# CHAPTER – 2

# REVIEW OF LITERATURE

To attain the task of Review of Literature we take a systematic approach that involves our research area there is a need to identify the relevant data which gives the awareness of current topic importance. In the part of literature review a particular topic is undertaken for identifying expertise resources, preparing questioners', integrating ideas which elevate the narrowness of the research area. Literature review enlights the use of methodologies and determine new innovative approaches.

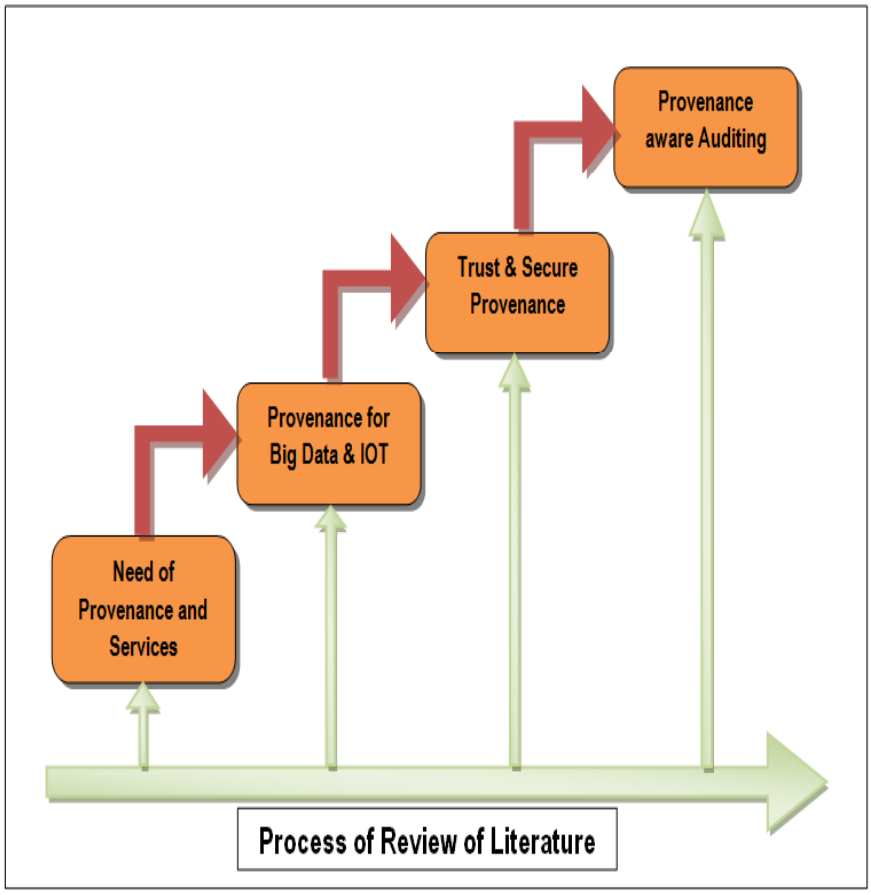


**Fig 2. Conceptual view of Process of Review of Literature**

The above diagram give an idea about our process of review of literature in a sequential manner in exhibiting of research gaps so, as to find an optimal solution. We encounter the different research articles about the need of provenance services, provenance for big data and IOT, Trust and Security in using provenance and provenance aware auditing.

## 2.1 Need of Provenance

As Provenance was used in many recent innovative technologies, the importance of provenance picturized the data lineage in the perspective of [what, why, where, when and how] a data object got modified. Tracking of provenance helps in understanding the problems and enhance the quality requirement level and reduce the cost to the proposed benchmark initiatives. Today most of the business priorities are integrated with provenance data as mandate implementations for dataflow checks and identifying risks.

### 2.1.1 Considerable facts of Provenance Services, Tracking, Flow dependencies

The adoption of provenance helps for audit trails, reliability and verification issues. Trust is the main factor which was enabled in cloud services that triggers retrieving of provenance data in a structured way. There are many frameworks which provenance aware cloud infrastructure that highlights the provocation issues of provenance. Most of the cloud paradigms fall under three categories SaaS, PaaS and IaaS, where the computing ability sharing and utilization of resources about the cloud architecture can be evaluated. Eucalyptus is an example for IaaS service where the usage of XML, SOAP protocols, REST protocol, Apache Axis 2/c framework are generally used. Enabling of provenance in the cloud environment represents many feasible factors like trust management, QoS, reliability, utilization of resources better economic cost approach and detection of faults.

There are some challenging issues in undertaking of provenance for cloud architectures are storage issue related to the granularity levels of provenance, usability of resources levels, identification of events and objects in the provenance, automatic and consistency levels in data processing, inter relation between cloud services etc. The attributes of the provenance data counter data to be processed with method name, service and flow control. [19]

The support of provenance for cloud computing depends on granularity levels in collecting of provenance data. The need of provenance addresses in the scenario of data sharing, healthcare sector, security faults and violation issues, context based search backup and audit ability. The storage system in a provenance represents data objects with different levels of abstraction which showcases the services as

provenance aware storage system. This prototype service model triggers with different functionalities like interceptor for system call information observer for tracking provenance records, LASAGNA for log management in provenance and WALDO for reading logs and records in a provenance. The above discussed functionalities are incorporated with the help of XEN hypervisor. In admitting this provenance aware storage some protocols are initialized in the transferring of data "ARE" PA-S3FS.

The PASS service deploys with protocol properties such as standalone process for cloud, cloud database for storage and messaging between cloud store and cloud database. The applicative scenario of provenance granularities exhibits applications of provenance, involvements of virtual machines in a provenance, involvement of physical machine of provenance and involvement of internet connectivity medium for a provenance. In specific provenance provocation issues raises on virtualization, reliability, interoperability, availability and scalability. The requirements to initiate the provenance deals the communications between virtual and physical machines inter phases that highlight the security with the factors of confidentiality, integrity and audit ability. The characteristics of the provenance should abide with atomicity rules, consistency, data independency, query performance etc. These PASS approaches synchronized with data PROVE approach which consists of five layers that are meant for policies, workflows, data system and regulations. This DATA-PROV approach complimented with tracking tools for events and networks. [20]

Data from different domains can be mined by using data mining principles which showcases uncertainties and complex similarities. In the part of investigating provenance data collection of data from provenance and security renders more benefit for integrity and information values. Adopting of provenance exhibits the pros and cons as two sides of a coin which helps the researchers to make decisions for their problem statements. A secure provenance schemes helps to protect consumer's confidentiality information for efficient fraud prevention. Message authentication code (MAC) may be used in data provenance by analyzing data packets in a network. The role of threat management in provenance was clearly judged with proprietorship of provider and the situation occurs in preventing of malicious entities. The categorization of provenance and its collections are differentiated with automatic

pruning and legitimacy. General provenance schemes are classified as location based, encryption based and key based. The security concerns of provenance undertake confidentiality, integrity, availability etc.[21]

Many frameworks proposed about the service of provenance and its advantages in cloud platform. There are many domains which rely on provenance data ensuring quality of data, reusability, information retrieval, benchmarking strategies and audit trails. The availability of resources and their threshold levels can be assessed with the workflow of provenance data. This provenance framework categorized business services as physical and logical services. The assumptions of this framework inputs the provenance data in an event based message at open application programming interface. This event based messages uses translators and adopters for messages and data. This framework model adopts service provenance model engine with different parameters and event sources. It also considers template model service engine which elevates the relationships of services and their thresholds. The framework incorporates with storage service provenance model which tells about the properties of services and events and time stamps of data.[22]

The increasing use of provenance data for security and trustworthiness synchronized with the elastic feature of cloud platform. The usage of cloud services can be adopted by different users in different geographical locations for accessing the resources. The proposed architecture for secure provenance in cloud platform delivers the role of cloud provenance system, host provenance system of auditor. The cloud provenance system consists of authority shell with provenance database which are coordinated by collector. The host provenance system represents provenance security and regards the provenance information and send to the common cloud storage. The enforcement of policies uses fine grained and attribute control mechanism for inter connecting to the consumer queries. The ensuring factors of integrity, confidentiality are determined by the auditor role as DRBAC (distributed role based access control) and provenance query language (PQL). [23]

Tracking and monitoring of data provenance outcomes the valuable information about the uncertainties in correlated computations. As provenance adoption is involved in business health and governance industries where

authentication of information without tampering is needed. Protecting of data provenance helps to retrieve the ownership of data, processing of DNA samples used at Shaah project, assurance of legal considerations for financial statements, digital forensic investigations etc. Provenance maintenance results trustworthiness, completeness, data integrity, availability, confidentiality and efficiency on data and its activities. Meta data of provenance was carried by the auditors for checking the integrity of provenance. the modifications of these documents are attached to the provenance chain. There are many secure provenance schemes which describes provenance chain, provenance records and identification fields. Secure provenance scheme evaluates confidentiality, threshold encryptions based on above discussed perspectives. In considering this secure provenance integrity checksum on provenance chain are implemented with some cryptographic hashes. The approaches like integrity spiral in provenance record addresses the features of spiral provenance chains with the synchronization of fine grained granularity towards confidentiality. Secure provenance model applications enlights with the aggregation of many theorems, propositions and proofs. In implementing of this secure provenance, S-PROV was introduced which significantly specify the functionalities of kernel layer, file system layer and application layer. The advantage of this S-PROV library is to add new entities for a provenance chain based on following system calls they are FREAD, FWRITE, PLOGIN, PCOPY, PDELETE, PMONITOR etc. [24]

Provenance information is an aggregated task of information flow dependencies and security concerns. Provenance exhibits history and evidence of data which can show the intensities of control policy access in developing a collaborative applications and their planning. SELLinks can be used in web management system for web based documents. The SELLinks programming language consists of labels with data and protected with restrictions and shows an independent high-level abstraction for user defined policies.

The organization of security policies with provenance data admit the style of blog which consists of blocks. The labels of SELLinks consist of groups for consumer to read and write. Policies are permitted in composite labels the security concerns of the provenance data enabled in SELLinks with PROV action labels. This label helps for non tampering and remits IP addresses for tracking and modification of

documents. CRED argument in SELLinks represents the credentials of user, apply right block used for modifying the document. The SELLinks admits the policy compositions in an easiest manner by considering the information analysis techniques. The drawback of SELLinks in tracking of documents is a burdensome programming for insertion of appropriate labels in considering appropriate data structures at provenance levels. [25]

## 2.2 Need of Provenance for Big Data and IOT

The evaluation of different latest sensors are enabled with memory and intelligent programming which are interoperable in enhancing of functionalities of resources with lower cost automation and high availability services. IOT methods need provenance data in adoption of fair IOT systems. The concept of provenance can also be used in big data as many IT enterprises transform in acquiring confidentiality and trust in business domains.

### 2.2.1 Considerable facts for the usage of Provenance towards Big Data and IOT

Big data is an emerging, innovative platform which enables data provenance for some operational tasks. The privacy and security concern of big data towards the use of provenance helps for data intensive systems. Data governance process involves distribution policy, ownership of data and their details. The annotation approaches of provenance helps for big data to attain confidentiality for the stake holders using data provenance. the usage of data provenance admits data reliability, data quality, data auditing and data validations which supports data intensive process with the use of Hadoop –Map reduce jobs.

The data provenance and its techniques elevate fine grained dependencies, workflow dependencies at heterogeneous layered architectures. CLOUDPROV architectures help in modelling and monitoring of data provenance in real time service paradigms. The provocation issues in achieving of big data provenance are usage of traditional methods, analyzing of data, multilevel abstractions of data sharing information and minimal overhead requirements. Constituting of big data provenance targets the challenges of query optimization, different data models, privacy preserving techniques, computing issues and visualization tools. [26]

The elevation of Big data technology aggregate different databases at higher computation levels to ensure reliability, reproducibility and security. MangoDB can be used for storing of data and their workflows in a consistent manner. In Big data the task of capturing provenance is a layered approach which exhibits different control mechanisms on data provenance. The layered architecture of provenance for Big data is deployed with storage, application and access layers. In the part of visualizing provenance data we should process some executable targets where maintaining of overhead, scalability, flexibility, annotations of user and visualization support. The procedural part about provenance information synchronized with administrator and user credentials. The collected provenance information stored in MongoDB and the large data is divided into small chunks stored in the form of files which can be retrieved with queries and visualized.[27]

In present day scenario IOT has vast demand towards the utilization of resources by consumer. The interconnectivity topologies of heterogeneous machines are coordinated with IOT to show case the service driven architectures. Provenance takes a major role in providing of security concerns for IOT resource utilizations. The aware of provenance discloses the ownership of data and its time stamp information at different granularity levels. The inclusion of provenance with an aggregation of IOT apprise secure provenance techniques, secure IOT device and initiation of RFID to IOT approach. The threat model elevates data provenance secure issues as congestion of data analysis, participating entity, packet dropping, legitimate note, intruder nodes and RFID tags. In the part of combining IOT and provenance the provocation issues hits more on RFID linked IOT devices, IP linked with IOT devices. The different levels of provenance collected node levels are categorized as phase based and level based provenance collections. Phase based deals with time and location parameters and Level based deals with energy consumptions. IOT with respect to WSN the showcase issues are integration approaches based on cloud, integration approaches based on proxy, integration approaches for internet communication protocols. In discussion towards IOT scenario with provenance enhance the issue of data storage, processing data, data binding, interoperability of data and fault tolerance issues. IOT for inclusion of security provenance addresses confidentiality, integrity, factors, chain integrity, privacy, availability, distribution of keys and access control. [28]

The emerging current day information trends rely on IOT. The growth in IOT usage minimizes the human interaction and human errors for enhancing the quality, automation and availableness of facilities. IOT can be defined as huge entities which are interconnected to one another. The entities can be classified as tangible and intangible which are efficient to generate large amounts of data. In discussion IOT security issues mobility is one of the issue that governs security levels with different parties. OWASP project addresses many faults in the segments of encryption, access control and role based issues of IOT. Many business domains encourage quality of data and trust worthiness which is a managing activity by the auditor to outcome the use of data reusability, reproducibility and performance issues. The completeness to showcase trust of data provenance with related to IOT was involvement of atomicity principle on actions, integrity of data cannot be modified, data confidentiality should be reserved personal data privacy etc. the triggering challenging points IOT integrated data provenance are bulk data, provenance indexing, multiple stake holders, query tools, interoperability and transformation of provenance tasks.[29]

Auditing process promotes for the new innovations of IOT to target privacy and policy factors basing on data provenance. Generally auditing flows are traced for the data which synchronized to the user requirements. The agenda of enforcing privacy and policy matters related to IOT gathering from the data provenance with the help of evidence. IOT was defined as an interoperable communications involved connecting physical or virtual devices with advanced improved services. The spectrum of IOT raises the point of ranging the customer, legal security concerns in service scenario. The protection concept in an IOT spectrum incorporated legal context in data protection laws and fair principles for information practice. According to federal trade commission of US, IOT system should possess detailed and user friendly approach. In enabling of auditing in IOT sector auditing has to contribute with some particular principles like maintaining transparency, secure policies, enabling, and consumer policies with rights expression.

The provocation inclusions for IOT states enforcing of uniform mechanisms to all the situations, constraints in deploying regulations tracing and managing approaches are not up to the maximum mark. The data flow in IOT is segmented into three planes as regulation constraint plane, mechanism enforcement plane and

auditing plane. In deploying of an audit mechanism transparency, trust and time bounds are to be achieved. Provenance graphs are generated with the help of W3C standard and broadly categorize the provenance into observed and disclosed provenance. The observed provenance deals the issues at system level where the disclosed provenance deals the issues at application level. The need of auditing declares in representing current system behaviour to the actual system behaviour. In the part of auditing maintaining confidentiality for control access, trust generating for integrity of data reach the levels of availability and scalability. Retrospective security approaches help in verification and policy monitoring at a higher level abstractions in distributed management. [30]

### 2.3 Need of Secure Provenance and Trust for Privacy Preserve

The cloud computing encompasses with heterogeneity of resources along with scalability and elasticity features addresses trust as a crucial issue in improving quality of service between different stake holders in cloud environment. The concept of Privacy preserving is a key feature in cloud computing which preferences and compensated to some extent by using provenance. Usage of cryptographic approach and encryption schemes provides confidentiality and integrity of data with the aware of provenance.

#### 2.3.1 Considerable facts for Secure Provenance and Trust for Privacy Preserve

A survey towards secure provenance addresses the importance of data, metadata and actions performed on that data. In reviewing different secure provenance schemes in cloud sampled the usage of data provenance with work flow provenance and cloud provenance. the granularity levels of data provenance exhibits virtual machines provenance, physical machines provenance and network provenance. Many secure data provenance frameworks are encouraged in deploying at service oriented architectures, operating systems trusted infrastructures, e-science grids etc. A data provenance should have the characteristics of confidentiality, compliance management and integrity management. The challenges involved for showcasing security data provenance are encryption, querying, auditing, provenance location and access control. [31]

The proposal of secure provenance with fine grained approach exhibits communication overheads and computational overheads for cryptographic operations. In the current scenario importance of digital provenance increased due to high computational segments, targets HPC and HTC levels. The proposed frameworks admit the provenance system architecture where it consists of an attribute authority that authorizes the users to access the cloud server by providing a key. Cloud users are authenticated by cloud servers and then users read write the data provenance to the cloud servers. A third party auditor audits on all the transactions on the cloud server data. The involved security models in this work enhance the parsing approach with bi linearity, non de generously and computability approach. Attribute signatures, group signatures and symmetric encryptions for cryptographic operations are achieved to the proposed extent. [32]

Trust plays a main role in data provenance towards security concerns in cloud environment. Most of the cloud infrastructures enable visualization and virtualization techniques which addresses trust and security factors. In the line of providing secure data provenance the addressing issues are confidentiality, integrity, privacy and availability. The issue of confidentiality determines sharing of sensitive information with enabled protection rules. Integrity issue represents un forgebility of data with time from any unauthorized attacks. The privacy issue represents protecting the user identity expressing trustworthiness. The availability issue considers data accessibility at any time for the demanding situations by authorized users. The scenario of proposed trust model involves with security mechanisms, authentication mechanism and key management mechanisms parameters. The sequence of process for the above discussed trust model parameters deploys in establishing of trust with secure data flows.[33]

The evaluations of policies and their management may be considered by SLA's in cloud platform with the aware of provenance. Data access policies have a huge variance in different machines is a complex task ensuring data policy security concerns. Garm tool enables the approach of policy assessment and monitoring in data provenance. This approach defines the provenance and its reference information from empirical executions. This tool makes the consumers to assign the data policies for manipulation and sharing of data. This Garm tool deploys with stream ciphers

which protects the policies without any violations. This framework offers protection of data enforcement of data policies and cross application support. This tool is loosely coupled towards buffer overruns which comprises integrity. This tool provides an independent approach to the consumer about their policies. Garm does not confidently restrict internal leaks which comprised trust worthiness. In showcasing of trust Garm tool describes completeness approach which showcases bios image, separate policy server and PCR register.

This completeness approach towards trust achievement incorporates encryption policies, decryption policies and authentication. Garm tool uses provenance and analyzes the state with memory shadow, temporary variables, files and registers. The outcome of this tool is a dynamic approach with instrumenting and computing provenance workflows. This tool represents base provenance strategy, composite provenance strategy and merging provenance strategies with cache and identity check. The experience of Garm in policy enforcements represents overheads in valgrind method, small blocks and big blocks. This Garm tool exhibits taint analysis from heterogeneous applications and minimization of flow in code considerations. [34]

Trust plays a main role in enabling of cloud services. In different service scenario the issue of trust was discussed in many framework models with the supported actors like broker's auditors and providers. Generally trust is different from security and privacy. Trust in cloud computing states the initial expected behaviours evidence with integrity and risk analyses. Trust involvement can be categorized with factors like reputation, SLA, service, audit standards. Trust mechanism basing on reputation is mostly used in e-commerce services which particularly help the performance of reliability between the stake holders. Trust based SLA validations enables comparisons of quality of service, QoS monitoring and self assessment feedbacks. Trust based on services attached with authority of service which maintains the profiles and identity of service. Trust based on audit is emphasized with different audit attestation services in ensuring trust worthiness between heterogeneous stakeholders of cloud.

Trust based on policies depends upon public key certifications and validation processes which are having specific functionalities that exhibit at that situation. Trust

based on evidence involves with unification of attributes with semantics of belief which helps in system performance assessment. In the session of trust establishment attribute assessment plays a vital role in considering the facts of consumer's opinions and observations. Trust establishment mainly rely on reputation strategy where the assessment test can be carried for provider, auditor and broker. Trust establishing in an integrated view aggregates decisions of auditor, broker and provider. The decision of auditor for enabling trust considers policies and accreditations. The policies accreditations, self assessments are the common considering points in the decision of broker and services. [35]

Trust computing information security and privacy preserving are the key factors for enabling of cloud migrations. The common categorized entities which evaluate security concerns in cloud platform is data access compliance and cloud infrastructural services. There are many implications which targets the security concern drive to enrich the research directions or comparisons of physical and software infrastructures, storage and network access environment data in the cloud which signified integrity, confidentiality, identity management and access control etc. Cloud security issues encompasses with abstraction of services, control execution, involvement of third party etc. A novel framework aggregates with security requirements review of security strengths and cloud security characteristics are disclosed.[36]

The need of auditing for privacy preserving models exhibits the uniqueness in elimination of possible threats from data privacy. The role of auditor can also helps in ensuring of quality of service and prevents from insider attacks in the integrity of data. Privacy preserving architecture proposed describes the detailed process on how the third party auditor submits an audit report to CSP and cloud user. The scenario starts with cloud user getting services from CSP and the same information is shared to TPA. TPA performs audit on the services by acquiring session's keys from CSP. The session key avoids auditor getting users private information with a time limit and audit report on SLA violations will be hand over to CU and CSP. In the reality the uncertainties involved as malicious from third party auditor, CSP and SLA violations. The proposed privacy model is endorsed with service type's insider malicious attacks of third party auditor's provider's authentication process and user's authentication

process for third party auditor (TPA). The validity of this model showcases provider's effectiveness, successful attempts of TPA malicious activities and auditing with reliability. [37]

## 2.4 Need of Provenance aware Auditing

Auditing is an independent assessment of cloud services and service providers. It triggers the compliance analyses and monitoring of resources to examine for the proposed benchmark. In general the performance audit initiates three tasks which resemble economy, effectiveness, and efficiency.

### 2.4.1 Considerable facts for Auditing QoS, Data, Accountability, intelligent decisions

The role of accountability in deploying of auditing focuses on different threats like malicious insiders, data leakages, uncertain risks, service hijackings, insecure programming interfaces etc. The comparative difference between accountability and audit ability shows that accountability deals with the events performed where as audit ability deals with logs maintenance. Enabling of cloud platform determines different physical and virtual machines which are monitored and enabled with advanced backup and operating system facilities to object the un authorized access for the utilization of services. The data accountability life cycle initialized with planning of policies, tracing, logging information, holding of logs safely, report generation, auditing and optimization. The above discussed life cycle phases concluded with three layers that is data layer, system layer, workflow layer which are synchronized with laws and policies. The data layer of this data accountability life cycle deploys the functionalities with provenance logger and consistency logger. In data accountable life cycle about workflow layer deals in managing of cloud with, automation auditing, and accountability of services with pre-processing post processing. In elevating of accountability, monitoring, tracking and domain segregations plays a main role. [38]

As the cloud computing is an aggregated task of services which are invoked by client from the provider that determines the quality of service. The highlighting point in measuring QoS endeavours the audit process in the cloud platform. The purpose of audit in cloud computing depends on several issues about the type of the

conflict, technical issues, evidence based documentations etc. There are some frameworks suggested for auditing task promoted by CSA, NIST, and ISACA which exhibit trust architectures. The role of auditor for the risk based approach delivers the assessment of risks, management objectives balancing. The need of auditing in IaaS and SaaS service in cloud effectively extracts the risk identities by an internal auditor. The participation of internal auditor in IaaS service targets connectivity issues, network service issues, security issues and data storage issues. The role of auditing in SaaS particularly concentrates on business workflow monitoring service level agreements and execution of processes.

In assessing of risk auditing task may be appended to the outsourcing third parties who determine the process of selection, defining and regulating workflows. Most of the auditing in cloud platform elevates the aspects on compliance monitoring governance of risks, security auditing, auditing database store, auditing business oriented agreements, auditing SLAs etc. The four major stakeholders in auditing scenario elevate their functionalities during audit process. DELOITE auditing framework enhances the role of internal audit and compliances which helps to identify risk based approach. It consists of tools proposed by NIST SP800. KPMG is an auditing approach that delivers more about risk relationships at every level of cloud. KPMG is auditing service that enables five levels of auditing by encompassing protection of data, risk technology, access management rules and operations.[39]

Auditing information flow represents the transparency of provenance in monitoring of compliance about personal data. Auditing of information flow is a pivot task where data controller takes a major role in audit trails. Traditional audit trails mainly focuses on internal working of applications rather than independent data and its involvements. The proposal of data centric audit shows a responsible solution for the above discussed demerit. The information flow control is achieved with trust computing base, operating system level and IFC polices with declassifies and endorsers. Data provenance defines where, when how and who based data involvement to target verification which outputs the quality.

A directed graph may be used for pictures the relationships between entities and information flows. The evaluated outcomes for information flow audit towards

provenance data showcases anonymous data can be transferred, explicit commitments on personal data, no more litigations when the data is deleted if the contract completes. The information flow audit can be picturized as a graph to validate compliance and regulations for the proposal of recommendations, third party services, authoritative decisions and aggregation process. The provocation issues for this information flow audit are visualization, data abstraction, policy based decision and legal relationships. [40]

The private data was exploited by un authorized stakeholders to check and regulate this issue auditing plays a prominent role by taking the base line evidence from the provenance. We know that data provenance is very helpful for auditing private data security concerns. An auditing structured with data protection act introduced by the European Union which consists of data controller, data subject and data processor entities. For the validity and reliability of audit data protection act is supported with eight principles which help private data protection. In practicing auditing principles, provenance states provenance aware applications with an interaction of p-assertions and relationships. The result of provenance queries is represented with directed acyclic graph which showcase the p-assertions on data. Provenance incorporating methodology (PrlMe) encompasses with three phases a) analysis phase deals with provenance capture, b) graph based approach by actor based decomposition and finally, c) involves with application adopting. [41]

Log analysis is a significant key feature for auditing in securing of provenance at cloud platforms. Accountability life cycles with auditing principles are evaluated based on logs which are enforced to determine credibility, privacy, audition and security concerns. This security auditing model addresses preventive controls and detective controls which aggregate a new accountable lifecycle. This accountable cloud life cycle deploys into five segments that is provision for legality, policy matters, mining of system data, storage for data prediction and application layer. The first segment promotes the law issues which incorporates log data protection and authorization. The policy segment undertakes the data behaviour and analysis for incorporation of methods. The data mining segment involves with the file system architectures and hardware topologies management. The data storage segment notices tracking of provenance, data consistency issues. The application segment targets

services auditing, file auditing and auditing in an automated manner. This auditing life cycle evaluates by comparison of post tracking, provenance security functions analyses, resistance from tampering, coupling factors, reliability and stability.[42]

As cloud computing is enabled with different services the philosophy of business commodities process as a service that takes a major role in promoting of remote auditing. In suggestion of this remote auditing for business process service tools should be deployed for maintaining good raise in economic values. Most of the business process services enable service oriented architectures where a business process was stated as set of activities with some specific services for a particular product to a consumer. The enabling this process service in cloud platform delivers in efficient management for supply chain activities.

The internal auditor has some restricted bounds where the services may not be useful in all decision making scenarios. The remote auditing is an outsourcing service which promotes business process service that eliminates location constraint. Information audit projects on dataflow issues where as operational audit targets the systematic operations evolution. There are some guidelines for business process service that is demonstrated by OECD, GLB ACT, HIPAA acts etc. In security concerns for this business process service auditing should be evaluated by the following factors authorized roles, differentiation of internal and intra instance for business process, guarantying approaches, privilege controlling, and four eye principle. The remote audit is enhanced with technical prerequisites like trusted computations, authenticated logs, and resistance for tampering. [43]

The promising use of big data adoption is more in present day computing environment where data analytics play a vital role in auditing. The research of this data analytics admits the auditing role and their significant improvements for the reality. Big data is termed with decisions evaluating on data having the features of high volume, velocity and variety with cost effective mode of processing information. The urgency of bulk volumes of data usage can be extracted through data mining tools to some extent. In the current day scenario public accounting firms want to polarize their audit computations by adopting the advantageous features of data accountability. Audit approach towards big data targets on two types of data like financial data and

non financial data. In non financial data the audit role is moderate and the usage of tools also is moderate for predicting the business outcomes and fraud detections. In financial data segment auditor takes an active part in collecting and testing data, moreover tools used in this segment by the auditor are up to the mark to extract the outcomes which show the benchmark. The provocation issues in undertaking of data analytics identifying of auditor, identification o of data availability expected level of integrity and expectations and regulations.[44]

Many complex computing systems are lagging for decision support for maintenance. The role of auditor in taking an intelligent decision is to run the machine in a consistent manner by overcoming the challenges. The usage of data provenance with the support of data analytics under the control of auditor exhibits the reliable and secure decision making for maintenance of cyber systems. DUNCAN and WHITTINGTON propose recommendation for audit trail in cloud computing are as follows data storage issues, log data migrations, information flows, marking digital entry points which are gathered by audit trail. An example of audit trail maintenance adopted for rail maintenance which discloses track the maintenance integrity factors of connected PDA devices, sensors monitoring, maintenance estimation and finally integrated with data mining AI machine learning approaches.[45]

Audit exhibits different patterns like performance audit, security audit and privacy audit. In deploying of performance audit the three characteristics that have to be undertaken are economy, efficiency and effectiveness. It is difficult to differentiate performance audit from traditional audit. A complete performance audit integrates audit on economy, audit on efficiency and audit on effectiveness where audit on economy incorporates planned resource utilization, audit on efficiency incorporates more optimal outcomes with simplified inputs, and audit on effectiveness incorporates perfect matches on desired input and acquired output.

An independent organization International Organization of Supreme Audit Institutions (INTOSAI) defines the standards for the above discussed performance audit concept. Organizing of performance audit evaluates with value chain strategy where policy programs, production models, economic performance are involved. There are different types of performance audits which are further classified as activity

audit, compliance audit and performance oriented audit. The performance audit organization depicts the role of auditor in an organization performance design as a dimensional pattern. The impact of performance audit reveals the prominence in many practices compared to the traditional auditing roles. [46]

## 2.5 Chapter Summary

We draw an attention about review of literature for solving our research problem in a planned sequential order. The first step need of provenance, considerable facts of provenance services and tracking were examined more from many research articles. Need of provenance for Big data and IOT with few considerable facts are chosen from many research scripts. Considerable facts for provenance security, trust and privacy are well interpreted and formed a baseline for scrutinizing provenance