RESEARCH & DEVELOPMENT

Facilitating the New Statewide GIS Metadata Standard through Training and Outreach

Timothy Mulrooney, Ph.D.
Department of Environmental, Earth and Geospatial Sciences
North Carolina Central University

NCDOT Project 2017-20

FHWA/NC/2017-20

April 2019

FINAL REPORT

Facilitating the New Statewide GIS Metadata Standard through Training and Outreach

North Carolina Department of Transportation Research Grant 2017-20

Submitted
by

Timothy Mulrooney
Department of Environmental, Earth and Geospatial Sciences
North Carolina Central University

Technical Report Documentation Page

1. Report No. FHWA/NC/2017-20	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Facilitating the New Statewide GIS Metadata Standard through Training and Outreach		5. Report Date 4/4/2019
		6. Performing Organization Code
7. Author(s) Timothy Mulrooney		8. Performing Organization Report No.
9. Performing Organization Name and Address Department of Environmental, Earth and Geospatial Sciences 2202 Mary Townes Science Complex 1800 N. Fayetteville Street Durham, NC, 27707		10. Work Unit No. (TRAIS)
		11. Contract or Grant No.
12. Sponsoring Agency Name and Address North Carolina Department of Transportation Research and Development Unit 104 Fayetteville Street Raleigh, North Carolina 27601		13. Type of Report and Period Covered Final Report 8/1/2016 – 3/21/2019
		14. Sponsoring Agency Code 2017-20
Supplementary Notes:		

16. Abstract

Under the supervision of the NCGICC (North Carolina Geographic Information Coordinating Council) and SMAC (Statewide Mapping Advisory Committee), an ad hoc metadata committee defined and developed a State and Local Government Metadata profile for GIS data intended for use in North Carolina. This standard is based on the ISO 191** format and is an improvement over prior metadata standards to account for evolving technologies such as remotely sensed imagery, online services and ontologies that did not exist when these original metadata standards were first published. It is necessary to educate all NCDOT GIS data users on the utility and application of this new GIS standard to ensure that information relating to a data set's creation, structure and administration is captured in a proactive, complete, uniform and timely manner. Under the direction of the NCDOT and NCGICC, the research team will work to help disseminate this metadata standard to the many developers and users of NCDOT GIS data and beyond through hands-on workshops, training tutorials and continued support of this metadata initiative.

17. Key Words GIS, Geographic Information Systems, Metadata, Metadata Education, Metadata Training, Spatial Metadata	18. Distribution Statement		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 22	22. Price

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

DISCLAIMER

The contents of this report reflect the views of the author(s) and not necessarily the views of the University. The author(s) are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of either the North Carolina Department of Transportation or the Federal Highway Administration at the time of publication. This report does not constitute a standard, specification, or regulation.

Acknowledgements
Timothy Mulrooney and the NCCU Research Team wishes to thank the NCDOT/NCDIT and NCGICC for the opportunity to work on a project of this magnitude. This project has provided valuable work experience for students getting ready to enter the workforce. NCCU is also grateful for input provided throughout the course of this project. While not fully complete, those who helped facilitate this project included: Sarah Wray, Chelsea Duncan, Jeff Brown and Mustansir Kadibhai. Timothy Mulrooney and NCCU are happy to continue offering training or assist those who have questions about this document. His contact information is below:

Timothy Mulrooney
Department of Environmental, Earth and Geospatial Sciences
North Carolina Central University
1246 Mary M. Townes Science Complex
1801 Fayetteville St., Durham, NC, 27707
Phone: 919-530-6575
E-mail: tmulroon@nccu.edu

Research Need Statement (RSN) ID: Researcher Generated Proposal
Preliminary Proposal Title: Facilitating the New Statewide GIS Metadata Standard through Training and Outreach

Executive Summary: Under the supervision of the NCGICC (North Carolina Geographic Information Coordinating Council) and SMAC (Statewide Mapping Advisory Committee), an ad hoc metadata committee defined and developed a State and Local Government Metadata profile for GIS data intended for use in North Carolina. This standard is based on the ISO 191** format and is an improvement over prior metadata standards to account for evolving technologies such as remotely sensed imagery, online services and ontologies that did not exist when these original metadata standards were first published. It is necessary to educate all NCDOT GIS data users on the utility and application of this new GIS standard to ensure that information relating to a data set's creation, structure and administration is captured in a proactive, complete, uniform and timely manner. Under the direction of the NCDOT and NCGICC, the research team worked help disseminate this metadata standard to the many developers and users of NCDOT GIS data and beyond through hands-on workshops, training tutorials and continued support of this metadata initiative.

This this project support was provided to the NCDOT and NCGICC in the form of 10 presentations at conferences at workshop, direct work with the NCDOT/NCDIT, the development of training videos and workshops for face-to-face training on campus and the development of Python programming scripts to automate metadata assessment and evaluation.

Contents

Introduction

GIS (Geographic Information Systems) data are the digital representation of the world in which we live. A GIS serves as the tangible and intangible means by which spatially related phenomena can be created, stored, analyzed and rendered. In the North Carolina GIS community, GIS is used to map transportation routes, delineate land ownership parcels, highlight patterns of crime and help make zoning decisions. The manner, however, in which we capture these data varies. Some methods include using a GPS (Global Positioning System) unit, extracting or improving existing GIS data, or creating data from an analog format using a process called digitization. Regardless of the method, the resources (e.g., the computers, time and people dedicated to the process of collecting and creating GIS data) are the most time-consuming portion of a GIS-related project. As a result, the GIS community needs to ensure that the quality of the GIS data created as a result of these methods is captured and assessed in a systematic way. While many only see the output of GIS data and analysis in the form of maps, the most costly component of any geospatial enterprise is the creation of these spatial data that contribute to these analyses and output.

Geospatial metadata serves as the formal framework to catalog descriptive (what the data are trying to show), administrative (who is in charge of the data) and structural (how the geo-information is stored and expressed) information about GIS data. GIS metadata is a subset of electronic metadata which catalogs electronic resources such as web pages and software applications. However, GIS metadata is inherently different than its electronic media counterpart because each metadata file can be applied to a spatial component that is not implicit with other forms of metadata.

While geo-information has proliferated because of technology such as the Internet, GPS units and GPS-enabled phones, the time dedicated to cataloging and assessing this information remains nearly the same. Research has shown that one of the major deterrents of GIS metadata is the time and energy required to maintain it. The FGDC (Federal Geographic Data Committee) GIS data standard, commonly referred to as the CSDGM (Content Standard for Digital Geospatial Metadata) allows for more than 400 individual metadata elements to be entered by the user. The GIS community needs to understand the requirements dictated by metadata standards so each piece of information can be fully understood and adequately populated. Any GIS project is only as good as the information on which it is based, while some go so far to say that GIS data and therefore the proceeding analysis is worthless without the accompanying metadata. Any conscientious GIS business model should budget adequate personnel resources to the creation and update of its metadata. The North Carolina GIS community has been proactive about understanding the importance of metadata. Under the supervision of the NCGICC (North Carolina Geographic Information Coordinating Council) and SMAC (Statewide Mapping Advisory Committee), an ad hoc Metadata Committee was created and developed a State and Local Government Metadata profile for GIS data intended for use in North Carolina.

The programmatic goal of the Metadata Committee is to "achieve statewide creation and maintenance of complete, consistent and reliable metadata records for published geospatial data in North Carolina to gain efficiencies in public access and in government data management." As it relates to the Training and Implementation phase of this standard, specific goals include 1) Getting a return on investment for metadata creation and maintenance 2) Understanding the role of metadata in the geospatial data archival processes 3) Developing best practices for integrating metadata creation into workflows 4) Evolving the existing state metadata standard through recommendations from the user community back to the Metadata Committee 5)

Implementation of this standard through training to improve the quality of compliant metadata in state and local government data management and 6) Assessment and evaluation of training activities to determine best practices for future training and support.

This standard is based on the ISO 191** format and is an improvement over prior metadata standards to account for domains and values specific to North Carolina GIS data users and evolving technologies such as remotely sensed imagery, online services and ontologies. These technologies were not considered when these original metadata standards such as the CSDGM (formally known as *FGDC-STD-001-1998)* were first published. Most importantly, it was created by North Carolinians for North Carolinians for use in North Carolina. Because of this, it is important to educate all North Carolina GIS data users on the utility and application of this new GIS standard to ensure that information relating to a data set's creation and administration is captured in a proactive, complete, uniform and timely manner. Under the direction of the NCDOT and NCGICC, it was important to disseminate this metadata standard to the many developers and subsequent users of NCDOT GIS data and beyond through hands-on workshops, training tutorials and continued support of this metadata initiative.

Goals of Project

GIS metadata serves as the means by which spatially-related phenomena can be catalogued within a formal framework. It is here where tacit information can be codified for use by the larger GIS community. Given the ever-increasing size of GIS data sets and the proficiency with which GIS data are created, there needs to be a mechanism to educate, assess and evaluate the human element to keep up with this proficiency. Programming techniques and software packages have allowed users to assess information that would take a human days or perhaps weeks to do. While much of this work serves to provide technical and hands-on support to the dissemination of a new metadata standard, there is a research impetus to this project as well, which includes:

1. Technical and material support to the implementation and education of the new metadata profile to the North Carolina GIS user community
2. Assessing the existing knowledge base of GIS users on the subject of GIS metadata through surveys given to training attendees
3. Understanding needs of GIS data mangers when it comes to metadata population and requirements for data under their purview through surveys given to GIS data managers
4. Determining best methods to deliver metadata training to the North Carolina GIS user community that close gaps between existing knowledge base and needs of GIS data managers and the larger NC GIS community
5. Assessment and evaluation of training activities to determine best practices for future training and future support through post-training quantitative and qualitative surveys

North Carolina Central University

The Department of Environmental, Earth and Geospatial Sciences (DEEGS) at North Carolina Central University (NCCU) performed the work for this project. North Carolina Central University is a Historically Black College and University (HBCU) with an enrollment of more than 8,500 students. NCCU has begun to integrate research with its teaching-based philosophy and has employed faculty and staff that not only excel in teaching, but serve as subject matter experts in their fields and integrate cutting-edge techniques and technologies into their research. The department began preliminary work in August, 2016, with much of the work being performed throughout the grant period until

March 2019. The DEEGS offers undergraduate programs in Environmental Science, Geography and GIS, as well as a graduate degree in Earth Science. The department's mission is to promote intellectual, professional, and personal excellence through the highest quality instruction, research, and service. Its vision is to be recognized as a regional, statewide, and national resource for students and society as well as professionals who work in the many fields that are encompassed by the environmental, earth, and geospatial sciences. The careers goals of recent DEEGS graduates has been a healthy combination of public sector professional work (EPA, State Agencies, City of Durham, etc.), private contractor work and the pursuit of Master's and Ph.D. degrees.

The Research Team

Timothy Mulrooney is an Associate Professor in the Department of Environmental, Earth and Geospatial Sciences (DEEGS) at North Carolina Central University (NCCU). The focus of his teaching and research is GIS and the application of GIS to a variety of disciplines that NCCU offers. Before his tenure at NCCU, he worked as a Senior GIS Analyst with the Army SRP (Sustainable Range Program) GIS Regional Support where he provided GIS analysis, support and database administration for Army assets throughout the world. In this research project, he served as the Principal Investigator and managed every aspect of this project, arranged for training, developed Python code and developed the reporting procedures for the project.

Craig Pederson (left) was a graduate student in the DEEGS at NCCU from 2016 through his graduation in Spring 2018. He worked with the NCDOT to help finalize documentation on the State and Local Government Metadata standard, delivered training and developed data in support of this project. He is currently a Ph.D. student at the University of North Carolina, Greensboro.

Richard Foster (right) is currently an undergraduate student in the DEEGS. After Craig Pederson graduated from NCCU in 2018, he performed Python programming and assisted in workshop delivery. He will be graduating in Fall, 2019.

Literature Review

Metadata serves as an organized means to describe a dataset, and it provides the formal framework for providing information about a dataset's lineage, age and creators. Metadata is composed of both qualitative and quantitative information and while metadata's original use was simply as a means to catalog data, its storage and assessment has become a science in itself.

The FGDC regularly meets to determine all possible values, parameters and domains that can be captured and expressed within the confines of GIS metadata. First formed in the early 1990s, the FGDC serves as a governing body for geospatial data and metadata in the United States. The FGDC defines metadata as the following:

A metadata record is a file of information, usually presented as an XML document, which captures the basic characteristics of a data or information resource. It represents the who, what, when, where, why and how of the resource. Geospatial metadata are used to document geographic digital resources such as Geographic Information System (GIS) files, geospatial databases, and earth imagery. A geospatial metadata record includes core library catalog elements such as Title, Abstract, and Publication Data; geographic elements such as Geographic Extent and Projection Information; and database elements such as Attribute Label Definitions and Attribute Domain Values.

FGDC metadata standards dictate that a plethora of individual entries (now more than 400 and counting) are populated for compliant GIS metadata. Thus, ensuring metadata integrity for large spatial data sets is a time-consuming process if done by hand. It is not uncommon for organizations to employ thousands of individual data layers within their digital warehouses. Since traditional GIS data are ever-evolving, metadata standards must be flexible enough to account for new technologies. Policy should dictate that these standards be revisited periodically to ensure adaptability that can be implemented through large-scale changes or the publishing of new metadata standards. The GIS community has employed a set of content standards to ensure compatibility across the entire GIS community. The updated State and Local Government Metadata profile developed by the NCGICC based on the ISO 191** discussed in this paper highlights this adaptability and is an example of one of these standards.

While regarded as a relatively new concept, both formal spatial and non-spatial metadata has existed in one form or another for the past 50 years. *MARC* (Machine Readable Cataloging) and its successor *MARC 21* are used by the Library of Congress to catalog bibliographic resources. This system has been in place since the 1960s, but it was not originally designed for computer interfacing, and the format is not very intuitive. A more popular format called the *Dublin Core* was created in 1995 for electronic recourses such as web pages and software applications. While the FGDC and GIS metadata standards described here actually predate this more generalized format, GIS metadata data contains a variety of geographically-explicit descriptors that may not be fully understood by the non-GIS community.

Dublin Core and FGDC generally share a base level of descriptive metadata elements. While Dublin Core is used to describe electronic resources and digital representations of physical resources such as artwork, GIS metadata adheres to FGDC and more recently ISO standards. These requirements are always changing as dictated by technology. Because of the spatial nature of GIS data, FGDC requirements dictate that information pertaining to absolute location be retained. These fields include datum, coordinate system, false easting, false northing and bounding coordinates. While Dublin Core does make accommodations for place keywords and spatial descriptors, it does not contain placeholders for elements that help describe geodetic elements associated with the quantitative representation of location with as much detail as GIS metadata.

Because of the different goals of each standard, a precipitous balance between MARC, Dublin Core and FGDC Metadata must be found. Crosswalking, a tedious and sometimes imprecise process where either people or algorithms find matching elements between the different standards may be necessary because various organizations use these popular formats interchangeably on a routine basis. Crosswalking methods have been used to match geospatial data to standards outside of FGDC, such as examining the feasibility of compatibility with the Dublin Core metadata standard.

Current research in the field of metadata can be closely associated with statistics and high-speed processing. Given the exponential increase in electronic resources and media, technologies must be able to accommodate the automation of resources that are viewed, accessed, and assessed. Research examined the role of metadata and its ability to be assessed, arguing that metadata for metadata's sake does no good. Metadata must have some utility as it needs to be assessed and have a role within the decision-making process. Metadata must ultimately serve a purpose and specifically the greater good of the user community. While other research proposes a quality assessment for metadata, it fails to do so with regards to changes in metadata quality, their accompanying values and the holistic structure used to store them. With the standardization of XML-based FGDC and ISO metadata standards, metadata can be compared from one time period to the next. One of these structures is through ontology, a semantic representation of a concept through various domains and properties. Most recently, e-learning technologies were applied to these ontological metadata structures. However, the lack of human cognition within these ontologies cannot eliminate unnecessary or ambiguous terms using results from previous analysis, sometimes referred to as semantic accuracy within the confines of GIS Quality Assessment/Quality Control (QA/QC) circles.

The role of metadata assessment can be seen in a variety of different fields. An Electronic Metadata Record (EMR), for example, is an emerging technology that is produced and edited when an electronic document is edited or created, such as a patient record or digital x-ray. A number of other related technologies for the medical industry have been developed to serve as a quality assurance and administrative tools. The process of accessing, viewing, and commenting on patient files or x-rays by physicians in electronic form can be documented and stored in a metadata file. Hardcopy records are often times time-consuming to complete, and they can be easily lost or destroyed. Thus, the ease of storing, accessing, and retrieving electronic metadata and files for medical data can help prevent litigation against malpractice lawsuits. For example, a complex statistical analysis was developed to retrieve biomedical articles from more than 4,800 journals to help support the decision-making process. It is impossible to scrutinize each of 14 million individual manuscripts. Clustering and classification methods performed on metadata derived from traditional statistical techniques are used to explore and retrieve related information within biomedical literature. If properly maintained, metadata serves as a capable surrogate when querying scanned imagery or hard copy information is not feasible and further validates in-situ decisions as they are reinforced by easily accessible support literature.

Metadata has the flexibility to capture many forms of qualitative and quantitative information stored as numbers, text strings, domain values and dates. However, it does have its drawbacks. In addition to the time, resources and expertise required to populate the information, ancillary concerns exist. Metadata can be applied to any electronic resource, but there are data privacy concerns, especially within the medical community. For example, metadata can be updated and collected to determine the number of times a medical professional has viewed patients' information within the EMR. Not only does this address privacy concerns by documenting access to particular records, but serves to report when, by whom and how long a digital record was viewed. In addition, EMR should not serve as an end-all diagnostic tool, especially when clinical data do exist. Metadata should aid in the evaluation and decision-making process. Other research used image sharing community to further reinforce this point and brought up more excellent points. Metadata for an image (date of image, place, context, etc.) is collected and stored with the image. Furthermore, social metadata not only explores information about the image itself within its place in the social media environment, but also tangential information related to an image such as comments about the image, information about people who have posted comments about the image and the user groups to which these commenters belong. Limiting this information greatly reduces the amount of analysis that can be performed on the accompanying image, decreasing the availability to knowledge in order to make sound business decisions. As this applies to GIS metadata, a happy medium must be found so privacy concerns can be satisfied while dutiful analysis can be performed. Given the relative infancy of these subjects and lack of established doctrine, the body of knowledge is still growing in this subject.

The very nature of spatial data dictates that a different approach must be taken for assessment and reporting within the digital environment. The proliferation of spatial technologies underscores the widely accepted and legitimate role of metadata within the GIS user community. All elements intrinsic to spatial data, such as those associated with position (e.g. latitude, longitude) as well as its representation (e.g., accuracy) must be carefully documented and recorded in GIS metadata. It is important that information about the data format, a description of the data, the processes by which the data were created, the areal extent of the data and the people who aided in data creation be retained. Formal controls may dictate specific tolerances for horizontal and temporal accuracy. This information is not only important from a legal standpoint, but it also validates GIS analysis by speaking directly to such necessary components as its horizontal and temporal accuracy. Since GIS analysis is only as good as the data on which it is based, metadata reinforces the data and ultimately the analysis and organizations which develop the GIS data.

As mentioned previously, metadata is important in helping to document dimensions of quantifiable GIS data quality such as attribute accuracy, horizontal accuracy and attribute completeness. Other forms of GIS data accuracy do in fact exist. FGDC and spatial data transfer standards (SDTS) also consider vertical accuracy (error in measured vs. represented elevation), data lineage (source materials of data) and logical consistency (compliance of qualitative relationships inherent in the data structure) as part of data quality. In some GIS circles, temporal accuracy (age of the data compared to usage date) and semantic accuracy or "the quality with which geographical objects are described in accordance with the selected model" are also considered elements of data quality. Placeholders within FGDC metadata exist to capture all of this information either quantitatively or qualitatively.

Early pioneers of GIS recognized the importance of data quality, not only from a cost efficiency standpoint, but because of the legal ramifications in publishing incorrect spatial information which may lead to accidents or the misuse of data. Even then, they understood the reconciliation between accuracy, the cost of creating accurate data and the eventuality that some error will occur. It is unreasonable to expect an organization such as the North Carolina Department of Transportation (NCDOT) to photo-revise and field check every single road in their GIS database, re-attribute it correctly and then verify them using another party in a timely manner given current personnel and financial constraints. This compromise is referred to as *uncertainty absorption*. Regardless of resource allocation, verification of data quality should be done by discipline experts with a long-term goal of developing data quality standards. This helps to protect the GIS data producer from the potential misuse of GIS data and metadata serves as the means to formally inform the data user of data quality measures applied to data, as well as protect GIS data stewards from its mismanagement.

In and of itself, data quality has no inherent value or worth, but is ultimately realized when action is taken on information pertaining to data quality. Along those same lines, the end goal of information quality is to satisfy customer needs, in this case being the many users who utilize GIS data with the understanding that the data have undergone some form of validation.

Quantitative measures related to this validation with qualitative processes needs to be highlighted in metadata.

Early research and commentary on the concept of geospatial metadata has touted its value as an effective decision-support tool, regardless of its native format [28].
These formats include Hyper Text Markup Language (HTML), Extensible Markup Language (XML) along with its various ISO standards (19115, 19115-1, 19139), TXT (Text File), Geography Markup Language (GML) and Standard Generalized Markup Language (SGML), as well as proprietary formats. Methodology has explored the ability to integrate spatial metadata to a stand-alone database long before GIS metadata was stored in a standardized format, as well as compiling statistics about metadata elements within the confines of specific software.

To that end, the population of geospatial metadata is a monotonous process and subject to error, although research has explored the large-scale production of standards-based metadata in order to alleviate these issues. Because of this, research maintains that human nature alone undermines the immediate and long-term goals of metadata for an organization and the GIS user community. While the omission of one minor element would not degrade a layer's metadata or invalidate the geospatial data on which it is based, it may compromise quantitative data quality measures captured from which decisions can be made. More recently, feature level metadata has been able to capture data quality information, but is typically limited to quantitative measures of positional accuracy and qualitative information related to data lineage within eight of the more than 400 entries that comprise a complete FGDC-compliant metadata file. Even now, the population of these metadata elements is not fully automated and some entries must be done by a GIS data steward. However, methodologies to explore its assessment and evaluation are evolving and this paper explores this notion within the confines of and applied to a particular standard.

Findings and Conclusions

Needs Assessment

In the North Carolina GIS community, GIS is used to represent transportation routes, elevation, delineate land ownership parcels, school attendance, highlight patterns of crime and help make zoning decisions. The manner in which geospatial data is captured varies. Some methods include using a Global Positioning System (GPS) unit, extracting or improving existing GIS data, downloading data from a web site, connecting to a service, the use of an Unmanned Aerial Vehicle (UAV) or some other remote sensing platform, or creating data from an analog format via digitization or georectification. Regardless of the method, the resources (e.g., the computers, time and people dedicated to the process of collecting and creating geospatial data) are the most time-consuming portion of a GIS-related project. As a result, the GIS community needs to ensure the quality of geospatial data created from these methods is captured, stored and assessed in a systematic way.

Geospatial metadata serves as the formal framework to catalog descriptive, administrative and structural information about geospatial data. Geospatial metadata is inherently different from other forms of electronic metadata because each metadata file can be applied a spatial component that is not implicit with other forms of metadata. These spatial components encompass a wide array of information to include the date, methods and sources by which geospatial information was captured, means to ensure that the geospatial information adheres to acceptable standards and/or aligns with other geospatial datasets to ensure seamless analysis, projection information of the dataset and bounding coordinates of the dataset. All of these entries, in addition to the data's non-spatial components can be queried within the confines of geospatial data portal such as one found at North Carolina OneMap, the geospatial data portal for the state of North Carolina.

Given the capricious rate at which all forms of geo-information can be created, formal metadata serves as a lifeline between the tacit knowledge of the data creator and current and future generations of geospatial data consumers. In the United States, the Federal Geographic Data Committee (FGDC) metadata standard, commonly referred to as the Content Standard for Digital Geospatial Metadata (CSDGM) allows for more than 400 individual metadata elements. The North Carolina GIS community has been proactive about understanding the importance of metadata.

Maintaining a complete and comprehensive metadata record is a continual and interactive process. GIS metadata is one of the most overlooked and underappreciated aspects of any GIS enterprise or project. If time or resources need to be sacrificed in the course of a project, it is usually at the expense of metadata. Information is key to an organization's vitality, sustainability and success. Metadata should be treated as an investment. Maps and analysis are only as good as the data on which they are based. Metadata is a direct reflection of this investment and the organization which makes this investment. Metadata captures important information related to data creators, data quality and the various accuracies (horizontal, vertical, temporal, attribute, semantic, etc.) with which we can quantitatively measure GIS data. These measurements help guide the decision-making process, especially in larger (hundreds of layers) spatial databases. Not only is good metadata a wise business practice, but saves time, money and resources in the long run. Unfortunately, metadata's true value is not realized until it is absent, and few studies have been done to place a direct monetary value on metadata.

As part of a needs assessment for this project, a survey was developed to help dictate and direct metadata needs within the state of North Carolina and to give the NCGICC and good idea about metadata about metadata knowledge in the state. This survey was developed in 2017 and distributed to the North Carolina GIS user community. Forty (40) respondents answered the survey, who ranged from GIS Technicians and Property Mappers to GIS Coordinators and Managers throughout the state of North Carolina. Questions were asked about respondents' experience with data development, metadata, as well as organizational requirements as it pertains to metadata.

Most prominent was the schism between respondents' experience with data development and experience with metadata, as shown in Figures 1 and 2. Respondents generally had an 'excellent working knowledge' of data development, but only 'some experience' or 'working knowledge' on the metadata created as a result of these data development techniques. These underscore technical experience in creating new data in support of analysis and projects, but less experience in cataloging these same data used for analysis and maps.

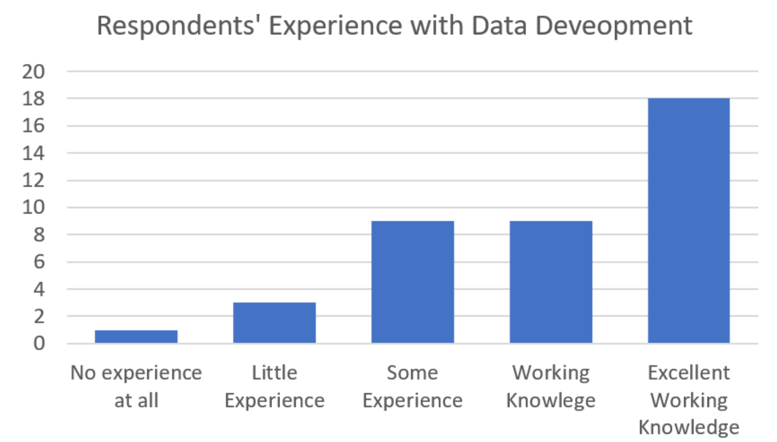

Figure 1: A survey of GIS professionals and experience with data development.

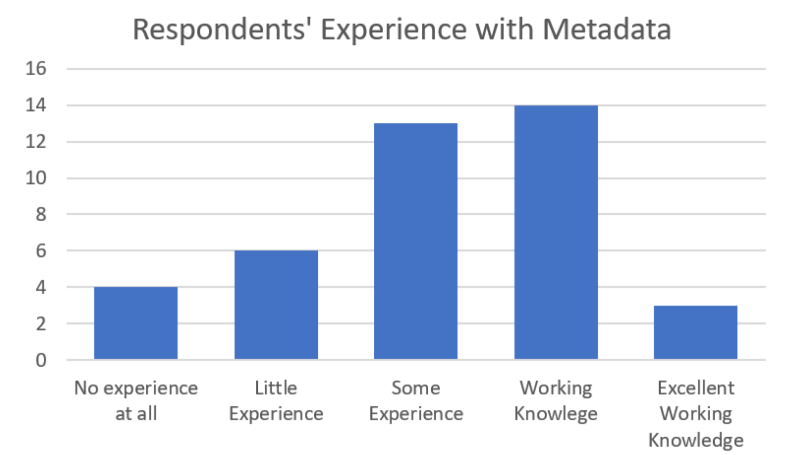

Figure 2: A survey of GIS professionals and experience with metadata.

Not only is this schism evidenced at the individual level, but also the organizational level. In a same survey of these 40 GIS professionals, they describe their organization's approach to metadata as shown in Figure 3. More than half of all respondents' organizations have no metadata requirement whatsoever and only five respondents work in an organization that has a firm metadata requirement. The rest do have a metadata requirement, but it is not upheld.

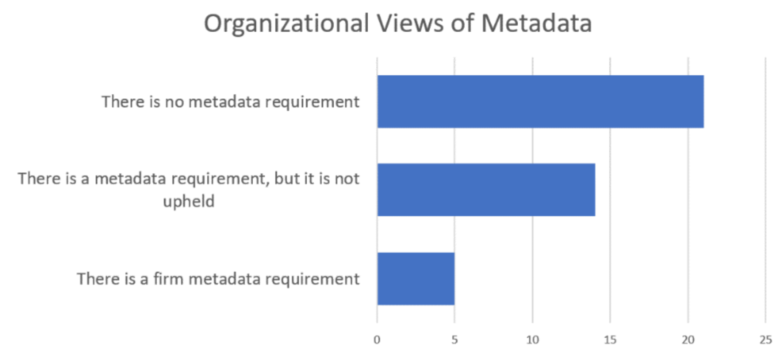

Figure 3: A survey of metadata requirements at the organization level.

The North Carolina State and Local Government Profile

Geospatial metadata standards serve as a cohesive and standardized means by which organizations can define, store and more importantly share information about geospatial data. It defines the categories of information that needs to be stored, individual entries, or tags, of individual elements within these categories and the types of data (text, date, number) and their lengths that can be stored while representing these tags. FGDC metadata is divided into 7 sections or divisions that transcend descriptive, administrative and structural components. They are: Identification Information, Data Quality Information, Spatial Data Organization Information, Spatial Reference Information, Entity and Attribute Information, Distribution Information, and Metadata Reference Information.

Within these high-level divisions, subdivisions and eventually individual metadata tags can be populated to catalog various forms of information about the GIS data layer. The hierarchy of these divisions and subdivisions are consistent with a standard. In addition to providing this structure, the FGDC also creates guidelines by dictating which metadata elements are to be populated. The FGDC requires seven metadata elements be populated for all GIS data. The FGDC also suggests that fifteen metadata elements be populated. These suggested and required elements are included in Table 1 below.

Table 1: Required and Suggested FGDC Elements

FGDC -Required Elements	***FGDC- Suggested Elements***	
Title	Dataset Responsible Party	Lineage Statement
Reference Date	Geography Locations by Coordinates (X and Y)	Online Resource
Language	Data Character Set	Metadata File Identifier
Topic Category	Spatial Resolution	Metadata Standard Name
Abstract	Distribution Format	Metadata Standard Version
Point of Contact	Spatial Representation Type	Metadata Language
Metadata Date	Reference System	
	Metadata Character Set	

Organizations actively create content standards for new technologies and manners in which geospatial data are collected and stored. One such example is the FGDC content standard for Remotely Sensed Data. This includes two divisions germane to the equipment and methods such as platform name, sensor information and algorithm information used to capture the imagery, in addition to the seven existing aforementioned divisions. In order to further elucidate descriptive, administration and structural information, additional addendums to existing metadata standards are also attached to specific geospatial-specific data such as addresses, biological data, shoreline data, and vegetation data. Standards such as these must be increasingly flexible and updatable to account for the evolving technologies in which geospatial data can now be captured (crowdsourcing, Unmanned Aerial Vehicle, large scale geocoding), processed (new geostatistical and interpolation algorithms) and ultimately delivered (web map service, web feature service) to the GIS user community.

In recent years, the North Carolina SMAC has recognized most GIS data managers lack the time and resources necessary to learn and apply a metadata standard that maintains dataset integrity and retains pertinent information for while not being too demanding on existing resources, most notably time and people. To address the problem of missing or incomplete metadata records among state and local data publishers, the SMAC chartered an ad-hoc Metadata Committee in October 2012 to "recommend ways to expand and improve geospatial metadata in North Carolina that are efficient for the data producer and benefit data users in the discovery and application of geospatial data." The Metadata Committee submitted a draft of this profile, based on the ISO 19115 (for Geographic Information – Metadata: 2003), ISO 19115-1 (for Geographic Information – Metadata – Part 1: Fundamentals: 2014) and ISO 19119 (Geographic Information – Services: 2016) standards. After review and modification by SMAC and its standing committees, the most current version of this standard has been in effect since December 30, 2016 and is available through the NCOneMap portal. While not entirely ground-breaking, North Carolina has been a forerunner in developing sub-country metadata standards. The SMAC worked with the Canadian Province of Alberta, who has already developed a standard germane to their province while states such as Missouri and Virginia have developed some level of uniform metadata available with their products.

Given seven required and fifteen recommended metadata elements are fairly ambiguous and less than ideal for many organizations whose data is integrated into the NCOneMap, the North Carolina state geospatial data portal, this profile provides explicit guidance on required/suggested metadata elements, wording for these elements, standardization of naming/date conventions and domain fields for topic categories for more than 75 metadata tags. A few examples of the rules for geospatial metadata include:

1. Title is required as a free-text entry.
2. Publication Date is required and the format for Publication Date is YYYY-MM-DD or YYYYMMDD. If day is not known, use YYYY-MM and use YYYY if month is not known.
3. Abstract is required as a free text entry. Do not use YYYYMM since it can't be distinguished from the incorrect, but still used YYMMDD.
4. Status is required and only possible values are 'historicalArchive', 'required', 'planned', 'onGoing' 'completed', 'underDevelopment' and 'obsolete'.
5. Topic Category is required and can be one of 23 possible values from domain table.
6. Use Constraints required as a free-text entry to describe any restrictions with using the data.
7. Online linkage is required to an URL address that provides access, preferably direct access, to the data.

In addition, given their nature and distinct differences between their geospatial data counterparts, the SMAC has defined rules for geospatial services to include the following:

1. Metadata Scope code must be 'service'.
2. Online Function code is required from domain of one of five possible values.
3. Title is required as a free-text entry.
4. Metadata Contact is required as a free-text entry, representing Organization Name of the agency that serves as a point of contact for the metadata record.

Description of Training Offered

Through this grant metadata training and presentations on the standard was offered on a number of occasions. This training and support includes:

- Metadata Training for State Government Data Managers, December 2016.
- North Carolina GIS Conference, "*Jumpstart Metadata Creation with the NC Metadata Profile*", February 2017.
- Face-to-Face Metadata Training at North Carolina Central University, December 2017.
- North Carolina State University Geospatial Studio Lecture Series, "*Jumpstarting Metadata Creation*", October 2017.
- Southeast Division of Association of American Geographers Annual Meeting, "*Application of a New GIS Metadata Standard in the State of North Carolina",* November 2017.
- North Carolina Geographical Society Annual Meeting, "*Best Practices of the New North Carolina State and Local Government GIS Metadata Profile", November 2017.*
- GIS-Transportation Conference, "*Application of the North Carolina State and Local Government GIS Metadata Profile",* March 2018.
- GeoProcessing 2018 Conference, "*Assessing and Evaluating Standard Compliance with a State and Local Government GIS Metadata Profile in Large Geospatial Databases*", March 2018.
- Face-to-Face Metadata Training at North Carolina Central University, October 2018.
- North Carolina GIS Conference, "*Facilitating the New Statewide GIS Metadata Standard Through Training and Outreach*", February 2019.

Figure 4: Face-to-face metadata training given at North Carolina Central University in 2018.

Data Development

Templates were developed for agency stakeholders such as state-owned buildings, roads, parcels and regulated facilities. The research team helped to develop data for these sample templates. The research team will work with the Metadata Committee to determine the best layers for data development. Data were also developed for virtual and face-to-face training purposes. These data layers included:

File Name	**Title**
Admin	World Administrative Divisions
Areacode	U.S. Telephone Area Code Boundaries
Counties	U.S. Counties (Generalized)
Institute	U.S. Institutions
	U.S. National Transportation Atlas
Intrstat	Interstate Highways
nc_camden_parcels_poly_2016_05_25	nc_camden_parcels_poly_2016_05_25
Railroads	World Railroads
Timezone	World Time Zones
spot_elevation_point	spot_elevation_point
Colluniv	Colluniv
corinst	Corinst
swlg	Swlg
schpl	Schpl

Facilitation of Online Resources

In addition to videos posted in support of his classroom instruction, the PI hosts YouTube tutorials (http://youtube.com/DEEGSNCCU) ranging from metadata basics and the use and application of the North Carolina State and Local Government Profile to more advanced topics such as XML translators and Python programming solutions in metadata assessment and evaluation. These tutorials have been utilized more than 3,000 times.

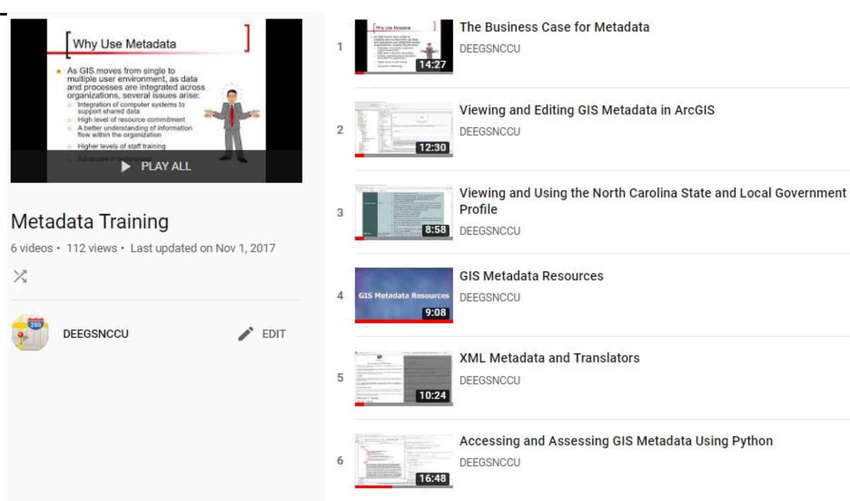

Figure 4: Screenshot from Metadata Training Tutorials on DEEGS YouTube page.

Development and Assessment of Metadata Templates

Given the ever-increasing size of GIS data sets and the metadata requirements for each data layer, there needs to be a mechanism to assess the quality of these metadata not seen in previous generations or documented in existing literature. There also needs to be a means by which individual metadata entries adhere to predefined profiles and standards. This is in support of Task 1 of the research tasks. Computer programming languages and templates have helped to streamline this process. Templates populate redundant features that are common throughout an entire GIS database such as the purpose, supplementary information, distribution liability statements and ordering instructions that can be specific to an agency or department. These templates can be imported one at a time, but programming techniques and software packages have allowed users to assess information that would take a human days or perhaps weeks to do. The NCGICC provides a number of templates through their web portal, NCOneMap. The themes for these templates are at the request of North Carolina GIS users, and include buildings, cadastre, municipal boundaries, school attendance districts, street centerlines, address points and orthoimagery. These templates contain much of the verbiage about a layer's description and creation processes, and can be easily imported and edited specific to the user's contact information. A sample of the identification information for the cadastral data template can be seen in Figure 5.

```
  Description:
    Abstract: Digital cadastral data describing the parcel dataset of ZZZZZ County,
  North Carolina.  This dataset includes attributes such as parcel boundaries,
  ownership, acreage, source references, assessments, and other core cadastral
  attributes.
    Purpose: This dataset was generated to publish a geospatial inventory of real
  property in ZZZZZ County, NC.  It serves to support and assist governmental agencies
  and others in resource management decisions. Additionally, these data provide a set
  of core attributes defined by North Carolina Content Elements for Statewide
  Publication of Core Geospatial Parcel Data and is used to facilitate the sharing,
  display, and use of cadastral data across the state.  The core data are intended to
  integrate local parcel information on a statewide level and eventually build a
  seamless parcel map for North Carolina.
    Supplemental_Information: This dataset is prepared for the inventory of real
  property found within this jurisdiction and is compiled from recorded deeds, plats,
  and other public records and data.  Assessed value of real property is based on
  market value as defined in the Machinery Act of North Carolina (Section 105-283). The
  latest assessment was completed in 88888.  Users of this dataset are hereby notified
  that the aforementioned public primary information sources should be consulted for
  variation of the information contained in thes data.  ZZZZZ County assumes no legal
  responsibility for the information contained in these data.
   Time_Period_of_Content:
    Time_Period_Information:
     Single_Date/Time:
      Calendar_Date: 20051207
```

Figure 5: Identification information metadata entries for cadastral template provided through NCOneMap web portal.

Development of Metadata Scripts

Open source programming solutions using Perl and R have been used to assess and evaluate metadata by traversing geospatial metadata stored in XML format as per FGDC requirements, resulting in quantitative metrics, graphs and reports regarding metadata compliance, as shown in Figure 6.

FGDC Compliancy Report

File Name	Layer Name	Required FGDC Features	Suggested FGDC Features	Missing Features
./control_point.xml	Monumented Benchmarks, BG Thomas Baker Training Site (Lil Aaron Strauss)		14	Metadata Standard Version
./elevations.xml	20 Meter Elevation Contour Line, Fort Knox	6	14	Metadata POC, Responsible Party
./extent.xml	Map Extent, Fort Knox			NONE

9 out of 12 layers (75.00%) had all of the FGDC Required metadata components
81 out of 84 individual FGDC required elements (96.43%) were adequately populated

7 out of 12 layers (58.33%) had all of the FGDC Suggested metadata components
175 out of 180 individual FGDC required elements (97.22%) were adequately populated

Figure 6: Sample of Metadata Compliance Report Generated Using Open Source Assessment Tool.

As applied to the NC State and Local Government Profile, one major challenge exists. Primarily, geospatial data and metadata is typically software specific. While optimal open source solutions could be used to gleam information from metadata stored in XML using an appropriate xPath, these software-agnostic solutions are typically loosely-coupled and not intuitive to the average user. As a result of reliance on Esri products throughout the state, the Python programming language is being used to run this iteration of an assessment and evaluation tool before open source solutions are explored.

Using the NC State and Local Government Profile as a guideline, the research team has been developing tools for data managers to access and evaluate metadata entries. At the current time, metadata entries are written to CSV (Comma Separated Values) files. While doing this, string operations are run to ensure that required entries are populated, date entries comply with required conventions and domain entries match those in the domain table, all while agglomerating results and statistics at the database, layer (record) and tag (attribute) level. This can provide GIS managers with insight on non-compliant metadata entries to determine relationships between non-compliant entries and the responsible data steward or particular attributes that are continually non-compliant.

While QC procedures need to be performed to determine if a metadata entry is accurate, below are a few examples of the many programming rules employed to determine if entries are populated properly.

1. *Title*, *Responsible Party Organization Name*, *Online Linkage*, *Abstract*, *Use Constraints*, *Feature Catalogues*, *Process Description*, *Spatial Reference Information* and *Metadata Contact Name* cannot be Null
2. *Data Type* can only have values of ‘creation’, ‘publication’, or ‘revision’.
3. *Publication Date*, *Temporal Extent of Data* and *Metadata Creation Date* must follow appropriate format. This entails:
   a. The date cannot be Null and must be populated.
   b. The date can only have a length of 10 (YYYY-MM-DD) 8 (YYYYMMDD), 7 (YYYY-MM) or 4 characters (YYYY).
   c. Besides the hyphens (‘-‘), the date can only contain numbers whose value range from 0 through 9. Letters and other characters are not allowed.
   d. If a date contains hyphens (‘-‘), there will be 2 hyphens in a string that has a length of 10 (YYYY-MM-YY) and there will be 1 hyphen in a string that has a length of 7 (YYYY-MM).
   e. Regardless of the format used, the first character of a string will either be ‘1’ or ‘2’ since the year of publication or creation will begin in only those 2 numbers.
4. *Metadata Contact Role Code* can only have values of ‘custodian’ or ‘pointOfContact’.

5. *Progress Code* can only have values of ‘completed’. ‘historicalArchive’, ‘obslete’, ‘onGoing’, ‘planned’, ‘required’ or ‘underDevelopment’.
6. *Maintenance and Update Frequency* can only have values of ‘continual’, ‘daily’, ‘weekly’, ‘fortnightly’, ‘monthly’, ‘quarterly’, ‘ biannually’, ‘annually’, ‘asNeeded’, ‘irregular’, ‘notPlanned’ or ‘unknown’.

The current application has a basic GUI (Figure 7) that allows for 4 input parameters: 1) an input database that contains the features classes for which metadata will be checked 2) an output folder to which XML metadata is written. Python cannot directly access metadata in geodatabase format, so this proprietary metadata is converted to XML format and traversed. Options exist so these XML files are immediately deleted. 3) The name and location of the output entries that stores all metadata entries that are checked from the North Carolina State and Local Government Profile, as well as a summary of the percentage of individual metadata elements that are correct and a summary of percentage of correct elements on a feature class by feature class basis and 4) the location of an error file that highlights errors within the metadata (Figure 8).

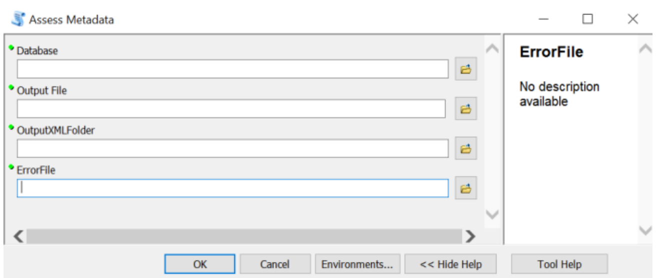

Figure 7: Metadata Assessment and Evaluation Tool.

```
The Ending Date for the Temporal Extent is missing from feature class:  areacode.
The Feature Catalogue is missing from feature class:  areacode.
The Map Projection is missing from feature class:  areacode.
The Online Linkage is missing from feature class:  counties.
The Beginning Date for the Temporal Extent is missing from feature class:  counties.
The Ending Date for the Temporal Extent is missing from feature class:  counties.
The Feature Catalogue is missing from feature class:  counties.
The Map Projection is missing from feature class:  counties.
The Online Linkage is missing from feature class:  institut.
The Publication Date in not in the proper format from feature class:  institut.
The Beginning Date for the Temporal Extent is missing from feature class:  institut.
The Ending Date for the Temporal Extent is missing from feature class:  institut.
The Feature Catalogue is missing from feature class:  institut.
The Map Projection is missing from feature class:  institut.
The Online Linkage is missing from feature class:  intrstat.
The Beginning Date for the Temporal Extent is missing from feature class:  intrstat.
```

Figure 8: Sample Error File Output from Metadata Assessment and Evaluation

While merely developing these standards does have some utility, implementing, disseminating and updating this useful information to a larger user community becomes problematic. While geospatial managers are familiar with these standards, that may not be enough. In many scenarios, especially in larger city and/or county governments, GIS technicians, analysts or independent contractors focus on the regular update and maintenance of metadata. Given the fickle nature of the human component, metadata entries can vary greatly from person to person and agency to agency if little guidance is provided on its proper population. As a result, a network to educate GIS users on the use of these standards to various applications and software is required. This research proposal was created due to the need for high-quality and standardized metadata training throughout the state as dictated by the Metadata Committee. The head of this committee, Sarah Wray, currently works as the Spatial Data Manager for the NCDIT in Raleigh.

Secondly, given the PI's existing burgeoning relationships with the NCGICC, state/federal agencies and private contractors on this ad hoc committee, news about these opportunities would spread quickly throughout the state and beyond. In addition to solidifying existing relationships, we foresee many new collaborations and connections as a result of this research, both locally and globally. The director of GeoDiscover Alberta (Canada) serves as a technical advisor to the ad hoc committee. It is anticipated that many more would be enthusiastic about endorsing training and subsequent activities related to this standard. This may include out-of-state agencies adopting our existing standard, increased viewership of online training videos or steering committee members serving as technical advisors as other agencies begin adoption of this new standard catered to their needs based on this training.

Recommendations to NCDOT

Despite this project formally ending, online and face-to-face training can still be provided by North Carolina Central University. NCCU has the resources in place to offer training on site and the digital infrastructure to travel to provide training on site. Please contact the PI about offering metadata training at your organization.

While a powerful and efficient tool, the programmatic assessment and evaluation of geospatial metadata still cannot altogether replace the human component. While these technologies can traverse metadata schema and extract tags to deem if they are complete, compliant or belong to a particular domain, it does not necessarily mean they are correct. For example, while the *Publication Date* tag may be properly populated (2016-02-29 for example) as per the rules dictated in the North Carolina State and Local Government Profile, it may not necessarily mean the data were published on that date. QA/QC techniques should be used to determine metadata quality across the entire dataset via ANSI (American National Standards Institute), ANSQ (American Society of Quality Control) or other institution-wide QA/QC procedures that best fit needs, resources and limitations.

While the level of attribution within metadata has improved with each new standard and this particular profile, it is in no way complete. As technologies improve and there include more diverse ways to collect, manipulate and create GIS data, metadata must be flexible enough to accommodate all of these techniques. For example, the standard CSDGM does not contain placeholders germane to the collection of data created via a GPS unit like the various DOP (Dilution of Position) measures such as vertical, horizontal and 3D. In addition, detailed information directly associated with the quality of data specific to GPS-collected data such as ephemeris can be entered via a free text field, but lacks the placeholders within the CSDGM as well as this standard. In addition, GIS data now extend well beyond the typical raster and data models that a GIS professional may have solely encountered only a decade earlier. GIS data may now include stand-along tables, Triangulated Irregular Networks (TINs), relationship classes and even topologies. They each have their own intrinsic qualities that make their creation and update difficult to encapsulate within a single catch-all metadata format. Stakeholders must be receptive to these various data formats and the Metadata committee should regularly convene to determine how, if and when new parameters must be required of the standard.

Implementation and Technology Transfer Plan

Who at NCDOT will use the Research Product

Anyone who uses NCDOT GIS data will hopefully see value in these research products. Particularly, those who are involved in GIS data development and the creation of GIS data (GIS Technicians, GIS Analysts) would see the greatest benefit this training in order to standardize the cataloguing of this data development. From a management perspective, managers can make better decisions as to where, how and when to allocated resources within their geospatial enterprise. While this work does not address the protocol that go into these different data development scenarios, this training ensured that these protocols are adequately catalogued for future users of the data.

How NCDOT will use the Research Product

Results of this research project hopefully provided consistency and uniformity across all NCDOT/NCDIT divisions and beyond who utilize NCDOT/NCDIT GIS assets. The NCDOT has working relationships with a variety of organizations to include the NCGICC, USDA (United States Department of Agriculture), Department of Fish and Wildlife, Department of Natural Resources and Wake County. This training and research would allow all of the organizations to work more efficiently with the NCDOT/NCDIT. An evolving set of templates, tools, training and online support based on input from all parties has been developed that best cater to ever-changing needs of the NCDOT GIS community.

Proposed Impact of Research Product

While merely developing these standards does have some utility, implementing, disseminating and updating this useful information to a larger user community becomes problematic. While geospatial managers are familiar with these standards, that may not be enough. In many scenarios, especially in larger city and/or county governments, GIS technicians, analysts or independent contractors focus on the regular update and maintenance of metadata. Given the fickle nature of the human component, metadata entries can vary greatly from person to person and agency to agency if little guidance is provided on its proper population. As a result, a network to educate GIS users on the use of these standards to various applications and software is required. This research proposal was created due to the need for high-quality and standardized metadata training throughout the state as dictated by the Metadata Committee. The head of this committee, Sarah Wray, currently works as the Spatial Data Manager for the NCDIT in Raleigh.

Secondly, given the PI's existing burgeoning relationships with the NCGICC, state/federal agencies and private contractors on this ad hoc committee, news about these opportunities would spread quickly throughout the state and beyond. In addition to solidifying existing relationships, we foresee many new collaborations and connections as a result of this research, both locally and globally. The director of GeoDiscover Alberta (Canada) serves as a technical advisor to the ad hoc committee. It is anticipated that many more would be enthusiastic about endorsing training and subsequent activities related to this standard. This may include out-of-state agencies adopting our existing standard or steering committee members serving as technical advisors as other agencies begin adoption of this new standard catered to their needs based on this training.

Cited References

Aronoff, Stan. 1989. *Geographic Information Systems: A Management Perspective*. Ottawa: WDL Publications.

Batcheller, J. 2008. Automating Geospatial Metadata Generation—An Integrated Data Management and Documentation Approach. Computers & Geosciences 34: 387 – 398.

Bedard, Y. 1987. Uncertainties in land information systems databases. *Proceedings of the 8th International Symposium on Computer Assisted Cartography (Auto Carto 8)*, (pp. 175-184). Baltimore, MD.

Bruce, T. and Hillman, D. 2004. The Continuum of Metadata Quality: Defining, Expressing, Exploiting. In: D. Hillman and E. Westbrooks, Editors, *Metadata in Practice*, ALA Editions, Chicago, pp. 238–256.

Chang, Kang-Tsung. 2008. *Introduction to Geographic Information Systems*. New York, NY: McGraw Hill Higher Education.

Dalcin, E.C. 2004. *Data quality concepts and techniques applied to taxonomic databases.* Thesis for the degree of Doctor of Philosophy, School of Biological Sciences, Faculty of Medicine, Health and Life Sciences, University of Southampton.

Department of Environmental, Earth and Geospaital Sciences YouTube Portal. *Metadata Training* [online]. Available from https://www.youtube.com/playlist?list=PL0JtGFJGnWTONvMk6dBOXqIN6pTlmoj5o. [retrieved Novmber 2018].

Dublin Core Metadata Initiative. *Dublin Core Metadata Element Set, Version 1.1: Reference Description*. [online], available from http://dublincore.org/documents/dces/ [retrieved September 2018].

English, L.P. 1999. *Improving data warehouse and business information quality: Methods for reducing costs and increasing profits*. New York: John Wiley and Sons.

Epstein, E. F. 1988. Litigation over information: The use and misuse of maps. *Proceedings, IGIS: The Research Agenda* 1 (pp. 177-184). Washington, D.C.: NASA.

FGDC (Federal Geographic Data Committee). 2000. *Content Standard for Digital Geospatial Metadata Workbook*. Washington D.C.: Federal Geographic Data Committee.

Greenberg, J., Spurgin, K. and Crystal, A. Functionalities for Automatic Metadata Generation Applications: A Survey of Metadata Experts' Opinions. International Journal of Metadata, Semantics and Ontologies 1(1): 3 – 21.

Giuliani, G., Guigoz, Y., Lacroix, P., Ray, N and Lehmann, A. 2016. Facilitating the production of ISO-compliant metadata of geospaital datasets. *International Journal of Applied Earth Observation and Geoinformation* 44: 233-243.

Hass, Stephanie, E. Henjum, M. O'Daniel and J. Aufmuth. 2003. Darwin and MARC: A Voyage of Metadata Discovery. *Library Collections, Acquisitions, & Technical Services* 27: 291 – 304.

Lanter, D. 1993. A Lineage Meta-Database Approach Towards Spatial Analytic Database Optimization. Cartography and Geographic Information Systems 20(2): 112-121.

Leiden, K., Laughery, K.R., Keller, J., French, J., Warwick, W., Wood, S.D.. 2001. *A Review of Human Performance Models for the Prediction of Human Error.* Moffett Field, CA : National Aeronautics and Space Administration.

Library of Congress. 2008. *MARC Standards*. Library of Congress Network Development and MARC Standards [website]. http://www.loc.gov/marc/ (last accessed 26 October 2015).

Limbach, T., Krawczyk, A., Surowiec, G. 2004. Metadata Lifecycle Management with GIS context. Proceedings of the 10th EC GI & GIS Workshop, ESDI State of the Art. Warsaw, Poland.

McLean, T., Burton, L., Haller, C. and McLean, P. 2008. Electronic Medical Record Metadata: Uses and Liability. *Journal of the American College of Surgeons* 206(3): 405 – 411.

Mulrooney, Timothy. 2009. *Turning Data into Information: Assessing and Reporting GIS Metadata Integrity Using Integrated Computing Technologies (PhD Dissertation).* Greensboro, North Carolina: University of North Carolina, Greensboro.

North Carolina One Map, *Geographic Data Serving a Statewide Community* [online]. Available from http://www.nconemap.gov/ [retrieved October 2018]

North Carolina Geographic Information Coordinating Council (NCGICC), *North Carolina State and Local Government Metadata Profile for Geospatial Data and Services* [online]. Available from http://www.nconemap.gov/DiscoverGetData/Metadata.aspx#iso. [retrieved February 2018]

North Carolina Geographic Information Coordinating Council (NCGICC). *North Carolina OneMap* [online]. Available from http://www.nconemap.gov. [retrieved February 2018].

Qi, L., Lingling, G., Feng, H., Yong, T. 2004. A Unified Metadata Information Management Framework For the Digital City. *Proceedings of IEEE's Geoscience and Remote Sensing Symposium*, Anchorage, Alaska, USA: 4422–4424.

Salgé, F. 1995. Semantic accuracy. In S.C. Guptill and J.L. Morrison *(*Eds.), *Elements of Spatial Data Quality (*pp. 139-152). New York: Elsevier Science Ltd.

Stvilia, B. and Gasser, L. Value-Based Metadata Quality Assessment. *Library & Information Science Research* 30(1): 67 – 74.

USGS. 1997. *Spatial Data Transfer Standard (SDTS): Logical specifications*. Reston, Virginia: United States Geologic Survey.