key assigned to every resource/subject

for which significant volume of knowledge can be presented to the end user from the knowledgebase.

With the growth of information technology and social network, foundation for “Social Resource Planning” has been laid for collaboration and information sharing.

Objective of this framework is to basically enable government and non-government sectors to process their operations more strategic protocolled manner.

Figure 1: In KAAS framework, data are collected from the heterogeneous sources and then went through various levels of ETL processes to get stored into various staging databases. Matching chromosome algorithm and information integrator modules are the heart of KAAS that plugs-in and plugs-out data source connections to perform various permutation/combination for generating highly processed information by coupling/decoupling the processed data.

This behavior enables the KAAS to generate most relevant information for optimal decision-making.

Figure 2: In the KAAS framework, we can see there are six major layers. Data are extracted, processed, transformed, and loaded at every layer. At every layer, different manifestation of information is available until it gets purified at the extreme level to generate knowledge for decision-making. At every staging databases, BI tools are integrated for further segregation, purification, processing, integration, and analytics. Once the knowledge information is collected into the centralized knowledgebase, URD ID is tagged with every resource/subject to provide unified resource description. All these are catered together into global KAAS datacenter to simulate Social Resource Planning for information neutrality and just-in-time decision-making capabilities.

KAAS provides the highest level of abstraction, scalability, and visualization along with security to maintain confidentiality and segregation of knowledge usage.

Figure 3: In the diagram, it is shown that how KAAS framework can be used by various types of users on subscription basis in real time. Knowledge about a resource can be searched by taking or scanning an image (including live image), videos/audios, and plain search data. Knowledgebase is continuously getting refreshed; therefore likelihood of getting real-time data on just-in-time basis is too high with least latency. Knowledgebase is secured and optimally encapsulated to maintain high degree of confidentiality and at the same time maintains degree of information segregation for business and government benefits.

## References

1. Minelli, M., Chambers, M., Dhiraj, A.: Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses. O'Reilly (2012)
2. Larman, C.: Agile and iterative development: a manager's guide. Addison-Wesley Professional (2013)
3. Thomas E., Zaigham M., Ricardo P.: Cloud Computing: Concepts, Technology & Architecture. Prentice Hall (2013)

4. Patel, B., Shah, D.: Meta-Search Engine Optimization. LAP Lambert Academic Publishing (2014)
5. Morabito, V.: Big Data and Analytics: Strategic and Organizational Impacts. Springer, Berlin (2014)
6. Hubert, C.: Knowledge Management: A Guide for Your Journey to Best-Practice Processes. APQC (2013)
7. Schlessinger. Infrared Technology Fundamentals. CRC Press (1995)
8. Sabherwal, R., Becerra-Fernandez, I.: Business Intelligence. Wiley (2011)

Big Data Analytics

Proceedings of CSI 2015

Aggarwal, V.B.; Bhatnagar, V.; Mishra, D.K. (Eds.)

2018, XVIII, 766 p. 280 illus., 178 illus. in color.,

Softcover

ISBN: 978-981-10-6619-1