

# Annotation for Transparent Inference (ATI):

Selecting a Platform for Qualitative Research Based on  
Individual Sources

IASSIST 2016, Bergen, Norway

Sebastian Karcher  
QDR, Syracuse University



# Slides Online

<http://doi.org/10.6084/m9.figshare.3409054>

QDR

# How Can We Make Qualitative Inference (More) Transparent?

# Quantitative Research: Matrix Data

RStudio

File Edit Code View Plots Session Build Debug Tools Help

Go to file/function Addins

gsu2010.data

Filter

X	condition	global_warming	human_induced	support_cap	smallercar	ce_ppi	ce_er	partyid	ideology
1	1	5	0.33	-0.33	0.33	-0.67	0.33	0.33	0.67
2	2	3	-1.00	-0.67	-0.67	-0.67	-0.33	0.00	-0.67
3	3	4	0.67	1.00	0.33	0.67	-0.33	-1.00	1.00
4	4	1	0.67	1.00	0.00	0.67	0.67	0.67	0.00
5	5	1	1.00	1.00	-0.33	1.00	1.00	1.00	0.67
6	6	4	0.67	0.67	1.00	-0.33	0.33	0.33	-0.33
7	7	5	0.33	0.67	0.00	0.33	0.67	1.00	0.67
8	8	2	-0.33	0.00	0.00	0.67	0.67	0.00	0.00
9	9	3	1.00	1.00	0.67	0.33	-0.67	-0.33	0.00
10	10	3	0.33	0.67	0.00	-0.67	0.33	0.33	0.67
11	11	1	1.00	1.00	1.00	1.00	1.00	-1.00	-1.00
12	12	2	0.67	0.33	0.00	0.33	0.67	0.33	-0.67
13	13	2	0.67	1.00	1.00	0.67	0.67	0.67	1.00
14	14	4	0.67	0.67	0.00	0.33	0.00	-0.33	0.67
15	15	2	0.33	0.67	0.67	0.33	0.33	0.67	1.00
16	16	5	0.33	-0.67	0.33	-0.33	0.33	0.33	0.33
17	17	4	0.67	0.67	0.00	0.33	0.00	0.00	0.67
18	18	2	1.00	0.33	1.00	-0.33	0.33	0.33	-0.67
19	19	4	0.33	0.67	0.67	-0.33	0.33	0.33	0.67

Showing 1 to 19 of 622 entries



Table 2. Treatment Effects (Experiment 1)

	Perceived Personal Influence	Expected Reciprocity	Believe Global Warming is Happening	Believe Global Warming is Human Induced	Support Cap on Carbon Emissions	Willingness to Take Personal Action
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0.38** (0.06)	0.23** (0.06)	0.37** (0.07)	0.25** (0.07)	0.28** (0.07)	0.20** (0.07)
Pro Descriptive (D)	-0.04 (0.08)	0.02 (0.08)	0.03 (0.09)	-0.03 (0.09)	-0.06 (0.09)	-0.02 (0.09)
Pro Descriptive + Injunctive (D+I)	-0.08 (0.08)	0.07 (0.08)	0.09 (0.09)	0.06 (0.09)	-0.02 (0.09)	-0.05 (0.08)
Con D	-0.18** (0.08)	-0.06 (0.08)	-0.06 (0.09)	-0.05 (0.09)	-0.16** (0.09)	-0.17** (0.09)
Con D+I	-0.10 (0.08)	0.11* (0.07)	0.03 (0.09)	-0.06 (0.09)	-0.08 (0.09)	-0.06 (0.09)
PartyID	0.21** (0.08)	0.21** (0.08)	0.28** (0.11)	0.31** (0.09)	0.21** (0.10)	0.13* (0.09)
PartyID * Pro D	-0.05 (0.11)	-0.04 (0.11)	-0.10 (0.14)	-0.04 (0.13)	-0.05 (0.13)	0.04 (0.13)
PartyID * Pro D+I	-0.10 (0.11)	-0.12 (0.11)	-0.06 (0.13)	-0.13 (0.12)	-0.14 (0.13)	0.04 (0.11)
PartyID * Con D	-0.10 (0.11)	-0.19** (0.11)	0.06 (0.14)	-0.10 (0.13)	0.03 (0.13)	0.11 (0.13)
PartyID * Con D+I	-0.13 (0.11)	-0.21** (0.10)	0.01 (0.13)	-0.15 (0.12)	0.01 (0.13)	0.04 (0.11)
SER	.48	.47	.52	.53	.51	.54
n	612	612	612	611	612	612

\*p < .10. \*\*p < .05. one-tailed test  
 Note: D = Descriptive, I = Injunctive. Cell entries are unstandardized linear regression coefficients with bootstrapped standard errors in parentheses. Baseline is the control condition. Party identification is coded with seven categories from Republican (-1) to Independent (0) to Democrat (+1). SER is the Standard Error of the Regression (Beck, 2010).

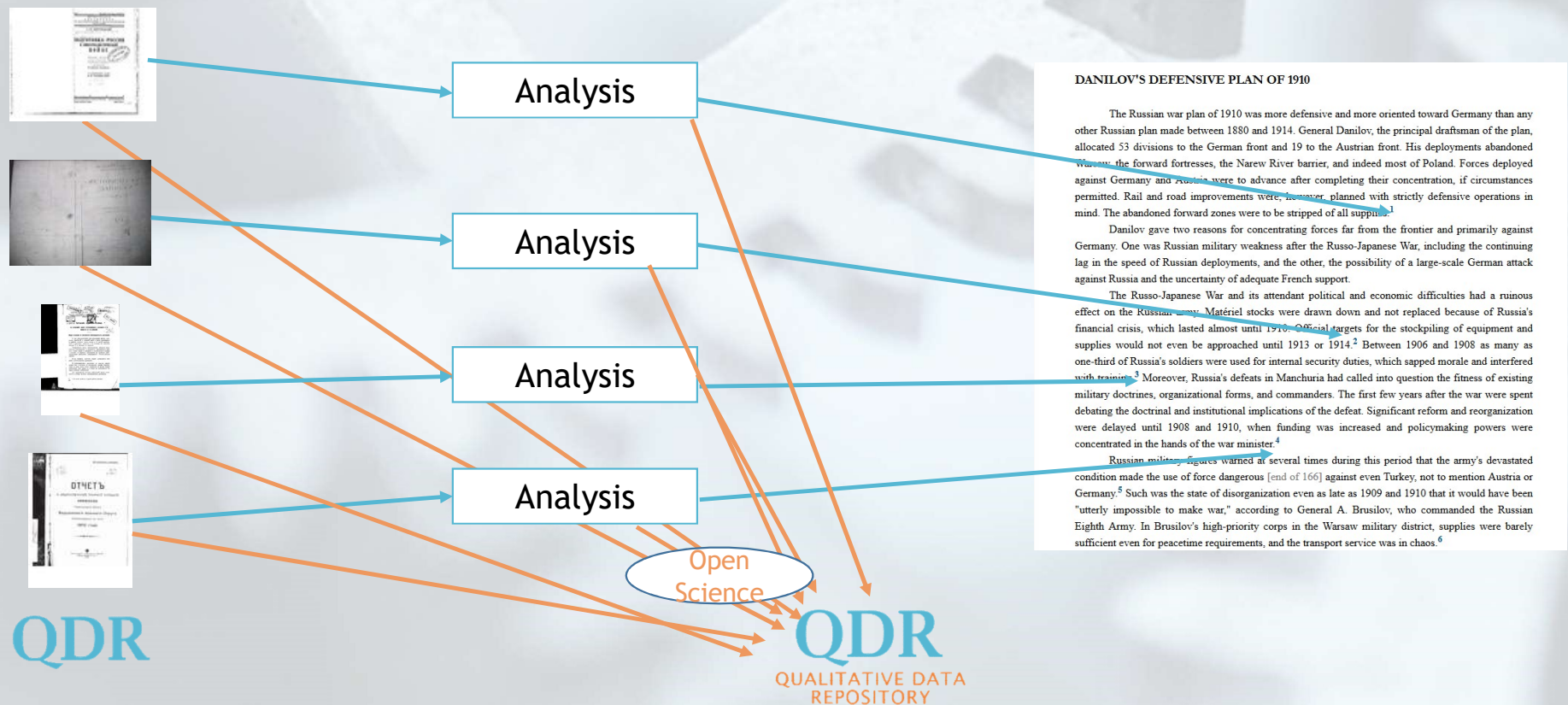
While we expected norms promoting action to influence both beliefs (Hypothesis 1a) and behavioral intentions (Hypothesis 2a), we find little support our hypotheses. Table 2 reports regression results with indicators for each treatment condition (relative to a control condition), including a control for party identification. Looking at the rows for the Pro Norm conditions, we do not see any effect on any outcome. Indeed for several outcomes (perceived personal influence and support for a carbon emissions tax), we find effects in the opposite direction of our

Toby Bolsen, Thomas J. Leeper, and Matthew Shapiro. 2014. "Doing What Others Do: Norms, Science, and Collective Action on Global Warming." *American Politics Research* 42(1): 65–89.

QDR

The  
**Dataverse**  
 Project

# Qualitative Research: Granular Data





# Active Citation in Practice

Table 4 documents performance indicators for the three cases under study in this paper. These figures are based on internal documents gathered during field research and in some cases contradict published “official figures” by the federal government.

For example, although the national census claims that Neza’s water coverage is 97% (INEGI, 2010), internal documents indicate that coverage is closer to 87%, and that only approximately 50% of the city has registered connections (CISA, 2007, p. 24). Municipalities throughout Mexico State have



Diagnostico del subsector agua potable, resumen ejecutivo. CISA Solutions, p. 24.

**Annotation:** *This document is a comprehensive report of the physical, financial and operational state of the Neza water utility undertaken by a private consulting company hired in 2007. This study was paid for by a higher tier of government. It was provided to me by an employee who had taken a copy and saved it for themselves.*

**Location:** shareable

**Excerpt:** EXCERPT IN ORIGINAL: La produccion es inferior a la demanda, que representa un deficit del 5.02% de la misma. Se tiene una cobertura en agua potable del 86.97%, sin embargo si comparamos el estimado de usuarios a partir de la poblacion contra el numero de usuarios registrados en el padron del Organismo Operador del Municipio, esta solo

<https://qdr.syr.edu/discover/projectcontentherrerar>

# Towards Annotations for Transparent Inference (ATI)



Automation

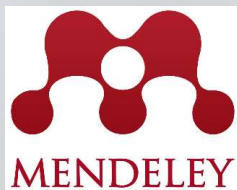


Annotation



Standardization

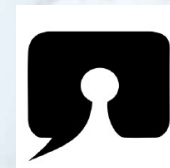
# Use Existing Tools & Technology



zotero



L<sup>A</sup>T<sub>E</sub>X



Collect / Store

Write

Display

QDR



# Getting the Data from Where They Are

## Online Database

<b>Item Type</b>	Report	
<b>Title</b>	Municipio de Nezahualcyotl, Estado de México: Diagnostico del subsector agua potable, resumen ejecutivo	HTML
<b>Author</b>	CISA	prehensive report of the physical, financial and y a private consulting company hired in ment. It was provided to me by an elves.
<b>Institution</b>	CISA Solutions	presenta un deficit del 5.02% de la misma. embargo si comparamos el estimado de arios registrados en el padron del 19.75%
<b>Date</b>	2007	ing to water utility figures] of 86.97%, to the city's population, coverage only
<b>URL</b>		
<b>Tags</b>		
<b>Notes and Attachments</b>		
 <a href="#">[Untitled]</a>		
 <a href="#">CISA_2007_Herrera.pdf (pdf, 939.9 KB)</a> (Attachment Details)		

# Regular Citation Workflow

For example, although the national census claims that Neza's water coverage is 97% (INEGI, 2010), internal documents indicate that coverage is closer to 87%, and that only approximately 50% of the city has registered connections { ADDIN ZOTERO\_ITEM CSL\_CITATION {"citationID":"21qfthcagg","properties":{"formattedCitation":"(CISA, 2007, p. 24)","plainCitation":"(CISA, 2007, p. 24)"},"citationItems":[{"id":9335,"uris": "<http://zotero.org/groups/483726/items/C88VJNIS>","uri":["http://zotero.org/groups/483726/items/C88VJNIS"], "itemData":{"id":9335,"type":"report","title":"Municipio de Nezahualcyotl, Estado de México: Diagnostico del subsector agua potable, resumen ejecutivo","publisher":"CISA Solutions","author":[{"literal":"CISA"}],"issued":{"date-parts":[["2007"]]},"locator":"24"},"schema":"https://github.com/citation-style-language/schema/raw/master/csl-citation.json"}] }.

# Published ATIs

Publisher  
Content

QDR




 Download PDF  Export   Advanced search

Table options ▼




For example, although the national census claims that Neza's water coverage is 97% (INEGI., 2010), internal documents indicate that coverage is closer to 87%, and that only approximately 50% of the city has registered connections (CISA, 2007, p. 24). Municipalities throughout Mexico State have been unable to make important service improvements due to insufficient cost recovery. A 2007 audit of Mexico state's 36 utilities show that half of all utilities and ¼ of metropolitan utilities are fiscally insolvent and that most do not recover enough revenue to invest in minimal service improvements (Poder Legislativo del Estado de México, 2008). While reliable data are difficult to attain and interpret, data triangulated through multiple sources suggest that the utilities analyzed in this study are representative of the types of service deficiencies found throughout the country.


The following section analyzes performance upgrades across Naucalpan, Toluca, and Neza and demonstrates the political and institutional challenges of implementing commercialization after decentralization.

### 5. Evaluating reform experiences in three cities

The three cases demonstrate the different pathways through which political conflict and institutional constraints reduce the likelihood that commercialization policies will be adopted and sustained over time. Naucalpan—due to a higher income consumer base and innovative mayoral leadership—was somewhat more successful than the other two cities in generating some internal revenue and making some service improvements, but the extent to which these reforms will endure over time is uncertain. Toluca lacked strong leadership, but also faced a greater amount of urban poor users, less industrial clients, and a greater amount of “non-billable” consumers such as public institutions. Neza failed to initiate any cost recovery reforms, encumbered by an urban poor consumer base. a

Repository  
Content

Qualitative Data Repository    2 mins


 QDR  
QUALITATIVE DATA  
REPOSITORY

(CISA, 2007, p. 24)

**Analytic Note:** Annotation: This document is a comprehensive report of the physical, financial and operational state of the Neza water utility undertaken by a private consulting company hired in 2007. This study was paid for by a higher tier of government. It was provided to me by an employee who had taken a copy and saved it for themselves.

**Excerpt:** La producción es inferior a la demanda, que representa un déficit del 5.02% de la misma. Se tiene una cobertura en agua potable del 86.97%, sin embargo si comparamos el estimado de usuarios a partir de la población contra el número de usuarios registrados en el padrón del Organismo Operador del Municipio, esta solo alcanza el 49.75%

**Translation:** There is an official coverage data [according to water utility figures] of 86.97%, however if we compare the number of registered users to the city's population, coverage only reaches 49.75%.







**MUNICIPIO DE NEZAHUALCOYOTL**  
**ESTADO DE MÉXICO**  
**DIAGNÓSTICO DEL SUBSECTOR AGUA POTABLE**  
**Análisis Técnico - Comercial**

Para la prestación del Servicio de agua potable, drenaje, alcantarillado, tratamiento y disposición de los aguas residuales el Municipio de Nezahualcóyotl cuenta con un Organismo Público Descentralizado, denominado ODAPAS, al cual según el Decreto Nº 201, expedido por la I Legislatura del Estado de México, publicado en el Gaceta de Gobierno, 26 de marzo de 1993.

La población que vive en Nezahualcóyotl, según proyecciones de la CONAPO al 2008 es de 1'105,731; con un índice de hacinamiento de 4.14 según el II Censo de Población 2005 del INEGI. El ODAPAS brinda el servicio de agua potable drenaje y alcantarillado sanitario en el la mayor parte del municipio.

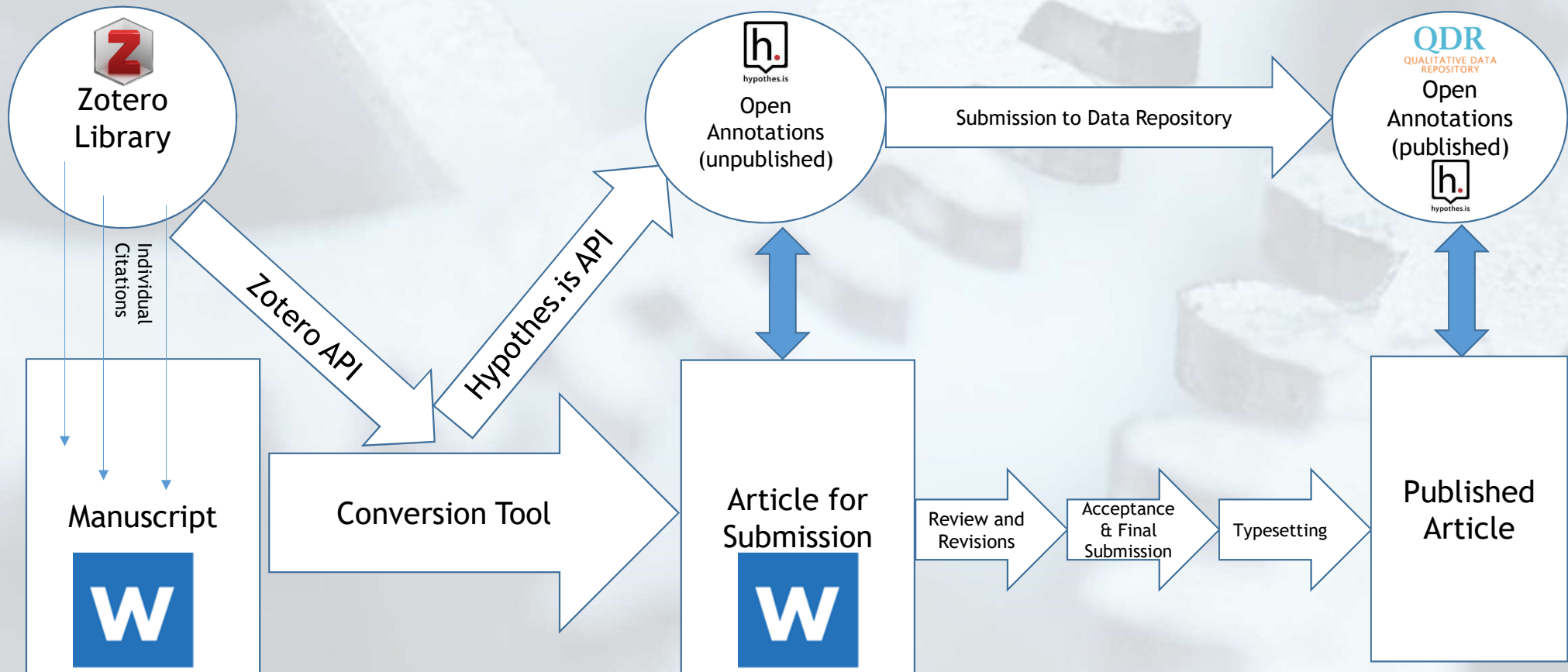
Para el abasto de agua potable el Organismo Operador cuenta con 8 pozos, los cuales otorgan un caudal promedio en conjunto de 901 lps.

El municipio de Nezahualcóyotl recibe agua en bloque, de fuentes Federales la cual se distribuye a través de la Planta Zaragoza, tanque Córdova y otras derivaciones para un

Less    

Some of  
this  
= Mock-up

# A Workflow for ATI





The background of the slide features a close-up, shallow depth-of-field photograph of several interlocking metal gears. The gears are light-colored, possibly silver or chrome, and their teeth are clearly visible. The focus is sharp on the gears in the foreground, while the ones in the background are blurred, creating a sense of depth and mechanical complexity.

# Slides Online

<http://doi.org/10.6084/m9.figshare.3409054>

QDR