

******* File Name Mapping *******

0: "before_Big_Data_Analytics_for_Security_Intelligence.txt"
1: "before_Big_Data_Taxonomy.txt"
2: "before_Comment_on_Big_Data_Future_of_Privacy.txt"
3: "before_CSA13-Top10Crypto_.txt"
4: "before_CSCC-Cloud-Customer-Architecture-for-Big-Data-and-Analytics.txt"
5: "iso_ISO-IECJTC1-WG9_N0087_N0087_WD_of_ISOIEC_20546_1st_Edition.txt"
6: "iso_N0147_ISO_IEC_20546_2nd_WorkingDraft_.txt"
7: "iso_N0154_ISO_IEC_20547-3_1st_Working_Draft.txt"
8: "iso_N0200_ISO-IEC_20546_Committee_Draft.txt"
9: "itu_ITU-T-A5-TD-new-Y.txt"
10: "itu_ITUbrochure.txt"
11: "itu_T-REC-Y.3600-201511-I!!PDF-E.txt"
12: "itu_T-REC-Y.Sup40-201607-I!!PDF-E.txt"
13: "nist_NIST.SP.1500-1.txt"
14: "nist_NIST.SP.1500-2.txt"
15: "nist_NIST.SP.1500-4.txt"
16: "nist_NIST.SP.1500-5.txt"
17: "nist_NIST.SP.1500-6.txt"
18: "nist_NIST.SP.1500-7.txt"

KeyWords extraction.

	KeyWord & Score	KeyWord & Score	KeyWord & Score	KeyWord & Score	KeyWord & Score	KeyWord & Score	KeyWord & Score	KeyWord & S
0	(data, [0.516541889402])	(security, [0.220244359146])	(big, [0.188203183691])	(informational, [0.151985197596])	(detecting, [0.140941495323])	(attacks, [0.14089581709])	(researchers, [0.132307792272])	(events, [0.1264091298])
1	(data, [0.577544795015])	(processed, [0.165415039801])	(databases, [0.142820997512])	(algorithm, [0.12581211284])	(time, [0.121299777294])	(application, [0.113367587131])	(compute, [0.111741544483])	(big, [0.1084440767])
2	(data, [0.496967186303])	(privacy, [0.278270545966])	(access, [0.176990130414])	(big, [0.174822980952])	(use, [0.163608677499])	(governing, [0.161987012153])	(analytics, [0.161424719651])	(policy, [0.1594968041])
3	(data, [0.5100164821])	(encryption, [0.230962982495])	(cloud, [0.184179919464])	(solution, [0.171719616893])	(privacy, [0.15861662025])	(filter, [0.152037548779])	(policy, [0.127373264338])	(movie, [0.1226692432])
4	(data, [0.593762613006])	(analytics, [0.217894497411])	(cloud, [0.215680561138])	(application, [0.169444256629])	(provided, [0.156683717022])	(enterprise, [0.143823337448])	(users, [0.140175383167])	(informational, [0.1253843652])
5	(data, [0.565269353882])	(iso, [0.185020558392])	(relation, [0.167607781644])	(standardization, [0.152570456168])	(documents, [0.138935313902])	(big, [0.13555996927])	(computing, [0.121699020753])	(terms, [0.1204895275])
6	(data, [0.542267133975])	(iso, [0.192048059691])	(nov, [0.170526881751])	(editors, [0.158867868942])	(standardization, [0.15703937458])	(relation, [0.148925353245])	(big, [0.1397050491])	(computing, [0.1348411905])
7	(data, [0.378453724529])	(big, [0.312244320071])	(activity, [0.297870757786])	(provided, [0.219864469599])	(architectural, [0.182263622278])	(function, [0.177137817984])	(role, [0.167319937494])	(component, [0.1453341455])
8	(data, [0.664894117785])	(big, [0.162541534481])	(standardization, [0.152758595375])	(processed, [0.14703125563])	(internal, [0.114806989313])	(relational, [0.107513767851])	(computing, [0.106490989409])	(need, [0.1064263305])
9	(itu, [0.402800819412])	(referred, [0.360240287017])	(information, [0.227548690584])	(documentation, [0.217633107565])	(group working party, [0.193492935108])	None	None	None
10	(data, [0.667562123292])	(big, [0.213799270076])	(http, [0.183513575698])	(itu, [0.115696523166])	(standards, [0.0969113669915])	(information, [0.0945041373647])	(technology, [0.0919568456265])	None
11	(big data, [0.380269549333])	(serviced, [0.209001585824])	(informative, [0.204565364474])	(provider, [0.134615900232])	(processing, [0.122352317656])	(activity, [0.118787919162])	None	None
12	(big data, [0.389696841798])	(networking, [0.199334999846])	(standardization, [0.170734778496])	(itu, [0.166929947114])	(services, [0.144366615411])	(provided, [0.13910237917])	(information, [0.136766443984])	None
13	(data, [0.650436361845])	(big, [0.226602651655])	(processed, [0.173575257332])	(analytic, [0.133654691798])	(nist, [0.115883821945])	(volumes, [0.115515083637])	(need, [0.114609716865])	(new, [0.1004441046])
14	(datas, [0.666833933905])	(provided, [0.170207995819])	(big, [0.169049746607])	(nists, [0.136057161334])	(requires, [0.126062336143])	(new, [0.110886527244])	(analytics, [0.104546203277])	(technologies, [0.0974222203])
15	(data, [0.510054607871])	(secured, [0.315191343493])	(big, [0.222358649375])	(privacy, [0.198409262475])	(inform, [0.151328613824])	(accessible, [0.12676174629])	(including, [0.11450360397])	(requirement, [0.1127388720])
16	(data, [0.558008578186])	(big, [0.188520052239])	(analytic, [0.180100871296])	(architectures, [0.15998715008])	(processed, [0.144312617853])	(management, [0.12918014896])	(infrastructure, [0.12894889771])	(supported, [0.1236705888])
17	(data, [0.532776389547])	(big, [0.188742708089])	(providing, [0.18070568099])	(processed, [0.160013107199])	(requirement, [0.159343418845])	(implement, [0.155728462359])	(management, [0.113301534221])	(application, [0.1129499920])
18	(data, [0.472503521358])	(standardizing, [0.233176435635])	(big, [0.190026778757])	(services, [0.184706638941])	(nist, [0.154213022748])	(processed, [0.131270892672])	(documents, [0.13048461892])	(specifications, [0.1273280356])

TF – IDF matrices and Search cosine similarity between documents.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	1.000000	0.663588	0.648583	0.379190	0.622026	0.555717	0.407898	0.530239	0.613775	0.031962	0.583664	0.522433	0.561803	0.660754	0.640200	0.693974	0.645668	0.634399	0.478013
1	0.663588	1.000000	0.648586	0.418512	0.632753	0.645309	0.468064	0.576972	0.718967	0.028837	0.600442	0.585616	0.612317	0.746822	0.723319	0.688153	0.721005	0.741063	0.540083
2	0.648583	0.648586	1.000000	0.447514	0.572629	0.559472	0.406272	0.554645	0.607149	0.021119	0.561516	0.524781	0.574831	0.667106	0.658167	0.721618	0.622470	0.638888	0.485913
3	0.379190	0.418512	0.447514	1.000000	0.429038	0.376429	0.272648	0.332974	0.423821	0.004452	0.365027	0.362849	0.363535	0.446500	0.425879	0.472189	0.419130	0.418562	0.312727
4	0.622026	0.632753	0.572629	0.429038	1.000000	0.595031	0.446737	0.562792	0.678017	0.025978	0.565411	0.613579	0.583802	0.700657	0.681959	0.651651	0.718971	0.709562	0.516782
5	0.555717	0.645309	0.559472	0.376429	0.595031	1.000000	0.782603	0.646314	0.839105	0.063326	0.616187	0.584619	0.649280	0.788180	0.731978	0.634002	0.703899	0.711811	0.576646
6	0.407898	0.468064	0.406272	0.272648	0.446737	0.782603	1.000000	0.510957	0.628386	0.051189	0.450132	0.428247	0.494276	0.570450	0.527878	0.463049	0.517770	0.520612	0.431896
7	0.530239	0.576972	0.554645	0.332974	0.562792	0.646314	0.510957	1.000000	0.664313	0.047301	0.555134	0.622874	0.643302	0.712921	0.752191	0.647357	0.728981	0.767596	0.579645
8	0.613775	0.718967	0.607149	0.423821	0.678017	0.839105	0.628386	0.664313	1.000000	0.049052	0.664465	0.650466	0.696707	0.869862	0.807042	0.698090	0.784735	0.793818	0.624503
9	0.031962	0.028837	0.021119	0.004452	0.025978	0.063326	0.051189	0.047301	0.049052	1.000000	0.114394	0.113093	0.162074	0.039584	0.038821	0.032145	0.035801	0.037791	0.048012
10	0.583664	0.600442	0.561516	0.365027	0.565411	0.616187	0.450132	0.555134	0.664465	0.114394	1.000000	0.587733	0.706000	0.704035	0.685037	0.630616	0.657741	0.658827	0.539268
11	0.522433	0.585616	0.524781	0.362849	0.613579	0.584619	0.428247	0.622874	0.650466	0.113093	0.587733	1.000000	0.691616	0.675580	0.675768	0.601975	0.655237	0.679386	0.521208
12	0.561803	0.612317	0.574831	0.363535	0.583802	0.649280	0.494276	0.643302	0.696707	0.162074	0.706000	0.691616	1.000000	0.715720	0.708639	0.566012	0.690613	0.708148	0.614303
13	0.660754	0.746822	0.667106	0.446500	0.700657	0.788180	0.570450	0.712921	0.869862	0.039584	0.704035	0.675580	0.715720	1.000000	0.929767	0.796897	0.877489	0.890085	0.709157
14	0.640200	0.723319	0.658167	0.425879	0.681959	0.731978	0.527878	0.752191	0.807042	0.038821	0.685037	0.675768	0.708639	0.929767	1.000000	0.826227	0.864023	0.912239	0.730423
15	0.693974	0.688153	0.721618	0.472189	0.651651	0.634002	0.463049	0.647357	0.698090	0.032145	0.630616	0.601975	0.666012	0.796897	0.826227	1.000000	0.770268	0.819563	0.673522
16	0.645668	0.721005	0.622470	0.419130	0.718971	0.703899	0.517770	0.728981	0.784735	0.035801	0.657741	0.655237	0.690613	0.877489	0.864023	0.770268	1.000000	0.879609	0.705076
17	0.634399	0.741063	0.638888	0.418562	0.709562	0.711811	0.520612	0.767596	0.793818	0.037791	0.658827	0.679386	0.708148	0.890085	0.912239	0.819563	0.879609	1.000000	0.741752
18	0.478013	0.540083	0.485913	0.312727	0.516782	0.576646	0.431896	0.579645	0.624503	0.048012	0.539268	0.521208	0.614303	0.709157	0.730423	0.673522	0.705076	0.741752	1.000000

Examples of Summaries by improved "TextRank" algorithm.

The length is 120 words.

Document ID	Summary
0	<p>The technological advances in storage, processing, and analysis of Big Data include (a) the rapidly decreasing cost of storage and CPU power in recent years; (b) the flexibility and cost-effectiveness of datacenters and cloud computing for elastic computation and storage; and (c) the development of new frameworks such as Hadoop, which allow users to take advantage of these distributed computing systems storing large quantities of data through flexible parallel processing.</p> <p>Big Data Analytics for Security Intelligence Analyzing logs, network packets, and system events for forensics and intrusion detection has traditionally been a significant problem; however, traditional technologies fail to provide the tools to support long-term, large-scale analytics for several reasons.</p>
4	<p>Provider cloud elements include: Data Integration Streaming Computing Data Repositories Actionable Insight Transformation and Connectivity. A cloud computing environment often allows provisioning decisions to be delayed until data volume, velocity and related processing requirements are better understood.</p> <p>Types of data repositories include: Copyright 2015 Cloud Standards Customer Council Catalog Data Virtualization Landing, Exploration & Archive Deep Analytics & Modeling Interactive Analysis & Reporting Results from discovery and IT data curation create a consolidated view of information that is reflected in a catalog.</p> <p>Key capabilities include: Copyright 2015 Cloud Standards Customer Council Page 10 Enterprise Security Connectivity Transformations Enterprise Data Connectivity API Management Monitors usage and secures results as information is transferred to and from the cloud provider services domain into the enterprise network to enterprise applications and enterprise data.</p>
9	<p>A.5 justification information for draft new Y.3600 (ex Y.BigData-reqts).</p> <p>Introduction According to ITU procedures, as described in ITU-T Recommendation A.5, any normative reference to documentation produced outside the ITU (other than ISO and IEC texts) needs to be evaluated by the study group or working party before a decision is made to incorporate the reference in an ITU-T Recommendation.</p> <p>This TD contains the A.5 justification information for new Y.3600 (ex Y.BigData-reqts).</p>

<p>13</p>	<p>Christine Hawkinson U.S. Bureau of Land Management Thomas Huang NASA Philippe Journeau ResearXis Pavithra Kenjige PK Technologies Orit Levin Microsoft Eugene Luster U.S. Executive Summary The NIST Big Data Public Working Group (NBD-PWG). Definitions and Taxonomy Subgroup prepared this NIST Big Data Interoperability Framework: Volume 1. Definitions to address fundamental concepts needed to understand the new paradigm for data applications, collectively known as Big Data, and the analytic processes collectively known as data science.</p> <p>To ensure that the concepts are accurate, future NBD-PWG tasks will consist of the following: Defining the different patterns of communications between Big Data resources to better clarify the different approaches being taken; taking into account the efforts of other working groups such as International Organization for Standardization (ISO) Joint Technical Committee 1 (JTC 1) and the Transaction Processing Performance Council; Improving the discussions of governance and data ownership; Developing the Management section; Developing the Security and Privacy section; and Adding a discussion of the value of data.</p>
<p>15</p>	<p>PII disclosure issues abound Various issues; for example, playing terrorist podcast and illegal playback Unknown Privacy-preserving data analytics Aggregate reporting to content owners Compliance with regulations Government access to data and freedom of expression concerns Data-centric security such as identity/policy-based encryption Policy management for access control Computing on the encrypted data: searching/ filtering/ deduplicate/ fully homomorphic encryption Audits Securing data storage and transaction logs Key management Security best practices for non- relational data stores Security against DoS attacks Data provenance User, playback administrator, library maintenance, and auditor Unknown Audit DRM usage for royalties Unknown N/A Traceability to data owners, producers, consumers is preserved Analytics for security intelligence Machine intelligence for unsanctioned Unknown Unknown use/access 37 NIST BIG DATA INTEROPERABILITY FRAMEWORK: VOLUME 4, SECURITY AND PRIVACY NBDRA Component and Interfaces Security and Privacy Topic Use Case Mapping Event detection Forensics Playback granularity defined Subpoena of playback records in legal disputes 6.2 NIELSEN HOMESCAN: PROJECT APOLLO Nielsen Homescan involves family-level retail transactions and associated media exposure using a statistically valid national sample.</p>

Summaries by algorithms: LSA, Kullback–Leibler, LexRank.

Document ID	Summary
0	<p><u>LSA algorithm:</u></p> <p>For example, Big Data analytics can be employed to analyze financial transactions, log files, and network traffic to identify anomalies and suspicious activities, and to correlate multiple sources of information into a coherent view.</p> <p>The human analyst is given the flexibility of combining multiple sensors according to known attack patterns (e.g., command-and-control communications followed by lateral movement) to look for abnormal events that may warrant investigation or to generate behavioral reports of a given users activities across time.</p> <p>By using a MapReduce implementation, an APT detection system has the possibility to more efficiently handle highly unstructured data with arbitrary formats that are captured by many types of sensors (e.g., Syslog, IDS, Firewall, NetFlow, and DNS) over long periods of time.</p> <p><u>Kullback–Leibler algorithm:</u></p> <p>Big Data Working Group Big Data Analytics for Security Intelligence September 2013 CLOUD SECURITY ALLIANCE Big Data Analytics for Security Intelligence v 2013 Cloud Security Alliance All Rights Reserved All rights reserved.</p> <p>Big Data differentiators The term Big Data refers to large-scale information management and analysis technologies that exceed the capability of traditional data processing technologies.¹ Big Data is differentiated from traditional technologies in three ways: the amount of data (volume), the rate of data generation and transmission (velocity), and the types of structured and unstructured data (variety) (Laney, 2001) (Figure 1).</p> <p>Experiments on a 2 billion HTTP request data set collected at a large enterprise, a 1 billion DNS request data set collected at an ISP, and a 35 billion network intrusion detection system alert data set collected from over 900 enterprises worldwide showed that high true positive rates and low false positive rates can be achieved with minimal ground truth information (that is, having limited data labeled as normal events or attack events used to train anomaly detectors).</p>

	<p><u>LexRank algorithm:</u></p> <p>Big Data Analytics for Security Intelligence Analyzing logs, network packets, and system events for forensics and intrusion detection has traditionally been a significant problem; however, traditional technologies fail to provide the tools to support long-term, large-scale analytics for several reasons: Big Data Analytics for Security Intelligence. The security data warehouse driving this implementation not only enables users to mine meaningful security information from sources such as firewalls and security devices, but also from website traffic, business processes and other day-to-day transactions. This incorporation of unstructured data and multiple disparate data sets into a single analytical framework is one of the main promises of Big Data. Big Data Analytics for Security Intelligence. The WINE Platform for Experimenting with Big Data Analytics in Security The Worldwide Intelligence Network Environment (WINE) provides a platform for conducting data analysis at scale, using field data collected at Symantec (e.g., anti-virus telemetry and file downloads), and promotes rigorous experimental methods (Dumitras & Shoue, 2011).</p>
4	<p><u>LSA algorithm:</u></p> <p>Another benefit is the ability to develop applications on dedicated resource pools in a hybrid cloud deployment that eliminates the need to compromise on configuration details like processors, GPUs, memory, networking and even software licensing constraints. Finally, because data is generally held in its original form for longer periods of time, it is possible to create multiple correlation and prediction algorithms to drive organizations towards better analytics and, ultimately, the best supported version of the truth. They promote better understanding of results by showing important areas of interest, highlighting outliers, offering innovative ways to refine and filter complex data, and by encouraging deeper exploration and discovery.</p> <p><u>Kullback–Leibler algorithm:</u></p> <p>Cloud Customer Architecture for Big Data and Analytics illustrates a simplified enterprise cloud architecture for big data and analytics.</p> <p>Provider cloud elements include: Data Integration Streaming Computing Data Repositories Actionable Insight Transformation and Connectivity A cloud computing environment often allows provisioning decisions to be delayed until data volume, velocity and related processing requirements are better understood.</p> <p>Copyright 2015 Cloud Standards Customer Council Data to be integrated can come from public network data sources, enterprise data sources, or streaming computing results.</p>

	<p><u>LexRank algorithm:</u></p> <p>These insights are used by users and enterprise applications as well as stored in data storage systems. Provider cloud elements include: Data Integration Streaming Computing Data Repositories Actionable Insight Transformation and Connectivity A cloud computing environment often allows provisioning decisions to be delayed until data volume, velocity and related processing requirements are better understood. Key capabilities include: Copyright 2015 Cloud Standards Customer Council Page 10 Enterprise Security Connectivity Transformations Enterprise Data Connectivity API Management Monitors usage and secures results as information is transferred to and from the cloud provider services domain into the enterprise network to enterprise applications and enterprise data.</p>
9	<p><u>LSA algorithm:</u></p> <p>A.5 justification information for draft new Y.3600 (ex Y.BigData-reqts).</p> <p>Introduction According to ITU procedures, as described in ITU-T Recommendation A.5, any normative reference to documentation produced outside the ITU (other than ISO and IEC texts) needs to be evaluated by the study group or working party before a decision is made to incorporate the reference in an ITU-T Recommendation.</p> <p>This TD contains the A.5 justification information for new Y.3600 (ex Y.BigData-reqts).</p> <p><u>Kullback–Leibler algorithm:</u></p> <p>A.5 justification information for draft new Y.3600 (ex Y.BigData-reqts).</p> <p>Introduction According to ITU procedures, as described in ITU-T Recommendation A.5, any normative reference to documentation produced outside the ITU (other than ISO and IEC texts) needs to be evaluated by the study group or working party before a decision is made to incorporate the reference in an ITU-T Recommendation.</p> <p>This TD contains the A.5 justification information for new Y.3600 (ex Y.BigData-reqts).</p> <p><u>LexRank algorithm:</u></p> <p>A.5 justification information for draft new Y.3600 (ex Y.BigData-reqts).</p> <p>Introduction According to ITU procedures, as described in ITU-T Recommendation A.5, any normative reference to documentation produced outside the ITU (other than ISO and IEC texts) needs to be evaluated by the study group or working party before a decision is made to incorporate the reference in an ITU-T Recommendation. This TD contains the A.5 justification information for new Y.3600 (ex Y.BigData-reqts).</p>

<p>13</p>	<p><u>LSA algorithm:</u> Six federal departments and their agencies announced more than \$200 million in commitments spread across more than 80 projects, which aim to significantly improve the tools and techniques needed to access, organize, and draw conclusions from huge volumes of digital data. Motivated by the White House initiative and public suggestions, the National Institute of Standards and Technology (NIST) has accepted the challenge to stimulate collaboration among industry professionals to further the secure and effective adoption of Big Data. While bounded in comparison to Big Data, past solutions considered legal, social, and technical requirements for privacy in distributed systems, very large databases, and in High Speed Computing and Communications (HPCC).</p> <p><u>Kullback–Leibler algorithm:</u></p> <p><u>LexRank algorithm:</u></p>
<p>15</p>	<p><u>LSA algorithm:</u> Six federal departments and their agencies announced more than \$200 million in commitments spread across more than 80 projects, which aim to significantly improve the tools and techniques needed to access, organize, and draw conclusions from huge volumes of digital data. Motivated by the White House initiative and public suggestions, the National Institute of Standards and Technology (NIST) has accepted the challenge to stimulate collaboration among industry professionals to further the secure and effective adoption of Big Data. Improved security software will include physical data correlates (e.g., access card usage for devices as well as building entrance/exit) and likely be more tightly integrated with applications, which will generate logs and audit records of previously undetermined types or sizes.</p> <p><u>Kullback–Leibler algorithm:</u> Mapping Web Traffic Analytics to the Reference Architecture Security and Privacy Topic Use Case Mapping NBDRA Component and Interfaces Data Provider Application Provider Application Provider Data Consumer Data Provider Framework Provider Framework Provider Fabric End-point input validation Real-time security monitoring Data discovery and classification Secure data aggregation Privacy-preserving data analytics Compliance with regulations Government access to data and freedom</p>

	<p>m of expression concerns Data-centric security such as identity/policy-based encryption n Policy management for access control Computing on the encrypted data: searching/filtering/deduplicate/fully homomorphic encryption Audits Securing data storage and transaction logs Key management Security best practices for non-relational data stores Security against DoS attacks Data provenance Device-dependent.</p> <p>Table 6: Mapping Pharmaceutical Clinical Trial Data Sharing to the Reference Architecture</p> <table> <tr> <th>Security & Privacy Topic</th><th>Use Case Mapping</th></tr> <tr> <td>End-point input validation</td><td>Real-time security monitoring</td></tr> <tr> <td>Data discovery and classification</td><td>Opaquecompany-specific Secure data aggregation</td></tr> <tr> <td>Privacy-preserving data analytics</td><td>Data to be reported in aggregate but preserving Opaquecompany-specific</td></tr> <tr> <td>None</td><td>Third-party aggregator NBDRA Component and Interfaces</td></tr> <tr> <td>Data Provider</td><td>Application Provider Application Provider Data Consumer</td></tr> <tr> <td>Compliance with regulations</td><td>Government access to data and freedom of expression concerns</td></tr> <tr> <td>Data Provider Framework</td><td>Provider Data-centric security such as identity/policy-based encryption</td></tr> <tr> <td>Policy management for access control</td><td>Framework Provider</td></tr> <tr> <td>Computing on the encrypted data: searching/filtering/deduplicate/fully homomorphic encryption</td><td>Audits</td></tr> <tr> <td>Securing data storage and transaction logs</td><td>Key management</td></tr> <tr> <td>Security best practices for non-relational data stores</td><td>Security against DoS attacks</td></tr> <tr> <td>Data provenance</td><td>42 potentially small-cell demographics</td></tr> <tr> <td>Responsible developer and third-party custodian</td><td>Limited use in research community, but there are possible future public health data concerns.</td></tr> </table> <p><u>LexRank algorithm:</u></p> <p>The scope of the Subgroups work includes the following topics, some of which will be addressed in future versions of this Volume:</p> <ul style="list-style-type: none"> Provide a context from which to begin Big Data-specific security and privacy discussions; Gather input from all stakeholders regarding security and privacy concerns in Big Data processing, storage, and services; Analyze/prioritize a list of challenging security and privacy requirements that may delay or prevent adoption of Big Data deployment; <p>2 NIST BIG DATA INTEROPERABILITY FRAMEWORK: VOLUME 4, SECURITY AND PRIVACY</p> <ul style="list-style-type: none"> Develop a Security and Privacy Reference Architecture that supplements the NBDRA; Produce a working draft of this Big Data Security and Privacy document; Develop Big Data security and privacy taxonomies; Explore mapping between the Big Data security and privacy taxonomies and the NBDRA; and Explore mapping between the use cases and the NBDRA. <p>Support both internal and third-party audits by unions, state agencies, responses to subpoenas Large enterprise security, transaction-level controlsclassroom to the fed</p>	Security & Privacy Topic	Use Case Mapping	End-point input validation	Real-time security monitoring	Data discovery and classification	Opaquecompany-specific Secure data aggregation	Privacy-preserving data analytics	Data to be reported in aggregate but preserving Opaquecompany-specific	None	Third-party aggregator NBDRA Component and Interfaces	Data Provider	Application Provider Application Provider Data Consumer	Compliance with regulations	Government access to data and freedom of expression concerns	Data Provider Framework	Provider Data-centric security such as identity/policy-based encryption	Policy management for access control	Framework Provider	Computing on the encrypted data: searching/filtering/deduplicate/fully homomorphic encryption	Audits	Securing data storage and transaction logs	Key management	Security best practices for non-relational data stores	Security against DoS attacks	Data provenance	42 potentially small-cell demographics	Responsible developer and third-party custodian	Limited use in research community, but there are possible future public health data concerns.
Security & Privacy Topic	Use Case Mapping																												
End-point input validation	Real-time security monitoring																												
Data discovery and classification	Opaquecompany-specific Secure data aggregation																												
Privacy-preserving data analytics	Data to be reported in aggregate but preserving Opaquecompany-specific																												
None	Third-party aggregator NBDRA Component and Interfaces																												
Data Provider	Application Provider Application Provider Data Consumer																												
Compliance with regulations	Government access to data and freedom of expression concerns																												
Data Provider Framework	Provider Data-centric security such as identity/policy-based encryption																												
Policy management for access control	Framework Provider																												
Computing on the encrypted data: searching/filtering/deduplicate/fully homomorphic encryption	Audits																												
Securing data storage and transaction logs	Key management																												
Security best practices for non-relational data stores	Security against DoS attacks																												
Data provenance	42 potentially small-cell demographics																												
Responsible developer and third-party custodian	Limited use in research community, but there are possible future public health data concerns.																												

	<p> eral government CSOs from the classroom level to the national level --- Standard NBD RA Component and Interfaces Data Provider Application Provider Application Pr ovider Data Consumer Compliance with regulations Government access to data and freedom of expression concerns Data Provider Framework Provider Data-centric s ecurity such as identity/policy-based encryption Policy management for access control Computing on the encrypted data: searching/filtering/deduplicate/fully homomorphic en crypton Audits Framework Provider Securing data storage and transaction logs Ke y management Security best practices for non- relational data stores Security against DDoS attacks. </p>
--	---

