interence intestation inflation inflicted influences informal information info

Metadata Quality Assurance: The University of North Texas Libraries Experience

Daniel Gelaw Alemneh & Hannah Tarver

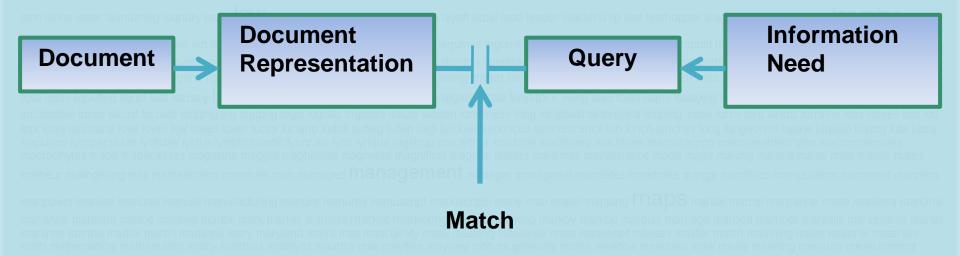
3rd annual Texas Conference on Digital Libraries (TCDL)

May 27-28, 2009

iunicipal

Inference infestation inflation inflicted influence influences informal informatics INTORMATION informed infrastructure ingersoll inheritance inhibitors initial initiative injured injurial information informati

Information Retrieval



Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review, 13*(5), 407-424.

murder murderers muriel murmansk muscle muscles muscular museum museums mushroom mushrooms MUSIC musical musici musicians musique mussolini musto mutants mutation mutton myers myosin myriophyllum mystery myth mythology myths mêlée nae naji named nanking nanocomposite nanocomposites nanocrystals nanoindentation



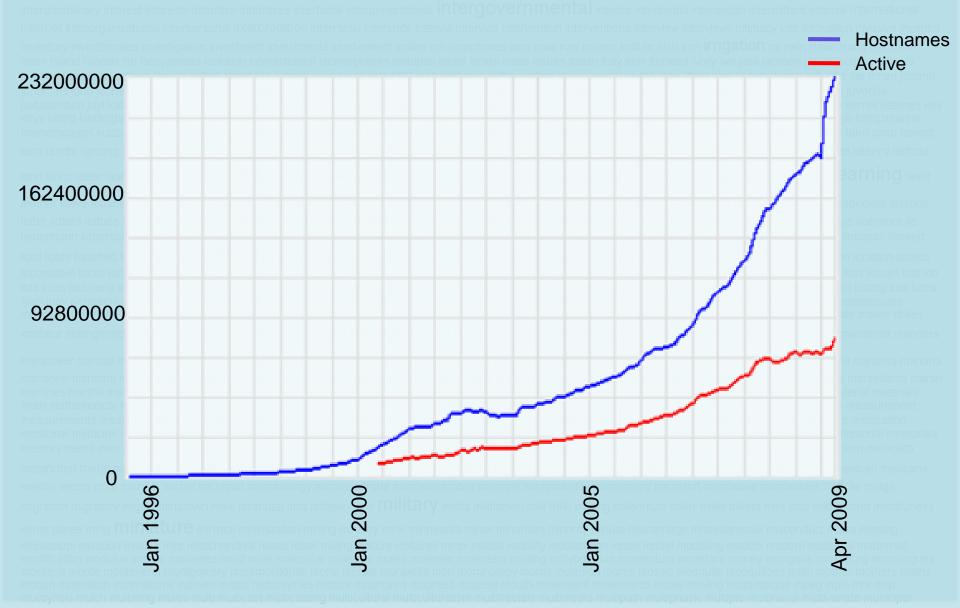
Trends



- Information creation, organization, retrieval, use, and preservation is becoming more complex
- User as creator, annotator, indexer, searcher, and eventual user of his/her content
- Visualization of the information space instead of a ranked list of search results

nuczynski mulch mulching mules multi multicast multicasting multicultural multiculturalism multifractals multimedia multipath multiphasic multiphasic multipate multivariate municipal

Total Sites Across All Domains August 1995-April 2009



UNIVERSITY OF NORTH TEXAS

LIBRARIES

DISCOVER THE POWER OF IDEAS

Digital Projects

- UNT Digital Collections
- Portal to Texas History
 - 100 + Collaborators
- Congressional Research Service Archives
- Other Statewide and National Projects

nuczynski mulch mulching mules multi multicast multicasting multicultural multiculturalism multifractals multimedia multipath multipath multiple multipacial multivariate municipa

murder murderers muriel murmansk muscle muscles muscular museum museums mushroom mushrooms MUSIC musical musici musicians musique mussolini musto mutan mutation mutton myers myosin myriophyllum mystery myth mythology myths mêlée nae naji named nanking nanocomposite nanocomposites nanocrystals nanoindentation

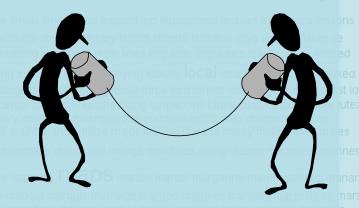
Factors Influencing Metadata Quality

Local Requirements:

- Objects
- Granularity
- Functionality

Collaborative Requirements:

- Diversity of Users
- Interoperability
- Digital Rights Issues



inference infestation inflation inflicted influence influences informal informatics information informed infrastructure ingersoll inheritance inhibitors initial initiative injured injuries

Poor Metadata Quality

Ambiguities

Poor recall

Poor precision

Inconsistency of search results



inference infestation inflation inflicted influence influences informal informatics **information** informed infrastructure ingersoll inheritance inhibitors initial initiative injured injuries injurious injury ink inlaid inlay inmates innocence innovation innovations inoculation inorganic inquiry insect insecticides insects insight insighia insomnia inspection installation

Common Errors

The data is:

- Incorrect
- Missing
- Ambiguous



murder murderers muriel murmansk muscle muscles muscular museum museums mushroom mushrooms MUSIC musical musici musicians musique mussolini musto mutants

Metadata Quality Assurance Mechanisms & Tools at UNT

Pre-Ingest

Post-Ingest

Training

Creation Tools

Proofing & Editing Tools

Analysis Tools

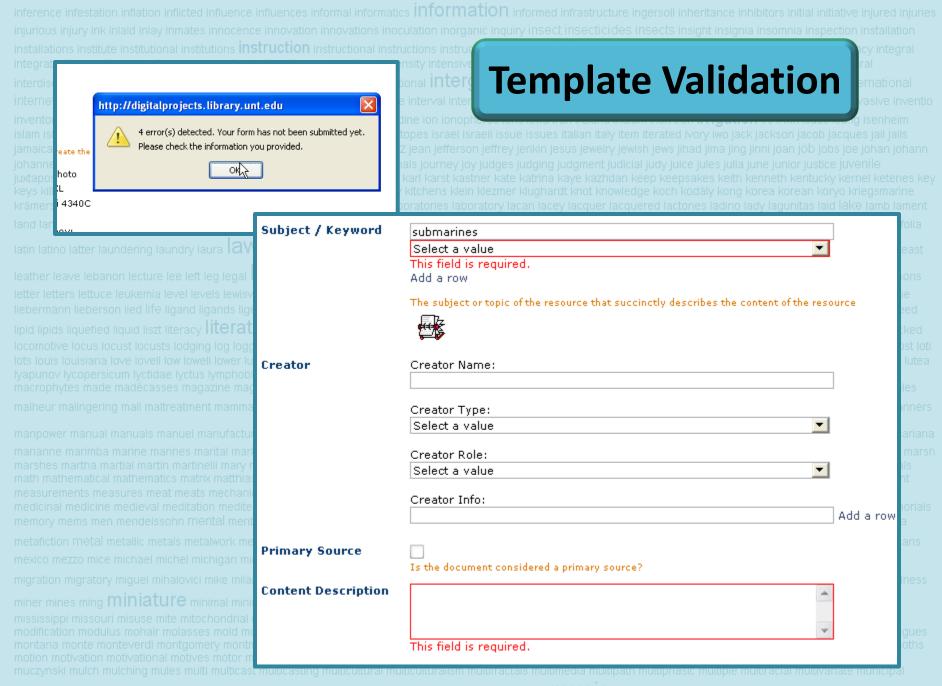
murder murderers muriel murmansk muscle muscles muscular museum museums mushroom mushrooms MUSİC musical musici musicians musique mussolini musto mutants.

Training Face-to-Face Instruction Metadata Schema & Documentation Internal Project Wikis Staff Support

murder murderers muriel murmansk muscle muscles muscular museum museums mushroom mushrooms MUSIC musical musici musicians musique mussolini musto mutants mutation mutton myers myosin myriophyllum mystery myth mythology myths mêlée nae naji named nanking nanocomposite nanocomposites nanocrystals nanoindentation

Title Add a row Select a value Metadata The official name given to the resource Identifier Creation Add a row Select a value A unique identifier or 'permanent name' for a resource **Template** Note ctones ladino lady lagurillas lalu lake lamb lament This element will serve as a catch-all note field Institution Add a row Select a value A consistent reference to the institution or administrating unit that owns the digital resource for which metadata was created Collection Add a row Select a value Collection title Subject / Keyword Add a row Select a value The subject or topic of the resource that succinctly describes the content of the resource Creator Creator Name: Creator Type: ath mathema ▾ Select a value Creator Role: ▼| Select a value Creator Info: Add a row **Primary Source** Is the document considered a primary source? **Content Description** A textual description of the content or purpose of the resource

murder murderers muriel murmansk muscle muscles muscular museum museums mushroom mushrooms MUSIC musical musici musicians musique mussolini musto mutants mutation mutton myers myosin myrlophyllum mystery myth mythology myths mêlée nae nail named nanking nanocomposite nanocomposites nanocyystals nanoindentation



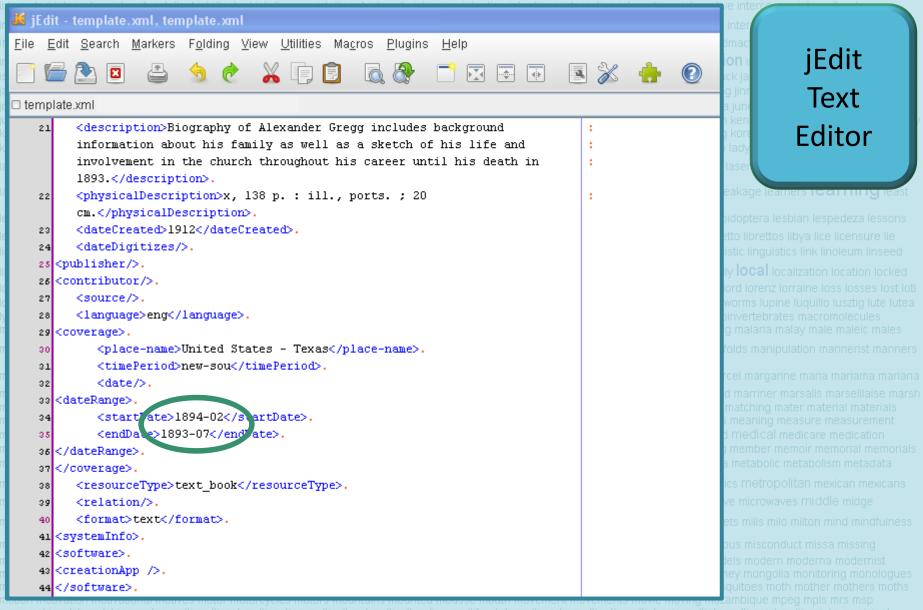
murder murderers muriel murmansk muscle muscles muscular museum museums mushroom mushrooms MUSIC musical musici musicians musique mussolini musto mutants mutation mutton myers myosin myriophyllum mystery myth mythology myths mêlée nae naji named nanking nanocomposite nanocomposites nanocrystals nanoindentation

influences informal informatics Information informed infrastructure ingersoll inheritance inhibitors initial initiative injured injuries **Template Reader**

interdisciplinary interest interests interface interfaces interfacial interden

Title(officialtitle):	Alexander Gregg: First Bishop of Texas		
Identifier(OCLC):	2629011		
Note:			
Institution:	UNT		
Collection:	тнс		
Creator	Name:	Gregg, Wilson	
	Role:	aut	
	Type:	per	
Subject(LCSH):	Bishops Texa	IS.	
Subject(LCSH):	Gregg, Alexander, 1819-1893.		
Subject(UNTL-BS):	People - Individuals		
Subject(UNTL-BS):	Religion - Denominations - Episcopal		
Subject(KWD):	biographies		
Description:	Biography of Alexander Gregg includes background information about his family as well as a sketch of his life and involvement in the church throughout his career until his death in 1893.		
Physical Description:	x, 138 p. : ill., ports. ; 20 cm.		
Date Created:	1912		
Date Digitized:			
Publisher			
Contributor:			
Source:			
Language:	eng		
Coverage:	Place:	United States - Texas	
	Era:	new-sou 1894-02 1893-07 e ror! End-date should be later than start-date	
	Start Date:	1894-02	
	End Date:	1893-07 e ror! End-date should be later than start-date	
Resource Type:	text_book		
Relation:			
Format:	text		

nference infestation inflation inflicted influence influences informal informatics INTORMATION informed infrastructure ingersoll inheritance inhibitors initial initiative injured injur njurious injury ink inlaid inlay inmates innocence innovation innovations inoculation inorganic inquiry insect insecticides insects insight insignia insomnia inspection installation



muczynski mulch mulching mules multi multicast multicasting multicultural multiculturalism multifractals multimedia multipath multipasic multipae multiracial multivariate municipal

murder murderers muriel murmansk muscle muscles muscular museum museums mushroom mushrooms MUSIC musical musici musicians musique mussolini musto mutants mutation mutton myers myosin myriophyllum mystery myth mythology myths mêlée nae naji named nanking nanocomposite nanocomposites nanocrystals nanoindentation

nference infestation inflation inflicted influence influences informal informatics INTOTII njurious injury ink inlaid inlay inmates innocence innovation innovations inoculation inor installations institute institutional institutions instruction instructional instructions instructional instructions instructional integrated integration integrity intellect intellectual intelligence intelligent intensity intension interdisciplinary interest interests interface interfaces interfacial intergenerational internet internet interorganizational interpersonal interpretation interracial interstate interval into internet invertebrates investigation investment investments involvement iodine ion iono slam island islands iso isocyanates isolation isomerization isomorphisms isotopes israel amaica james jamestown jan jap japan japanese japonica japs jars java jazz jean jeffer ohannes john johnson joint jon jones joseon joseph josquin journalism journals journey uxtaposition jund kabinet kafir kaiser kaki kale kang kansas kar kangdom kirg kalt kansa karapada king kangdom kirg kalt kansa karapada kangakan kangakan karapada kangakan
latin latino la

leather leave letter letters liebermann li lipid lipids liq locomotive lo lots louis loui lyapunov lyo macrophytes

manpower m marianne ma marshes ma math mather Metadata
Analysis
Tools:
Post-ingest

UNT

ing labor laboratories lat ngland **language** lang

rence laws lavered lavoff

I legislative lego legume te liability liberation libertie like lime limestone liming la lograph lithographs litigation logistics lolium london lor lolf ludwig lufkin luigi lumbe tric lyrique løgstrup macart magnetite magnificat magneties.

nures manuscript manuscr rs market marketing mar mas masculinity mask mas

medium medtner meerbaum meet

metafiction metal metallic metals metalwork metaphor meteorology method method mexico mezzo mice michael michel michigan microbiology microelectrode microelectrodes

migration migratory miguel mihalovici mike milanuzzi mild mildew miles ${\sf military}$ militia

miner mines ming **miniature** minimal minimization mining ministry mink minnesota m mississippi missouri misuse mite mitochondrial mixed mixer mixing mixture mixtures mmpi modification modulus mohair molasses mold mole molecular molecule moles molester mo montana monte monteverti montgomery montmorillonite moral moraxella mori momotion motivation motives motor motorcycles motors mountain mounted moraxella motives multipasting multipultural multipultural multipultural motives motors motors multipultural
Reports:

- NULL Values
- List All Values
 - List Values by Institution
 - List Values by Collection
- Browse All Values
 - Browse Values by Institution
 - Browse Values by Collection
- List All Records
 - List Records by Institution
 - List Records by Collection
- Authorities Values

Graphs:

- Records Added Over Time
- Records Added Per Month
- Files Added Over Time
- Map of Texas
 - by Collection
 - by Institution

Other Stuff:

- Search PTH System
 - Search a Collection
 - Search a Institution
- Word Clouds
 - by Collection
 - by Institution
- Word List
 - by Collection
 - by Institution

insurgency integral intercultural ernal international nation invasive inventional accisang isenheim ob jacques jail jails jobs joe johan johann stice juvenile ucky kernel ketenes key koryo kriegsmarine laid lake lamb lament at late latency latifolia

s learning least

an lespedeza lessons
bya lice licensure lie
link linoleum linseed
ization location locked
aine loss losses lost loti
uquillo lusztig lute lutea
macromolecules
ay male maleic males

tion mannerist manners

maria mariama mariana rsalis marseillaise marsh er material materials isure measurement dicare medication noir memorial memorials etabolism metadata

middle midge

hilton mind mindfulness

oderna modernist nonitoring monologues mother mothers moths g mpls mrs msp

murder murderers muriel murmansk muscle muscles muscular museum museums mushroom mushrooms MUSIC musical musici musicians musique mussolini musto mutants mutation mutton myers myosin myrjophyllum mystery myth mythology myths mêlée nae naji named nanking nanocomposite nanocomposites nanocyystals nanoindentation

Values that are more than 90% similar to the previous value will be highlighted.

off/on

ignore qualifiers | use qualifiers

all qualifiers

value	count
aut: Hare, Maud Cuney	1
aut: Hartley, Oliver Cromwell, 1823-1859	5
aut: Hartley, Oliver Cromwell, 1823-1859.	13
aut: Hartley, R. K.	12
aut: Hartnett, Minerva	1
aut: Hawker, H. W.	2
aut: Hawker, H. W. (Herman William)	12
aut: Hayes, F. A.	1
aut: Hays, Hugh	1
aut: Hazzard, J. L.	1
aut: Hearn, W. Edward (Williamson Edward)	3
aut: Henderson, J. P.	1
aut: Hendrickson, B. H. (Bertram Higbie)	4
aut: Hendrix, R.	1

migration migratory migr

Enhanced by Qualifier - Use/Ignore

Enhanced by Highlighter - On/Off

ignore qualifiers | use qualifiers

UNTL-BS - (64394)

value

UNTL-BS: Business, Economics and Finance - Medicine - Dentist

UNTL-BS: Business, Economics and Finance - Medicine - Dentists

UNTL-BS: Business, Economics and Finance - Medicine - Doctors

UNTL-BS: Business, Economics and Finance - Medicine - Hospitals

UNTL-BS: Business, Economics and Finance - Medicine - Nurses

UNTL-BS: Business, Economics and Finance - Medicine - Nursing Homes

UNTL-BS: Business, Economics and Finance - Medicine - Red Cross

Metadata Analysis: NULL

DC Element	Number that are NULL
DC Title	22
DC Creator	26390
DC Subject	1516
DC Description	0
DC Contributor	39967
DC Publisher	41885
DC Identifier	4942
DC Source	43613
DC Relation	45914
DC Language	271
DC Coverage	1241
DC Date	2125
DC Resource Type	287
DC Format	194
DC Rights	16171
UNTL Degree	46826
46828 Total Reco	ords

structional instructions instructors instrument instrumental instrumentation instruments insurance insurgency integral intelligent intensity intensive intention intentions interaction interactions interactive intercourse intercultural ial intergenerational interpovernmental international interactions in

rphisms isotopes israel israeli issue is
ars java jazz jean jefferson jeffrey jenli
nalism inurnals journey joy judges jud

ir varakul karg karl karst kastner kate kat ı kir ydom kirby kitchens klein klezmer klu p belinin abor laboratories laboratory lacan

ngland language languages lapidar

ON legislative lego legumes leguminous eyte liability liberation liberties libertine lib

nograph lithographs litigation little livesto ic logistics lolium london loneliness long lt lolf ludwig lufkin luigi lumber lumbricus lu yric lyrique løgstrup macarthur machine m e magnetite magnificat magritte mahler m

ged **management** manager mar

anures manuscript manuscripts many man ers market marketing markets marking m a mas masculinity mask masking maslanka urice max mayflies mayview mbti mcgehear

ledicinal medicine medieval meditation mediterranean medium medtner meerbaum meeting meeting lemory mems men mendelssohn mental mentally mentor mentoring menus merchant mercier merge

metafiction metal metallic metals metalwork metaphor meteorology method methodology me

Null Value Analysis Tools

y militia militiamen m

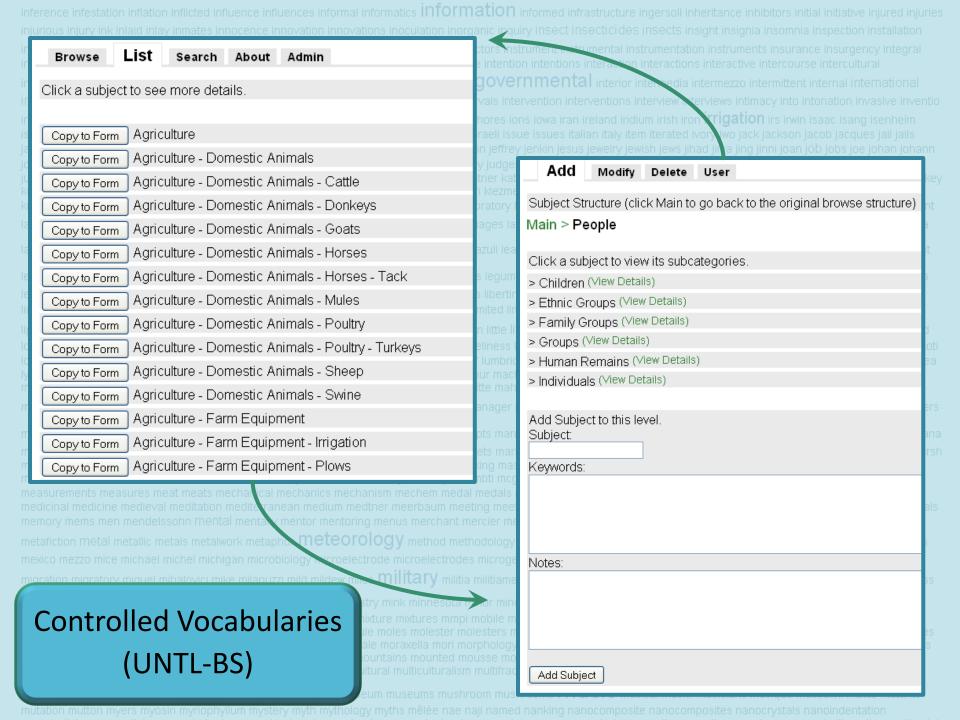
es mmpi mobile mobilit ester molesters moltini mori morphology mon inted mousse mouth n Ituralism multifractals n

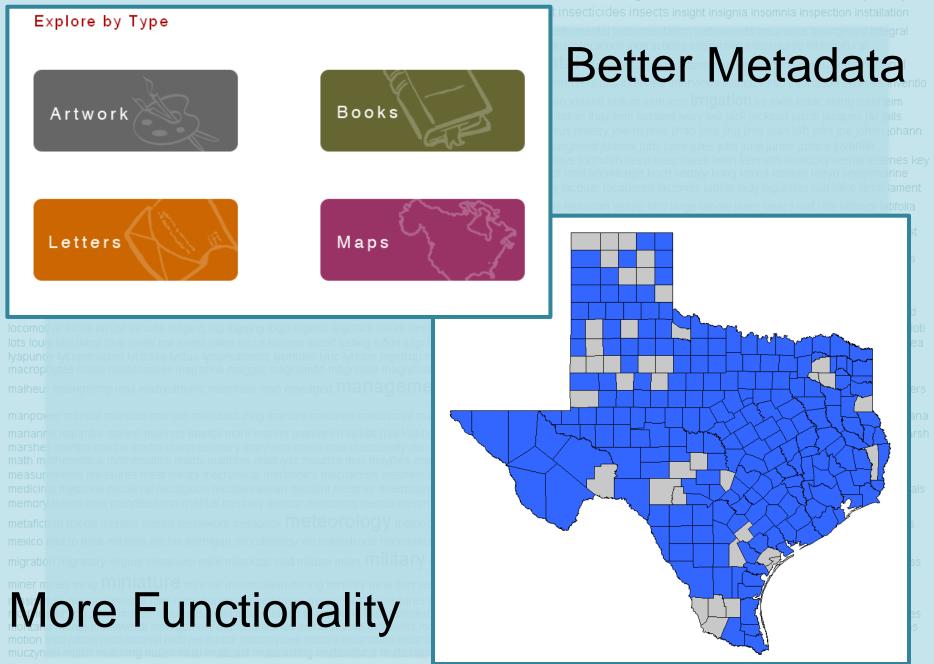
Metadata Analysis: NULL dc_title

details | thumbnails

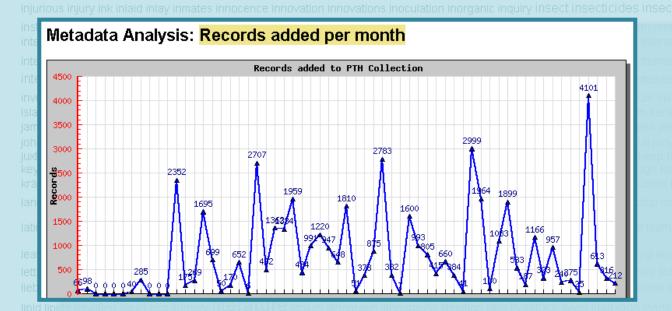
Value:	Count:	
1	meta-pth-6111	<u>edit</u>
2	meta-pth-6577	<u>edit</u>
3	meta-pth-6580	<u>edit</u>
4	meta-pth-6581	<u>edit</u>
5	meta-pth-6582	<u>edit</u>
6	meta-pth-9528	<u>edit</u>
7	meta-pth-9546	<u>edit</u>
8	meta-pth-9547	<u>edit</u>
9	meta-pth-9617	<u>edit</u>
10	meta-pth-14408	<u>edit</u>
11	meta-pth-14409	<u>edit</u>
12	meta-pth-14410	<u>edit</u>
13	meta-pth-14411	<u>edit</u>
14	meta-pth-14412	<u>edit</u>
15	meta-pth-14413	<u>edit</u>
16	meta-pth-14414	<u>edit</u>
17	meta-pth-14415	<u>edit</u>
18	meta-pth-15485	<u>edit</u>
19	meta-pth-39258	<u>edit</u>
20	meta-pth-39259	<u>edit</u>
21	meta-pth-39260	<u>edit</u>
22	meta-nth-39261	edit

utation mutton myers myosin myriophyllum mystery myth mythology myths mêlée nae naji named nanking nanocomposite nanocomposites nanocrystals nanoindentation



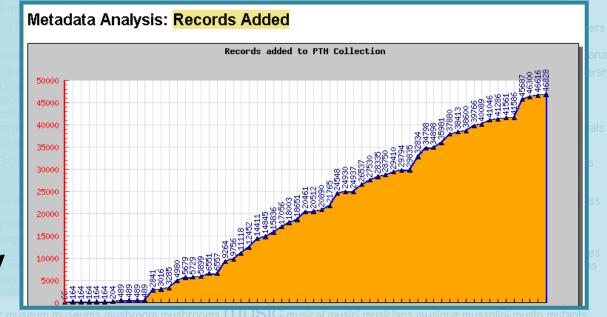


murder murderers muriel murmansk muscle muscles muscular museum museums mushroom mushrooms MUSIC musical musici musicians musique mussolini musto mutants mutation mutton myers myosin myrioohyllum mystery myth mythology myths mêlée nae nail named nanking nanocomposite nanocomposites nanocystals nanoindentation



Better Metadata

More Functionality

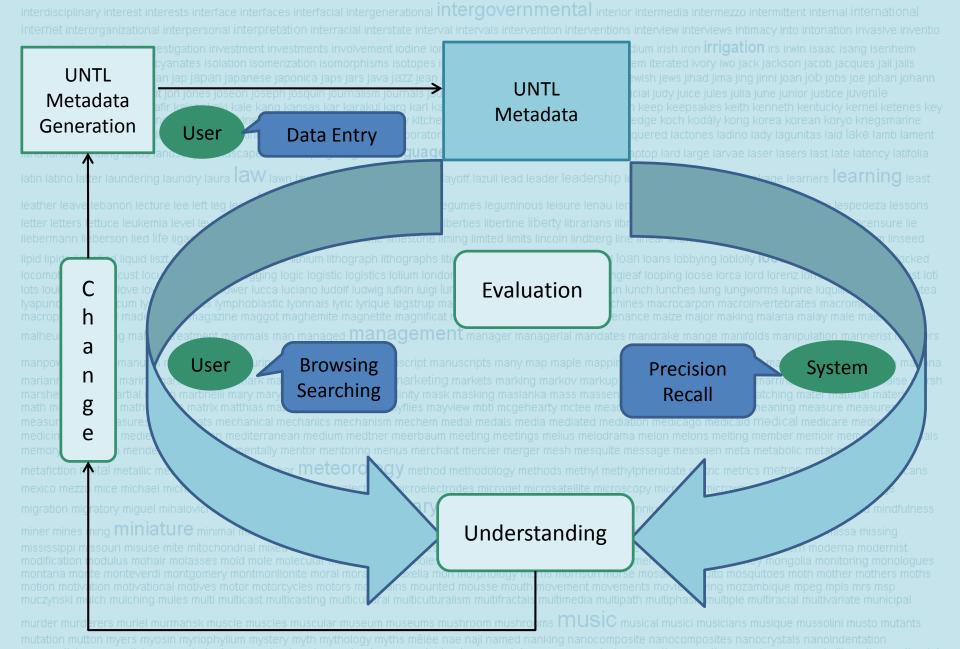


Summary

- **Determine level of quality required**
- Determine nature of gap and how to close it
- Machine verses human error handling
- **Compromise**
- **Prioritize**
- Test the workflow



Measure Quality and Usefulness of UNT Metadata



inference infestation inflation inflicted influence influences informal informatics if it of it informed infrastructure ingersoll inheritance inhibitors initial initiative injured injuries injurious injury ink inlaid inlay inmates innocence innovation innovations inoculation inorganic inquiry insect insecticides insects insight insight insighia insomnia inspection installation

References & Web Sites Consulted

- Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. Online Review, 13(5), 407-424.
- Netcraft (2009). April 2009 Web Server Survey. Retrieved May 19th, 2009 from http://news.netcraft.com/archives/web_server_survey.html
- OCLC (2007). Sharing, privacy and Trust in our Networked World. Retrieved May 19th,
 2009 from http://www.oclc.org/reports/pdfs/sharing.pdf
- TechSmith, Co. (2008). "UX 2.0: Any User, Any Time, Any Channel." Retrieved May 19th, 2009 from: http://download.techsmith.com/morae/docs/UserExperience2_0.pdf
- UNT Libraries Metadata Initiative page. Retrieved May 19th, 2009 from: http://www.library.unt.edu/digitalprojects/metadata

inference infestation inflation inflicted influence influences informal informatics information information information infrastructure ingersoll inheritance inhibitors initial initiative injured injuried injuried information.
installations institute institutional institutions instruction instructional instructions instructors instrument instrumental instrumentation instruments insurance insurgency integral integrated integration integrity intellect intellectual intelligence intelligent intensity intensive intention intentions interaction interactions interactive intercourse intercultural
interdisciplinary interest interests interface interfaces interfacial intergenerational intergovernmental interior intermedia intermezzo intermittent internal international
inventory invertebrates investigation investment investments involvement iodine ion ionophores ions iowa iran ireland iridium irish iron irrigation irs irwin isaac isang isenheim islam island islands iso isocyanates isolation isomerization isomorphisms isotopes israel israeli issue issues italian italy item iterated ivory iwo jack jackson jacob jacques jail jails jamaica james jamestown jan jap japan japanese japonica japs jars java jazz jean jefferson jeffrey jenkin jesus jewelry jewish jews jihad jima jing jinni joan job jobs joe johan johann johannes john johnson joint jon jones joseon joseph josquin journalism journals journey joy judges judging judgment judicial judy juice jules julia june junior justice juvenile juxtaposition juyi kabinet kafir kaiser kaki kale kang kansas kar karakul karg karl karst kastner kate katrina kaye kazhdan keep keensakes keith kenneth kentucky kernel ketenes ke keys killing kindergarten kindertotenlieder kinder ook sa king kingdom kirby kitchens klais kle Oer klughardt knot knowledge ko and salvy kong korea korean koryo kriegsmarine krämerspiegel kudzu kurtuluş kuwait königski ker lab selembelim lah and sarakul karakun jarakultan jara
land landfill landing lands landscape landscape lands up land land land land land land land land
krämerspiegel kudzu kurtuluş kuwait königski der lab dielembeli milat para araba kartında koniga kurtuluş kuwait königski der lab dielembeli milat para araba kartında kartında kartında kartında laba kartında ka
leather leave lebanon lecture lee left leg legal legislation legislative lego legumes leguminous leisure lenau lending leo leopold lep lepidoptera lesbian lespedeza lessons
lipid lipids liquefied liquid liszt literacy literature lithium lithograph lithographs litigation little lives. Ck living load loan loans lobbying loblolly local localization location locked locomotive locus locust locusts lodging log logging logic logistic logistics lolium london longliness log longitudinal longleaf looping loose lorca lord lorenz lorraine loss losses lost lot lots louis louisiana love lovell low lowell lower lucca luciano ludolf ludwig lufkin luigi lumbar localizations luminescence lun lunch lunches lung lungworms lupine luquillo lusztig lute lutear lyapunov lycopersicum lyctidae lyctus lymphoblastic lyonnais lyric lyrique løgska amacan and hine machinery machines macrocarpon macroinvertebrates macromolecules macrophytes made madécasses magazine maggot maghemite magnetite magn
malheur malingering mall maltreatment mammals man managed manage. em managerial mandates mandrake mange manifolds manipulation mannerist manners
manpower manual manuals manuel manufacturing manure manures manuscript ma many map maple mapping Maps marble marcel margarine maria mariama marian
marianne marimba marine marines marital mark marker markers market marketing marshes martha martine marines marinelli mary maryland maría mas masculinity mas marshes martha martinal martin martinelli mary maryland maría mas masculinity mas marshes martha martinal martinelli mary maryland maría mas masculinity mas marshes martha martinal martinelli mary maryland maría mas masculinity mas marshes market mar
metafiction metal metallic metals metalwork metaphor meteorology method et dology methods methyl methylphenidate metric metrics metropolitan mexican mexicans
mexico mezzo mice michael michel michigan microbiology microelectrode microelectri des icrogel microsatellite microscopy microsoft microwave microwaves middle midge
migration migratory miguel mihalovici mike milanuzzi mild mildew miles military mil a meamen milit milk milking millennium miller millet millets mills milo milton mind mindfulness
miner mines ming miniature minimal minimization mining ministry mink miss at a minor minorities minority minute miscarriage miscellaneous misconduct missa missing mississippi missouri misuse mite mitochondrial mixed mixer mixing mixture mixtures may mobility mobility mobility mobility model model modeling models modern moderna modernist modification modulus mohair molasses mold mole molecular molecule moles molester molesters molting molybdenum momentum monetary money mongolia monitoring monologues montana monte monteverdi montgomery montmorillonite moral morale moraxella mori morphology morris morrison morse mosaic mosquitoe mosquitoes moth mother mothers moths motion motivation motivational motives motor motorcycles motors mountains mounted mousse mouth movement movements movie moving mozambique mpeg mpls mrs msp muczynski mulch mulching mules multi multicast multicasting multicultural multiculturalism multifractals multimedia multiphasic multiphasic multiple multiracial multivariate municipal