Best Practices of Privacy
Preserving Computation
1 Overview

Privacy Preserving Computation (PPC): Link Data Islands under the Premise of Privacy Protection

It has been a major problem in the development and application of big data technology to link data islands and realize multiparty collaboration to release the value of data elements under the premise of ensuring data security, privacy protection and regulatory compliance. PPC is a technical solution to this problem. As a widely-recognized information technology, PPC can effectively protect data privacy in the circulation and integration of data elements.

With the increasing of laws and regulations on privacy protection and the awakening of people's awareness, countries and regions have accelerated their development and commercialization of PPC in recent years, including Europe, the US, China, Japan and South Korea among others. PPC has been applied to over 10 industries such as finance, scientific research and medical treatment, to protect data privacy in joint machine learning, joint statistics, joint scientific research, data publishing, outsourced computation, outsourced data query and etc.

With complex and fast-developing technologies, PPC has become a rapidly developing field. Its technologies and applications are still not yet well known by many organizations and individuals expecting to use it, and it's not easy for them to select the technologies and products in this field. This paper will focus on scenarios, technology selection, product selection and other topics in PPC and provide potential users with practices of PPC for future reference.
PPC can be used in three typical applications of data circulation, i.e. joint computing, data publishing and outsourced computing, to protect sensitive data. In joint computing, multiple parties jointly complete specified computations based on their data, such as modeling and statistical analysis based on multiparty data. In data publishing, the data provider offers its data to one or more users, and the users can perform analysis and computations on the data. In outsourced computing, the data provider uploads its data to a computing service provider (e.g. a cloud service provider), and the users can process the data by using rich computing and storing resources on the cloud. “Data” and “computation” are both broad concepts here. “Data” includes data sets, query conditions, model parameters in machine learning, computational logic among others. “Computation” includes model training, model inference, statistics and query among others.

Among these three major applications, joint computing, with its wide demand and difficulty in privacy protection, has attracted the most attention in the industry. This paper will mainly talk about PPC in joint computing.

The applications of PPC will be introduced in the scenario of finance, medical treatment and public services as follows.

2.1 Financial Scenario

Data elements are intensive in the financial industry and there are also many pain points in multiparty data collaboration and data value release.

**PPC prevents multiplatform loan.** In the internet financial industry, the credit overdue risk of multiplatform loan is 3 to 4 times that of ordinary loans. For every additional institution the borrower applies for credit in, the default possibility increases by 20%. Therefore, it’s an important step of industry risk control to accurately assess the multiplatform loan risk. Through PPC, institutions can jointly build an alliance for risk control in this industry, allowing members to obtain risk control data including risk blacklist, long-term loans, long-term overdue and long-term queries through secure query and supporting multiparty joint security modeling and joint risk prediction without outputing detailed data, and formulate a joint solution of risk prevention and control in the industry to greatly reduce the business risk of enterprises. PPC makes sure that members can only obtain statistical data within the alliance but can’t obtain the detailed data of any other member, ensuring the data security of all parties.

**Joint risk control improves risk identification capability in the financial industry.** To improve the risk identification capability in personal credit of the bank, it’s necessary to fully tap the value of its own data and introduce external data at the same time. In the offline modeling stage, aligning with the data service provider, the bank will use PPC to combine the samples from both parties for machine learning training to get a risk control model. In the online inference stage, the bank will encrypt the model to ensure the security of the model and query results.
2.2 Medical Scenario
Being the most sensitive part of personal private information, medical data are of high value and strong privacy. There has always been a contradiction between the privacy protection and the full use of medical data. However, PPC resolves the contradiction and shows its broad prospects in the medical field.

PPC supports joint medical research.
In-depth tapping into large-scale medical data by machine learning can improve the efficiency of medical research and disease inference as well as the accuracy and efficiency of medical service on the whole. However, the data samples of a single medical institution are limited, and it’s necessary to combine medical data from multiple institutions for joint modeling. Medical data belong to personal private information, the collection of which is prone to face the violation risk of laws and regulations. At the same time, competition in intellectual property rights in the medical industry is fierce. For the protection of commercial interests, data sharing will cause data-holding institutions to lose their advantages. Hospitals and medical research institutions can deploy PPC nodes in their private domains and connect the data for joint computing to their own nodes. Through joint privacy statistics, joint statistical analysis of diseases can be done under the premise that medical data won’t leave the private domains. In addition, the efficiency and accuracy of medical service can be improved through joint machine learning for disease inference.

2.3 Public Services Scenario
The scenario of public services is also important in PPC. Through PPC, first, data of different authorities can be integrated, to provide more convenient and intelligent services for the people. Second, data of multiple parties can be integrated into the scenario of public services, such as the scenario of the city brain. More data provide a strong impetus for the city to improve its governance of transportation, municipal facilities planning, security, commercial development and other aspects. Third, data of the government can be safely opened to the industry to help with its development.

3 Recommendations on PPC Technology Roadmap Selection
PPC is not a single technology, but involves a variety of privacy-preserving techniques, including cryptography, secure hardware, information theory, distributed computation, etc. Based on the principle of privacy protection, PPC can be divided into four types of technology roadmaps—Cryptography, Trusted Execution Environment (TEE), Data Obfuscation and Desensitization, and Distributed Computation. These technologies differ significantly in terms of the computation they can support, the dimensions of privacy protection, the strength of privacy protection, security, and performance, and apply to different applications and scenarios. Sometimes, it is also necessary to use a combination of two or more of these technologies to meet the application requirements.

- The representative technology of the Cryptography roadmap is MPC (secure Multi-Party Computation). This roadmap is to compute and retrieve data in an encrypted state, in which the input data and intermediate results are not exposed and only the final results are output.

- The Trusted Execution Environment (TEE) approach is to build a trusted and secure environment with trusted tamper-resistant hardware and software, in which data is processed by trusted programs.

- The representatives of Data Obfuscation and Desensitization techniques are Data Anonymization and Differential Privacy. This approach processes the data by means of information processing such as noise addition, sensitive info deletion, and data generalization before output.

- The Distributed Computation roadmap is represented by Federated Learning. Most distributed computation tasks do not require the output of the original data, but only intermediate results computed locally based on the original data. Compared with the centralized computation approach that aggregates and computes data from all parties, the leakage of original data is greatly reduced, but there is still exposure of intermediate results.

MPC and TEE are the primary technologies used for joint computation. Both of them can support various computations on multi-party data. In addition to protecting the data privacy of each party, they can also allow only the specified party to obtain the result or a part of the result of the computation. Moreover, each computation requires the cooperation and collaboration of all parties, so each party can have a fine control over what computation tasks the data is involved in, how many computations the data is involved in and which fields are involved, and easily implement the principle of minimal data usage to prevent data misuse and unauthorized use.

The following is the summary and comparison of the representative technologies of each technology roadmap for PPC, and technology selection recommendations are provided. Please refer to the Appendix for specific analysis of each technology roadmap.

In general, MPC and TEE can be used for various types of joint computation and outsourced computation, with respective strengths and weaknesses in terms of security and performance, Federated Learning is suitable for joint modeling which has no requirement for result control, Data Obfuscation and Desensitization can be used for data publishing and can also assist MPC and TEE for joint computation.

Recommendations on Technology Roadmaps. Starting from the function richness of joint computing, PPC products for joint computing must adopt MPC or TEE technical routes, supplemented by information desensitization and obfuscation technology to enhance privacy protection. In some modeling situations (models do not need to be controlled, information leakage is acceptable), federated learning can be adopted as a supplement. How to choose between MPC and TEE? MPC and TEE have
<table>
<thead>
<tr>
<th>Roadmaps</th>
<th>Privacy Protection of Inputs</th>
<th>Privacy Protection of Intermediate Results</th>
<th>Privacy Protection of Final Results</th>
<th>Control of Data Usage</th>
<th>Control of Results</th>
<th>Performance</th>
<th>Strengths (S) and Weaknesses (W)</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Secure Multi-Party Computation (MPC)</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Low to Medium</td>
<td>S: high security, high privacy protection, high control</td>
<td>1. High-security joint computation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>W: relatively low performance</td>
<td>2. Outsourced computation (requires multiple computation service providers)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. Federated Learning: gradient secure computation to reduce data leakage</td>
</tr>
<tr>
<td>Trusted Execution Environment (TEE)</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>High</td>
<td>S: high privacy protection, high control, high performance</td>
<td>1. Medium-security, high-performance joint computation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>W: Reliance on the rustworthiness of specific hardware and hardware vendors</td>
<td>2. Outsourced computation</td>
</tr>
<tr>
<td>Data Obfuscation and Desensitization</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>High</td>
<td>S: high performance</td>
<td>1. Data publishing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>W: low control, difficulty in balancing privacy protection and data availability</td>
<td>2. Joint computation: in combination with other roadmaps to reduce the exposure of sensitive information</td>
</tr>
<tr>
<td>Federated Learning</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Medium</td>
<td>S: development and implementation easiness</td>
<td>Joint modeling (all participants are allowed to obtain the model)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>W: low control over results, exposure of intermediate information (requires combination with other roadmaps to reduce the exposure)</td>
<td></td>
</tr>
</tbody>
</table>
their own advantages in security strength and performance. MPC has higher security strength while its performance is poorer than TEE. In order to support various scenarios flexibly, an ideal PPC product should support twin-engine of MPC and TEE. It allows users to choose engine flexibly, and to operate the PPC product in the same way.

4 Recommendations on PPC Product Selection

With the promulgation of Data Security Law of the People’s Republic of China and Personal Information Protection Law of the People’s Republic of China, many organizations begin to pay attention to PPC and hope to solve the problem of privacy protection in joint computing among organizations with the help of PPC. Facing the uneven PPC products among the market, how to choose? The selection recommendations are given below from the five dimensions of function, technical route, audit, integrated delivery, performance and security.

4.1 Whether the Function Covers Common Joint Computation

Joint Machine Learning: You need to pay attention to whether the PPC product provides the capability of feature processing and analysis, model algorithm support, and model effect evaluation indicators. If there is a demand for model privacy, you need to pay attention to whether the model can be kept secret. Also, to ensure the stability of model inference service, the PPC product should monitor the stability of results and underlying feature.

Joint Statistical Analysis: You need to pay attention to whether the support SQL operator meets the requirements. If there is a demand for scheduled tasks, you need to pay attention to the capability of scheduling tasks. In addition, it is critical to pay attention to the capacity of security verification on SQL scripts because the flexibility of SQL is quite high.

Custom Script: For various and customized computing requirements, you need to pay attention to whether PPC product supports custom computational logic, and compile them into secure computing scripts for execution.

Secure Matching: You need to pay attention to the amount of data supported and its performance.

Anonymous Query: You need to pay attention to whether the PPC product supports batch query and accounting by the quantity of return records.

4.2 Whether Support Audit

In many application scenarios, it is necessary to track and audit the computation data, processes, and results. For example, the Regulations for Financial Application Evaluation of Artificial Intelligence Algorithms issued by the Central Bank on March 2021, requires that the data used, the models adopted, the parameters of the models, and the computation results in the artificial intelligence algorithms of banks are traceable and auditable. Under the scenario of use of personal information, organizations must obtain user authorization and also prove the authorization. For scenarios that require multi-party consensus or regulatory audit, organizations need to pay attention to whether the PPC product had the ability blockchain ledger and audit when choosing PPC products.

4.3 System Integration and Delivery Capability

Account System Interconnection: Organizations generally have its own account management system, so you need to pay attention to whether the internal account of your organization can be used in PPC products easily and audit the account operation. First, it could improve internal user experience. Second, it could enable unified privilege management over the account.

Log System Interconnection: If there is a demand for your organization to manage, monitor and maintain your product operation and business logs of operators, you need to pay attention to the capacity of log system interconnection so that organizations can build log management and audit platforms according to their own needs.

Data Connection: When conducting data cooperation, you need to pay attention to how the data of organization can be easily connected to the PPC product and how different secure levels of data can be managed with different authorizations. The PPC product should support multiple forms of data access, such as database, files and APIs.

Delivery Capacity: You need to pay attention to whether the PPC product has diversified delivery capabilities for different delivery demands of your organization. For example, it is necessary for the organizations with self-built PPC platform to confirm whether the PPC product has encapsulated PPC APIs. And it is also critical for the organizations with higher security and performance requirements to confirm whether the PPC product has all-in-one machine to delivery.

4.4 Performance

Computation Performance Offline: Offline PPC functions generally include joint machine learning, joint statistical analysis, joint strategy, security matching and so on. As PPC scenarios become more and more abundant, and the requirements for computing efficiency become higher and higher, you need to pay attention to whether the PPC product can support large-scale distributed secure computing and hardware acceleration.

Computation Performance Online: For online production application scenarios for PPC, such as model prediction, strategy service, and anonymous query, you need to pay attention to the amount of requests and time lag, and also whether the online supporting Monitoring and Early-warning mechanism of the PPC products is complete to ensure the stability of online services.

4.5 Privacy Protection

Would you not worry about leakage of private data as long as the use of PPC technologies? The fact is that the risk of data leakage may still exist. There are four main reasons:

- The privacy protection capabilities of the PPC technologies adopted do not match the demand. As mentioned before, the privacy protection capabilities of each PPC technology are different. You need to adopt an appropriate technology or
sometimes a combination of multiple technologies.

- The security strength of the PPC algorithm used is insufficient, and it is easy to be destroyed. For example, some PPC algorithms cannot resist collusion attacks by participants, and any two participants colluding can obtain input data from other participants. If the possibility of collusion between the participants is high, technical means need to be adopted to prevent this attack, or alternatively, a more secure algorithm should be used.

- The authority is not strictly controlled. If the control is not strict, the partner may initiate a joint computing task beyond the scope and deadline.

- Lack of defense against malicious scripts and input. Taking joint statistics through SQL scripts as an example, the script may output the original data of the other party instead of the statistical results based on the data of both parties due to the flexibility of SQL. Even though the SQL is executed by PPC, data privacy is also destroyed.

In response to the above risks, the PPC products should reasonably select PPC technologies and algorithms, implement strict and precise authority control, and adopt technical means to detect and prevent malicious scripts and input.

5 Summary and Prospect

There are a wide range of industries and scenarios for PPC applications, but its applications also face huge challenges. The PPC technology is complicated, including a variety of technical routes and specific algorithms that are different in function, performance, privacy protection strength, and security strength, so reasonable technology selection is required. As a tool for linking data islands, the PPC products must serve applications well. So, expect for privacy protection capabilities, features such as rich computing types, high performance, auditability, easy integration, easy delivery, security and credibility are also required for a PPC product. To achieve these features at the same time is very challenging, especially it is almost impossible to achieve rich computing types, high performance, and high security at the same time. In order to improve these features, in addition to continuing to improve and develop PPC technologies, the industry is also actively developing methods that combine multiple PPC technologies, combining PPC with blockchain, adopting distributed cluster acceleration and all-in-one hardware and software acceleration.

Appendix—Analysis of PPC Technology Roadmaps

Cryptography Roadmap

The Cryptography roadmap includes technologies such as MPC, Homomorphic Encryption, Searchable Encryption, etc. The representative is MPC. This roadmap is to compute and retrieve data in an "encrypted state", in which the input data and intermediate results are not exposed and only the final results are output. In this process, MPC uses cryptographic techniques such as Secret Sharing, Garbled Circuits, Homomorphic Encryption, etc. to convert the plaintext data from multiple parties to the encrypted state for computation, and eventually converts the encrypted result back to the plaintext result.

MPC can support various computations on multi-party data. In addition to protecting the data privacy of each party, they can also allow only the specified party to obtain the result or a part of the result of the computation. Moreover, each computation requires the cooperation and collaboration of all parties, so each party can have a fine control over what computation tasks the data is involved in, how many computations the data is involved and which part of the data are involved, and easily implement the principle of minimal data usage to prevent data misuse and unauthorized use. MPC also has an optional feature to verify the validity of the computation.

Homomorphic Encryption, in addition to building MPC and achieving privacy protection for multi-party joint computation, can also be used for secure outsourced computation, so that the outsourced sensitive private data is invisible in the cloud. The weakness of Homomorphic Encryption is that it has a relatively poor performance and is not yet practical for computation with complex logic. Although MPC can also be used for secure outsourced computation, it requires multiple computing service providers (e.g., cloud providers) -- the original data will be split by means of Secret Sharing and sent to multiple computing...
service providers for joint computation using MPC techniques.

**Trusted Execution Environment Roadmap**

The Trusted Execution Environment (TEE) approach is to build a trusted and secure environment through trusted tamper-resistant hardware and software, in which data is processed by trusted programs. The environment has a certain resistance to external data theft, data tampering, and program tampering. The existing relatively mature TEE solutions include SGX, TrustZone, HyperEnclave, etc.

Like MPC, TEE also supports arbitrary computation logic for data privacy protection, controlled result output, and result validity checking. The differences between the two stem from the different implementation mechanisms, with TEE relying on the security of specific hardware and software, and MPC relying on the security of cryptographic algorithms. The former is relatively weak in security and requires a TEE hardware environment, but since the plaintext computation is performed in a secure environment, its performance is much higher than that of MPC in the encrypted state.

**Data Obfuscation and Desensitization Roadmap**

The representatives of Data Obfuscation and Desensitization techniques are Data Anonymization and Differential Privacy. This approach processes the data by means of information processing such as noise addition, sensitive info deletion, and data generalization before output. Data Anonymization is the processing of identifier information and quasi-identifier information in the data before output to prevent the association of the results to specific individuals. Differential Privacy is generally used in the computation of multi-person data. It prevents the individual information from being derived from the results by introducing noise in the computation results. The introduction of noise can be done by adding noise directly to the results or by adding noise to the input or intermediate results of the computation.

Data Obfuscation and Desensitization is suitable for data publishing scenarios. For example, medical authorities anonymize and desensitize the collected personal data and epidemic data for public release for the general medical institutions to study and for other institutions and individuals to do their epidemic prevention work. In such scenarios, the data is very wide in application and very large in quantity. The high cost of MPC or TEE and the strict control of the access will restrict the circulation and utilization of the data. On the contrary, the Data Obfuscation and Desensitization roadmap has the advantages of low cost, high performance, and easy implementation. Its weakness is the difficulty in balancing privacy protection and data utility—data obfuscation and desensitization removes certain information or reduces the accuracy of certain information that may be needed for certain computation and analysis. In this case, MPC or TEE is also needed.

Another major use of Data Obfuscation and Desensitization is the combination with other privacy-preserving technology roadmaps (e.g., MPC, TEE) to reduce the amount of sensitive information obtained from the results. MPC and TEE can protect the input data and intermediate computation results from leakage. However, if the final computation results contain sensitive information, then Data Obfuscation and Desensitization is further required before outputting the final results.
Distributed Computation Roadmap

The Distributed Computation roadmap is represented by Federated Learning. Most distributed computation tasks do not require the output of the original data, but only intermediate results computed locally based on the original data. Compared with the centralized computation approach that aggregates and computes data from all parties, the leakage of original data is greatly reduced. However, its performance in privacy protection is relatively weak compared with MPC and TEE methods, and there is additional exposure of intermediate results.

Federated Learning follows the parameter server-work server architecture of traditional distributed machine learning. In this architecture, a central server acts as the parameter server to coordinate multiple data source servers (as work servers) for joint machine learning model training, and the gradient information computed by each work server based on local data is delivered to the parameter server for aggregation, then the parameter server sends the latest iteration of model parameters to each work server. This architecture lacks control over the computation results, i.e., the model parameters are available to any participant.

Another weakness of Federated Learning is the leakage of the intermediate computation information (gradients). Several studies have shown the risk of exposing sensitive information in the original data in this leak. To reduce the gradient information leakage, Federated Learning generally uses MPC or Differential Privacy techniques for aggregating the gradients of all parties. Compared to the pure MPC roadmap where the whole modeling process is computed in an encrypted state, Federated Learning only introduces MPC's encrypted state computation in the gradient computation, which makes it easier to implement in research and development.
Privacy-enhancing computation enables data monetization and privacy protection scenarios that were not possible with previous approaches. IT leaders need to understand how PEC can support their confidentiality and privacy use cases.

Overview

Opportunities

• Privacy-enhancing computation (PEC) implements protection of data in use. It supports use cases where data must remain confidential during processing and analytics, and also where the algorithms must remain confidential even if the data itself does not. It enables data monetization and privacy protection scenarios that were not possible with previous approaches.

• PEC is an overarching term for many techniques, each with different security and privacy guarantees. They can be used individually, but may also be combined for greater efficacy. The exact choice of techniques depends on the use case and implementation requirements.

• Techniques have matured to the point that commercial implementations are available. Trusted execution environments are available in computer processors and hyperscale cloud providers. Secure multiparty computation and homomorphic encryption products are available from an increasing number of vendors.

Recommendations

IT leaders responsible for technology, information and resilience risk should:

• Identify candidates for PEC by assessing data processing activities that require the use of highly sensitive personal data for data monetization, fraud analytics and business intelligence use cases.

• Factor in PEC opportunities for more general use cases that have requirements to keep data and/or algorithms secure during data processing and analytics.

• Assess the differences in effectiveness and implementation needs of differential privacy, homomorphic encryption, secure multiparty computation, trusted execution environments and other approaches to ensure they are appropriate for the use case and required security and privacy guarantees.

Strategic Planning Assumption

By 2025, 50% of large organizations will adopt privacy-enhancing computation for processing data in untrusted environments and multiparty data analytics use cases.

Analysis

What You Need to Know

This research is part of Gartner’s Top Strategic Technology Trends for 2021.

Executive Guide to Privacy-Enhancing Computation (download).

The increased demand for data sharing, an unprecedented desire and opportunity to unlock value from data, and international data residency restrictions are powerful drivers to protect data in use. Following legal developments in a variety of countries, including in the EU, U.K., and U.S., identifiable data can’t be shared or transferred cross-border without additional privacy protection mechanisms.

In addition, many data analytics and business intelligence use cases serve secondary purposes — that is, other than the primary purpose for which personal data was obtained — which often leads to the need for anonymous data handling instead. Other confidential data, such as trade secrets or export-restricted information, has similar requirements with respect to confidentiality, even though broader privacy requirements do not apply. PEC techniques can enable privacy and confidentiality protection for data in use, ensuring expanded business activity and facilitating analytics and international transfers of data. It also reduces existing compliance and other privacy risks that may currently hinder (public) cloud adoption.

PEC techniques (see Figure 1) do not provide a singular approach to enhance privacy protection and secure data confidentiality. Instead, it’s a consolidation term for various techniques that can be applied in isolation or in combination with others, depending on the use case at hand. Many of the PEC techniques are listed in Hype Cycle for Privacy, 2020; stand-alone, these techniques have not yet passed the Peak of Inflated Expectations, yet combined and together they make for a consistent trend for the coming years.

Description

PEC comprises three types of technologies that protect data while it’s being used to enable secure data processing and data analytics.

The first provides a trusted environment in which sensitive data can be processed or analyzed. It includes trusted third parties and hardware-trusted execution environments (also called confidential computing).

The second performs processing and analytics in a decentralized manner. It includes federated machine learning and privacy-aware machine learning.

The third transforms data and algorithms before processing or analytics. It includes differential privacy, homomorphic encryption, secure multiparty computation, zero-knowledge proofs, private set intersection and private information retrieval.

Each technology provides specific secrecy and privacy guarantees, and some can be combined for greater efficacy.

Why Trending

Global data protection legislation is maturing and, with the unstoppable pervasiveness of personal data, every organization that processes personal data faces ever-higher privacy and noncompliance risks. At the same time, organizations now realize the economic potential of their data repositories.
Privacy-Enhancing Computation Techniques

The demand for processing data in untrusted environments and performing multiparty data sharing and analytics is rapidly growing. The increasing complexity of analytics engines and architectures mandates a by-design privacy capability, rather than a bolt-on approach. Unlike common data-at-rest security controls, privacy-enhancing computation protects data in use, thus enabling the use cases described previously, while maintaining secrecy or privacy. As a result, organizations can implement data processing and analytics that were previously impossible because of privacy or security concerns.

Implications

Hyperscale cloud providers have started to offer trusted execution environments. Organizations can use them to provide enhanced security and privacy in cloud environments. Other technologies, such as homomorphic encryption, secure multiparty compute and private information retrieval have transitioned from academic research projects to commercial solutions. Adoption is nascent, but implementations exist in fraud analysis, intelligence operations, finance and healthcare. These implementations follow one of two use-case patterns:

A third-party organization analyzes data from one or more organizations, without these organizations providing full access to their data.

Multiple organizations pool their data for joint analytics without having access to each other’s underlying data.

Actions

Identify candidates for privacy-enhancing computation by assessing data processing activities that require the use of highly sensitive personal data for data monetization, fraud analytics and business intelligence use cases.

Assess the differences in effectiveness and implementation needs of differential privacy, homomorphic encryption, secure multiparty computation, trusted execution environments and other approaches for these use cases.

About Gartner’s Top Strategic Technology Trends for 2021

This trend is one of our Top Strategic Technology Trends for 2021. The trends and technologies don’t exist in isolation, they
reinforce one another to enable people centricity, location independence and resilient delivery (see Figure 2). You should explore each of these trends for their applicability to your organization.

Evidence

1. Frequently Asked Questions on the Judgment of the Court of Justice of the European Union in Case C-311/18 — Data Protection Commissioner v Facebook Ireland Ltd. and Maximillian Schrems, EDPB.

2. Aside from the existing CLOUD Act, new laws may continue to disrupt level playing fields or influence assessments and viable technical alternatives. Examples include the EARN IT Act and the LAED 2020 Act.

Source: Gartner Research Note G00740641, Ramon Krikken, Bart Willemsen, 12 January 2021