

**Request for Information (RFI) 19-601 Data Visualization Solution for Complex Ownership Structures  
Questions and Answers**

#	Vendor Question	Department Response
1	Is this considered an urgent need for the department? If there is an RFP that may be forthcoming, is there an estimated release timeframe? Or a project start or implementation date?	This Request for Information (RFI) is issued solely for the research and gathering of information to learn about the options available in the marketplace regarding Data Visualization Solutions for Complex Ownership Structures – It is not a Request for Proposals (RFP) or a promise to issue an RFP in the future.
2	Has funding been secured for this new solution?	This is not applicable as this is a research and information gathering document only.
3	Our company specializes in creating solutions that include open data, big data, custom application development and data visualizations, however, our solutions are typically provided as work-for-hire, not as a licensed product. Your questions seem to indicate an expectation of an existing tool or commercial-off-the-shelf solution. Are you amenable to a custom solution that can work with your data to visualize complex ownership structures?	The Department does not anticipate entering into a work-for-hire contract. However, the Department will accept responses for all solutions.
4	a) Are you able to provide sample data (redacted of course), that we could see before preparing a response to this RFI?  b) Do you have a single source of truth for data and existing master data management solutions in place?	a) Not at this time.  b) We have data of varying types, depending on the type of information captured.
5	Can you describe some of the unique data points that can connect across multiple data sets (e.g. tax ids such as EIN or SSN come to mind, but are you looking for softer "potential" matches such as multiple and historical address information, additional taxpayer relationships (that also consider variations in spelling, married vs maiden names etc. that may surface when cross-referencing multiple data sets?	We are looking to include soft or potential matches accompanied with confidence levels that can be evaluated by users.
6	a) Could this system also integrate and cross-reference 3rd party, private and or public data sources or data sources from other agencies such as DMV records, legal records, USPS records, department of state records, etc?  b) Could this system also consider data from other states or federal records in order to discover potential connections and relationships?	a) We are looking to integrate our own data, and data from other agencies with which we have exchange agreements.  b) Where permissible by law.

**Request for Information (RFI) 19-601 Data Visualization Solution for Complex Ownership Structures  
Questions and Answers**

#	Vendor Question	Department Response
7	Will the Department of Taxation and Finance be providing all data sources regardless of origin?	Yes, with the potential for external data.
8	a) Are you expecting to integrate data from multiple places in order to use this tool for visualizing complex owner structures? b) Would this be a combination of real-time integrations, scripted/scheduled ETL jobs and/or actual import/export uploads of structured data (fixed-width, csv, etc.) or a combination thereof?	a) Yes. b) To be determined.
9	a) Do you envision this tool existing as a stand-alone application with it's own Single Sign on for authorized users? or could this extend or add onto existing systems that are internal to the Department of Taxation and Finance today--such as an intranet or auditor's case management system. b) If the latter, what technologies and technology stack does the department use today and how might we integrate our solution with those systems?	a) Either. b) Information pages are built on a custom framework with Java being the main language used.
10	What is the preferred platform for hosting this solution? Would it be on premise, public or private cloud, or Software as a Service?	On-premise.
11	Does the application need to be browser based?	Not necessarily but would need to run in a VDI Windows 10 environment.
12	a) On what platform(s) is the data currently hosted? b) How is it accessed currently for analysis?	a) DTF has a Data Warehouse that houses 4.8 TB of data in a DB2 format. b) Various.
13	Is there a requirement to save off and export reports?	Yes.
14	a) Is there a requirement to generate alerts? b) If so, what types of alerts are required and how frequently is the data updated?	a) Yes, several types. b) Data is updated at different rates depending on the source system and use cases the data supports.

**Request for Information (RFI) 19-601 Data Visualization Solution for Complex Ownership Structures  
Questions and Answers**

#	Vendor Question	Department Response
15	a) What is the approximate number of entities that need to be present for the analysis and what is the size in KB of the data stored for each entity?  b) Does the entire tax return need to be stored for analysis, or can the system point to the location of the tax return for viewing by the analyst?	a) To be determined.  b) Unable to answer at this time.
16	a) Where is the data coming from, and how frequently does it change?  b) What is the requirement for timeliness of the transactional data being analyzed? Is it near real-time, or are daily updates sufficient?	a) Our data comes from our own data warehouse and updates on a regular schedule.  b) To be determined.
17	How many years of historical data can DTF provide for all the data sources identified?	The potential for multiple timeframes.
18	Where does the data reside?	DTF has a Data Warehouse that houses 4.8 TB of data in a DB2 format. Most, if not all, of the data needed for this solution resides there.
19	How many data sources, types and volume?	Billions of data elements spanning multiple data sources of varying structures.
20	What is your current data architecture?	Unable to answer at this time.
21	What are the current tools used for ETL and Visualization? If no, any preferred tools?	Unable to answer at this time.
22	Proposed solution should be in Cloud or on-premise?	On-premise.
23	Kindly elaborate the below points with examples(wherever applicable) for better understanding <ul style="list-style-type: none"> <li>• The solution has a user-friendly and flexible user interface that allows any user to search for filers, display connections and activities, and customize searches and results. Additional options can be added, modified, or removed from the user interface.</li> </ul>	The tool will be used by employees of varying skillsets in working with data. The user interface needs to be easily understandable and usable by employees that don't necessarily work with data analytics every day.  The tool must be able to:

**Request for Information (RFI) 19-601 Data Visualization Solution for Complex Ownership Structures  
Questions and Answers**

#	Vendor Question	Department Response
	<ul style="list-style-type: none"> <li>The solution produces visualizations and final work-products that can be stored on a user's hard drive and shared amongst other users and non-users.</li> <li>The solution makes data connections using exact matches and near-matches (fuzzy matches), if needed. It displays confidence rates of each match.</li> <li>The solution is reliable and effective when employed on extremely large volumes of data elements (hundreds of millions of rows and columns across multiple databases and datasets).</li> <li>The solution can be structured to grant various levels of access and permissions to users and admins. The access to federal or sensitive data can be easily toggled on and off based on a user's need for the information, and flag visualizations as federal or sensitive when it uses that data.</li> <li>The solution has ad-hoc reporting capabilities. Please describe the reporting capabilities and provide samples.</li> <li>What technologies are used to analyse, connect, and draw conclusions on the data it ingests? Does the solution provide graph database capabilities or similar technologies? Please provide any examples of how that technology can be used for the requirements listed in this RFI.</li> </ul>	<ul style="list-style-type: none"> <li>be flexible enough to modify criteria, filters, and elements included in the visualization without the need for Office of Information Technology Services (ITS) intervention;</li> <li>produce a PDF, excel file, word document, etc. that can be shared with other employees electronically or printed on demand. Not all employees will have use of the data visualization tool but will need the documentation produced from it;</li> <li>map known connections with 100% confidence and verify discovered connections that have less than 100% confidence;</li> <li>handle large volumes of data;</li> <li>grant different tiers of access; and</li> <li>generate reports of any results.</li> </ul> <p>Please describe in your RFI response how your tool would make connections in our data.</p>
24	Is power user same as admin user? If not kindly elaborate.	No. The admin user will be our developers and ITS staff. The power user will be for users working primarily with data analytics.
25	<p>Below points will require the client's regular inputs to evaluate whether solution is compliant:</p> <ul style="list-style-type: none"> <li>The solution is substantially compliant with the provisions of the New York State Tax Law.</li> <li>The solution is substantially compliant with the secrecy and confidentiality requirements of the New York State Department of Taxation and Finance, New York State Law, and the IRS. This also includes any confidentiality</li> </ul>	The contractor must comply with our secrecy and confidentiality rules when dealing with sensitive data.

**Request for Information (RFI) 19-601 Data Visualization Solution for Complex Ownership Structures  
Questions and Answers**

#	Vendor Question	Department Response
	agreements with other parties that provide the Department data through MOUs, including the IRS.	
26	Is there a retention limitation, i.e., does data age out and need to be removed?	Retention limitations are based on data sources.
27	How many concurrent users will access the system at peak times?	Unable to be determined at this time. Your response should notate any solution limitations in this aspect.
28	Do different user types have different role-based privileges on what they are allowed to see?	Yes.
29	a) Do any DTF staff (who would be available to support the system) have experience and training developing analytical models? b) Do these employees have any experience with graph or network analytics?	a) Yes b) Unable to answer at this time.
30	What are the tax return types (and associated demographics) that would be integrated by DTF into the system in the short term (e.g. corporate income tax returns, personal income tax returns, withholding, sales tax, other)?	All tax return types.
31	a) What are the other internal or external data sources (beyond tax account data) that DTF would like eventually integrated into the corporate structure analysis? b) Is this data structured, unstructured or both?	a) Yes, we are looking to integrate multiple sources of data. b) Structured.
32	Will NY DTF be responsible for getting all the MOUs for the data needed from other third parties?	Yes.
33	Does DTF have rough estimates of the approximate amount of data that will ultimately be housed (over time) in the system (e.g. < 1 TB, 1-5TB, 5+ TB) and an estimated annual growth rate?	Not at this time.
34	Does DTF want a cloud-based or on-premise solution?	On-premise.
35	Can DTF provide an estimate on the data volume that will have to be processed by this solution?	Not at this time.

**Request for Information (RFI) 19-601 Data Visualization Solution for Complex Ownership Structures  
Questions and Answers**

#	Vendor Question	Department Response
36	<p>a) Would DTF want to build corporate structure visualizations for all New York taxpaying businesses, or only a specific pre-identified subset?</p> <p>b) Can DTF provide an approximate number of businesses that would be modeled in this manner?</p>	<p>a) Visualization tool for multiple types of filers. Broad based not pre-determined.</p> <p>b) Not at this time.</p>
37	Is DTF looking for fraud detection/alerting capabilities, or just the network visualization and analysis? Or both?	Primarily network visualization analysis.
38	How will/can the data be provided - what is the format, frequency, and refresh rate of the data?	DTF has a Data Warehouse that houses 4.8 TB of data in a DB2 format. Scheduled refresh rates depend on datasets.
39	Does the department currently have a data visualization solution in place? If so, was this provided by an outside vendor, and are they currently under contract?	Unable to answer at this time.
40	Is DTF using any graph database technologies today? If so, what are they?	Unable to answer at this time.
41	Has DTF tried previously to do corporate structure visualization and analysis? If so, what software tools were used? Would DTF be willing to share a bit more about the approaches, successes, and challenges of those attempts in order to help vendors make the best recommendations for your organization?	Unable to answer at this time.
42	What type of access control is used currently by DTF? Would integration with active directory be required/preferred?	Active Directory is used for most every security implementation and would be preferred.
43	Is the data that will be integrated into the system considered "FTI" or is it state-only tax data?	We would like to give employees the option of viewing FTI and state data, or state-only data.
44	Is there an existing audit or case management system that the Corporate Structure Visualization solution will need to directly integrate with, or is the Corporate Structure Visualization system envisioned as standalone solution?	Standalone solution.
45	Are there data privacy requirements?	Yes.

**Request for Information (RFI) 19-601 Data Visualization Solution for Complex Ownership Structures  
Questions and Answers**

#	Vendor Question	Department Response
46	<p>a) Is there a more detailed list of compliance requirements for provisions within New York State Tax Law, or secrecy/confidentiality requirements within the New York State Department of Taxation and Finance, New York State Law and the IRS that this RFI seeks to address?</p> <p>b) Can one be provided to ensure that the proposed solution addresses these compliance requirements?</p>	See response to Question #48.
47	If a corporate structure has sensitive data in it (e.g. FTI, relationships with VIP taxpayers), and DTF wants to restrict certain employees from viewing it, would “minimization” of the sensitive data be required at the row level (e.g. hide only the specific transactions that are FTI, but still show rest of network) or is stopping an employee from viewing the entire structure sufficient?	We would like to give employees the option of viewing Federal Tax Information (FTI) and state data, or state-only data.
48	Can you provide the secrecy and confidentiality requirements of the NY DTF?	<p>Vendors who enter into a contract with the Department are required to agree to New York State Tax Law and Internal Revenue Code Tax Secrecy Provisions. Minimally, the contractor would have to agree to language substantially similar to the following:</p> <p>The various secrecy provisions of the Tax Law (e.g., Tax Law §§ 487, 697(e) and 1825) prohibit independent contractors from disclosing tax information in any manner and provide for misdemeanor prosecution for violations. The secrecy provisions of the Internal Revenue Code (26 USC § 6103) provide for felony prosecution for unauthorized disclosure of Federal tax information in the possession of the Department.</p> <p>All other information about the Department's operations not covered by the preceding provisions of law must be kept confidential as if it were so covered. Contractor representatives must comply with the administrative procedures enforcing these rules.</p> <p>The Contractor, all staff members and Subcontractors shall agree to view, access, and use only that information relevant and necessary to provide Services to the State</p>

**Request for Information (RFI) 19-601 Data Visualization Solution for Complex Ownership Structures  
Questions and Answers**

#	Vendor Question	Department Response
		<p align="center">under the Agreement; and to subscribe to the provisions of §§ 73 and 74 of the Public Officers Law.</p> <p>The above should not be construed to be the entirety of secrecy and confidentiality requirements the contractor may need to adhere to. There may be other confidentiality requirements in addition to the above contract language should the Department move forward with an actual contract. The Department is unable to delineate any additional requirements at this time.</p> <p>Please describe in your RFI response your capacity to safeguard confidential information.</p>
49	Are test / dev environments required? If so, how many?	Yes, Dev and UTM environments are required. The number of environments is not known at this time.
50	Can NY Tax provide more details around Compliance and how is it defined? Is it compliance in how the data is stored? Technology compliance? Please provide more information so that we can better address the requirement.	IRS Publication 1075 (PUB1075), National Institute of Standards and Technology (NIST) 800-53, and the IRS Safeguard Computer Security Evaluation Matrix (SCSEM) based on the technology used.
51	Again similar to #10 Can NY Tax provide more details around Compliance and how is it defined? Is it compliance in how the data is stored? Technology compliance? Please provide more information so that we can better address the requirement.	IRS Publication 1075 (PUB1075), National Institute of Standards and Technology (NIST) 800-53, and the IRS Safeguard Computer Security Evaluation Matrix (SCSEM) based on the technology used.
52	<p>Please provide guidance related to the relevant provisions of the New York State Tax Law referenced in the below item. If possible, please provide the relevant provisions of the tax law.</p> <p><i>The solution is substantially compliant with the provisions of the New York State Tax Law.</i></p>	The Department would like to clarify that it is the contractor who needs to be compliant with any secrecy and confidentiality requirements. Please see the response to Question # 48 above.
53	Please provide guidance related to the relevant provisions of the New York State Tax Law referenced in the below item. If possible, please provide the relevant provisions of the tax law.	The Department would like to clarify that it is the contractor who needs to be compliant with any secrecy and confidentiality requirements. Please see the response to Question # 48 above.

**Request for Information (RFI) 19-601 Data Visualization Solution for Complex Ownership Structures  
Questions and Answers**

#	Vendor Question	Department Response
	<i>Has your solution been proven to be substantially compliant with the provisions of the New York State Tax Law? If so, where is it currently employed? Please provide client contact information.</i>	