Content analysis 2: Designs, reliability & validity

LQRPS

Frederik Hjorth
fh@ifs.ku.dk
fghjorth.github.io
@fghjorth

Department of Political Science
University of Copenhagen

February 10th, 2017

1. 3. Coding
   - Principles
   - Examples
2. 4. Reliability
   - Defining
   - Assessing
   - Krippendorff's alpha
   - Examples
3. 5. Analysis
4. Automated content analysis
   - General principles
   - King et al.

1. 3. Coding
   - Principles
   - Examples
2. 4. Reliability
   - Defining
   - Assessing
   - Krippendorff's alpha
   - Examples
3. 5. Analysis
4. Automated content analysis
   - General principles
   - King et al.

Unit of sampling $\neq$ unit of analysis

- Contrast w. survey research
- Ex. 1: sampling articles ctr. analyzing paragraphs
- Ex. 2: Comparative Manifesto Project
- Sampling units typically $>$ analysis units, but not always (e.g. King et al.)

Unit of sampling $\neq$ unit of analysis

- Contrast w. survey research
- Ex. 1: sampling articles ctr. analyzing paragraphs
- Ex. 2: Comparative Manifesto Project
- Sampling units typically $>$ analysis units, but not always (e.g. King et al.)

Unit of sampling $\neq$ unit of analysis

- Contrast w. survey research
- Ex. 1: sampling articles ctr. analyzing paragraphs
- Ex. 2: Comparative Manifesto Project
- Sampling units typically $>$ analysis units, but not always (e.g. King et al.)

Unit of sampling $\neq$ unit of analysis

- Contrast w. survey research
- Ex. 1: sampling articles ctr. analyzing paragraphs
- Ex. 2: Comparative Manifesto Project
- Sampling units typically $>$ analysis units, but not always (e.g. King et al.)

Unit of sampling $\neq$ unit of analysis

- Contrast w. survey research
- Ex. 1: sampling articles ctr. analyzing paragraphs
- Ex. 2: Comparative Manifesto Project
- Sampling units typically $>$ analysis units, but not always (e.g. King et al.)

»[Coding instructions] must delineate the phenomena of interest and define the recording units to be described in analyzable terms, the categories relevant to the research project, and their organization into a system of separate variables—also called a data language.« (Krippendorff, p. 351)

Codebooks serve three purposes:

a. Instruct coders
b. Link structured and unstructured data
c. Document the research process

»[Coding instructions] must delineate the phenomena of interest and define the recording units to be described in analyzable terms, the categories relevant to the research project, and their organization into a system of separate variables—also called a data language.« (Krippendorff, p. 351)

Codebooks serve three purposes:

a. Instruct coders
b. Link structured and unstructured data
c. Document the research process

»[Coding instructions] must delineate the phenomena of interest and define the recording units to be described in analyzable terms, the categories relevant to the research project, and their organization into a system of separate variables—also called a data language.« (Krippendorff, p. 351)

Codebooks serve three purposes:

a. Instruct coders
b. Link structured and unstructured data
c. Document the research process

»[Coding instructions] must delineate the phenomena of interest and define the recording units to be described in analyzable terms, the categories relevant to the research project, and their organization into a system of separate variables—also called a data language.« (Krippendorff, p. 351)

Codebooks serve three purposes:

a. Instruct coders
b. Link structured and unstructured data
c. Document the research process

»[Coding instructions] must delineate the phenomena of interest and define the recording units to be described in analyzable terms, the categories relevant to the research project, and their organization into a system of separate variables—also called a data language.« (Krippendorff, p. 351)

Codebooks serve three purposes:

a. Instruct coders
b. Link structured and unstructured data
c. Document the research process

»[Coding instructions] must delineate the phenomena of interest and define the recording units to be described in analyzable terms, the categories relevant to the research project, and their organization into a system of separate variables—also called a data language.« (Krippendorff, p. 351)

Codebooks serve three purposes:

a. Instruct coders
b. Link structured and unstructured data
c. Document the research process

»[Coding instructions] must delineate the phenomena of interest and define the recording units to be described in analyzable terms, the categories relevant to the research project, and their organization into a system of separate variables—also called a data language.« (Krippendorff, p. 351)

Codebooks serve three purposes:

a. Instruct coders
b. Link structured and unstructured data
c. Document the research process

»[Coding instructions] must delineate the phenomena of interest and define the recording units to be described in analyzable terms, the categories relevant to the research project, and their organization into a system of separate variables—also called a data language.« (Krippendorff, p. 351)

Codebooks serve three purposes:

a. Instruct coders
b. Link structured and unstructured data
c. Document the research process

Labor

»Three words describe good coder preparation: train, train, train (Neuendorf, p. 133)«

»Content analyst have reported spending months in training sessions with coders, during which time they refined categories, altered instructions [...]« (Krippendorff, p. 129)

Labor

»Three words describe good coder preparation: train, train, train (Neuendorf, p. 133)«

»Content analyst have reported spending months in training sessions with coders, during which time they refined categories, altered instructions [...]« (Krippendorff, p. 129)

Labor

»Three words describe good coder preparation: train, train, train (Neuendorf, p. 133)«

»Content analyst have reported spending months in training sessions with coders, during which time they refined categories, altered instructions [...]« (Krippendorff, p. 129)

Labor

»Three words describe good coder preparation: train, train, train (Neuendorf, p. 133)«

»Content analyst have reported spending months in training sessions with coders, during which time they refined categories, altered instructions [...]« (Krippendorff, p. 129)

Labor

»Three words describe good coder preparation: train, train, train (Neuendorf, p. 133)«

»Content analyst have reported spending months in training sessions with coders, during which time they refined categories, altered instructions [...]« (Krippendorff, p. 129)

Labor

»Three words describe good coder preparation: train, train, train (Neuendorf, p. 133)«

»Content analyst have reported spending months in training sessions with coders, during which time they refined categories, altered instructions [...]« (Krippendorff, p. 129)

Labor

»Three words describe good coder preparation: train, train, train (Neuendorf, p. 133)«

»Content analyst have reported spending months in training sessions with coders, during which time they refined categories, altered instructions [...]« (Krippendorff, p. 129)

Blind coding

»Ideally, the individuals who take part in the development of the recording instructions should not be the ones who apply them, for they will have acquired an implicit consensus that new coders cannot have and that other scholars who may wish to use the instructions cannot replicate.« (Krippendorff, p. 131)

»Blind coding, in which the coders do not know the purporse of the study, is desirable, to reduce bias that compromises validity« (Neuendorf, p. 133)

→ potential tradeoff btw. training and blindness

Blind coding

»Ideally, the individuals who take part in the development of the recording instructions should not be the ones who apply them, for they will have acquired an implicit consensus that new coders cannot have and that other scholars who may wish to use the instructions cannot replicate.« (Krippendorff, p. 131)

»Blind coding, in which the coders do not know the purporse of the study, is desirable, to reduce bias that compromises validity« (Neuendorf, p. 133)

→ potential tradeoff btw. training and blindness

Blind coding

»Ideally, the individuals who take part in the development of the recording instructions should not be the ones who apply them, for they will have acquired an implicit consensus that new coders cannot have and that other scholars who may wish to use the instructions cannot replicate.« (Krippendorff, p. 131)

»Blind coding, in which the coders do not know the purporse of the study, is desirable, to reduce bias that compromises validity« (Neuendorf, p. 133)

→ potential tradeoff btw. training and blindness

Blind coding

»Ideally, the individuals who take part in the development of the recording instructions should not be the ones who apply them, for they will have acquired an implicit consensus that new coders cannot have and that other scholars who may wish to use the instructions cannot replicate.« (Krippendorff, p. 131)

»Blind coding, in which the coders do not know the purporse of the study, is desirable, to reduce bias that compromises validity« (Neuendorf, p. 133)

→ potential tradeoff btw. training and blindness

Example: open-ended responses about reasons for voting for/against euro adoption

	KATEGORI	FORKLARING	EKSEMPEL
1	Generel	Generel, uspecificeret modstand mod EU/euro	"Jeg bryder mig ikke om euroen"
2	Suverænitet	Landets suverænitet/selvbestemmelse	"National selvstændighed"
3	Afstand	Afstand til Bruxelles/centralisme/demokrati	"Demokrati i EU"
4	Usikker	Generel usikkerhed/foretrækker at vente	"Vi ved ikke hvad vi får"
5	Priser	Frygt for højere prisniveau	"Priser"
6	Proces	Utilfredshed med en udemokratisk proces	"Politikerne snyder befolkningen"
7	Identitet	Vil ikke miste national identitet, dansk/svenskhed	"Vil forblive dansk/svensk"
8	Kronen	Specifikt ønske om at bevare kronen	"Vil bevare kronen"
9	Andet	Residualkategori - andre svar	-

Reliability: $\alpha \approx .4 \rightarrow$ what was the problem?

Example: open-ended responses about reasons for voting for/against euro adoption

	KATEGORI	FORKLARING	EKSEMPEL
1	Generel	Generel, uspecificeret modstand mod EU/euro	"Jeg bryder mig ikke om euroen"
2	Suverænitet	Landets suverænitet/selvbestemmelse	"National selvstændighed"
3	Afstand	Afstand til Bruxelles/centralisme/demokrati	"Demokrati i EU"
4	Usikker	Generel usikkerhed/foretrækker at vente	"Vi ved ikke hvad vi får"
5	Priser	Frygt for højere prisniveau	"Priser"
6	Proces	Utilfredshed med en udemokratisk proces	"Politikerne snyder befolkningen"
7	Identitet	Vil ikke miste national identitet, dansk/svenskhed	"Vil forblive dansk/svensk"
8	Kronen	Specifikt ønske om at bevare kronen	"Vil bevare kronen"
9	Andet	Residualkategori - andre svar	-

Reliability: $\alpha \approx .4 \rightarrow$ what was the problem?

Example: open-ended responses about reasons for voting for/against euro adoption

	KATEGORI	FORKLARING	EKSEMPEL
1	Generel	Generel, uspecificeret modstand mod EU/euro	"Jeg bryder mig ikke om euroen"
2	Suverænitet	Landets suverænitet/selvbestemmelse	"National selvstændighed"
3	Afstand	Afstand til Bruxelles/centralisme/demokrati	"Demokrati i EU"
4	Usikker	Generel usikkerhed/foretrækker at vente	"Vi ved ikke hvad vi får"
5	Priser	Frygt for højere prisniveau	"Priser"
6	Proces	Utilfredshed med en udemokratisk proces	"Politikerne snyder befolkningen"
7	Identitet	Vil ikke miste national identitet, dansk/svenskhed	"Vil forblive dansk/svensk"
8	Kronen	Specifikt ønske om at bevare kronen	"Vil bevare kronen"
9	Andet	Residualkategori - andre svar	-

Reliability: $\alpha \approx .4 \rightarrow$ what was the problem?

Example: open-ended responses about reasons for voting for/against euro adoption

	KATEGORI	FORKLARING	EKSEMPEL
1	Generel	Generel, uspecificeret modstand mod EU/euro	"Jeg bryder mig ikke om euroen"
2	Suverænitet	Landets suverænitet/selvbestemmelse	"National selvstændighed"
3	Afstand	Afstand til Bruxelles/centralisme/demokrati	"Demokrati i EU"
4	Usikker	Generel usikkerhed/foretrækker at vente	"Vi ved ikke hvad vi får"
5	Priser	Frygt for højere prisniveau	"Priser"
6	Proces	Utilfredshed med en udemokratisk proces	"Politikerne snyder befolkningen"
7	Identitet	Vil ikke miste national identitet, dansk/svenskhed	"Vil forblive dansk/svensk"
8	Kronen	Specifikt ønske om at bevare kronen	"Vil bevare kronen"
9	Andet	Residualkategori - andre svar	-

Reliability: $\alpha \approx .4 \rightarrow$ what was the problem?

	A	B	C	D	E	F	G	H	I
1	**Artikel-ID**	**Outlet**	**Date**	**Relevance**	**Frame**	**Winningorloosing**	**Strategy**	**Polconse**	**Coalition building**
2	e0bff6d4	24timer	13.11.2007	1	99	2	2	2	2
3	e0bff529	24timer	13.11.2007	1	99	2	2	2	2
4	e0bfc5c9	24timer	12.11.2007	1	3	1	2	1	1
5	e0bfc573	24timer	12.11.2007	1	2	2	1	1	1
6	e0bfc3bb	24timer	12.11.2007	1	3	2	1	1	1
7	e0bfc39c	24timer	12.11.2007	1	1	2	1	1	1
8	e0bfc387	24timer	12.11.2007	1	3	2	1	2	1
9	e0bf1f96	24timer	09.11.2007	1	1		2	2	2
10	e0bf1f95	24timer	09.11.2007	1	1	2	2	2	2
11	e0bf1ea9	24timer	09.11.2007	1	3	2	1	2	2
12	e0bf1e97	24timer	09.11.2007	1	1	2	1	1	2
13	e0bffc5e	BT	13.11.2007	99					
14	e0bffbbd	BT	13.11.2007	1	99	2	2	2	2
15	e0bffbb8	BT	13.11.2007	1	3	2	1	1	1
16	e0bfc7f3	BT	12.11.2007	1	99	2	2	2	2
17	e0bcf7f1	BT	12.11.2007	1	3	2	2	1	1
18	e0bfc7b8	BT	12.11.2007	1	99	2	2	2	2
19	e0cd1b29	BT	11.11.2007	1	99	2	2	2	2
20	e0bfa912	BT	11.11.2007	1	99	2	1	2	2
21	e0bfa8eb	BT	11.11.2007	1	3	2	1	2	2

Ark4 Ark2 Ark1 Ark3

Exercise 3

What would be your coding strategy for the data gathered in Ex. 2? Which categories would you ask coders to classify observations into?

Exercise 3

What would be your coding strategy for the data gathered in Ex. 2? Which categories would you ask coders to classify observations into?

»Reliability can be defined as *the extent to which a measuring procedure yields the same results on repeated trials*« (Neuendorf p. 141, emhp. added)

Conceptually, potential tradeoff btw. reliability and validity:

- e.g. measuring 'readability'
  - subjectively assessed
  - Flesch-Kincaid: $206.835 - 1.015\left(\frac{\text{total words}}{\text{total sentences}}\right) - 84.6\left(\frac{\text{total syllables}}{\text{total words}}\right)$
- but: badly conceptualized coding scheme $\rightarrow$ low reliability *and* validity

»Reliability can be defined as *the extent to which a measuring procedure yields the same results on repeated trials*« (Neuendorf p. 141, emhp. added)

Conceptually, potential tradeoff btw. reliability and validity:

- e.g. measuring 'readability'
  - subjectively assessed
  - Flesch-Kincaid: $206.835 - 1.015 \left(\frac{\text{total words}}{\text{total sentences}}\right) - 84.6 \left(\frac{\text{total syllables}}{\text{total words}}\right)$
- but: badly conceptualized coding scheme $\rightarrow$ low reliability *and* validity

»Reliability can be defined as *the extent to which a measuring procedure yields the same results on repeated trials*« (Neuendorf p. 141, emhp. added)

Conceptually, potential tradeoff btw. reliability and validity:

- e.g. measuring 'readability'
  - subjectively assessed
  - Flesch-Kincaid: $206.835 - 1.015 \left(\frac{\text{total words}}{\text{total sentences}}\right) - 84.6 \left(\frac{\text{total syllables}}{\text{total words}}\right)$
- but: badly conceptualized coding scheme $\rightarrow$ low reliability *and* validity

»Reliability can be defined as *the extent to which a measuring procedure yields the same results on repeated trials*« (Neuendorf p. 141, emhp. added)

Conceptually, potential tradeoff btw. reliability and validity:

- e.g. measuring 'readability'
  - subjectively assessed
  - Flesch-Kincaid: $206.835 - 1.015 \left(\frac{\text{total words}}{\text{total sentences}}\right) - 84.6 \left(\frac{\text{total syllables}}{\text{total words}}\right)$
- but: badly conceptualized coding scheme $\rightarrow$ low reliability *and* validity

»Reliability can be defined as *the extent to which a measuring procedure yields the same results on repeated trials*« (Neuendorf p. 141, emhp. added)

Conceptually, potential tradeoff btw. reliability and validity:

- e.g. measuring 'readability'
  - subjectively assessed
  - Flesch-Kincaid: $206.835 - 1.015\left(\frac{\text{total words}}{\text{total sentences}}\right) - 84.6\left(\frac{\text{total syllables}}{\text{total words}}\right)$
- but: badly conceptualized coding scheme $\rightarrow$ low reliability *and* validity

»Reliability can be defined as *the extent to which a measuring procedure yields the same results on repeated trials*« (Neuendorf p. 141, emhp. added)

Conceptually, potential tradeoff btw. reliability and validity:

- e.g. measuring 'readability'
  - subjectively assessed
  - Flesch-Kincaid: $206.835 - 1.015 \left(\frac{\text{total words}}{\text{total sentences}}\right) - 84.6 \left(\frac{\text{total syllables}}{\text{total words}}\right)$
- but: badly conceptualized coding scheme $\rightarrow$ low reliability *and* validity

»Reliability can be defined as *the extent to which a measuring procedure yields the same results on repeated trials*« (Neuendorf p. 141, emhp. added)

Conceptually, potential tradeoff btw. reliability and validity:

- e.g. measuring 'readability'
  - subjectively assessed
  - Flesch-Kincaid: $206.835 - 1.015 \left(\frac{\text{total words}}{\text{total sentences}}\right) - 84.6 \left(\frac{\text{total syllables}}{\text{total words}}\right)$
- but: badly conceptualized coding scheme $\rightarrow$ low reliability *and* validity

The naive approach: percent agreement

Example: dichotomous coding 0/1, 2 coders

	1	2	3	4	5	6	7
Lars	0	0	1	0	0	0	0
Solrun	0	0	0	0	1	0	0

Percent agreement $= \frac{5}{7} = 71$ percent

- problem: does not correct for chance
- most severe with presence of high-frequency, theoretically unimportant categories

The naive approach: percent agreement

Example: dichotomous coding 0/1, 2 coders

	1	2	3	4	5	6	7
Lars	0	0	1	0	0	0	0
Solrun	0	0	0	0	1	0	0

Percent agreement $= \frac{5}{7} = 71$ percent

- problem: does not correct for chance
- most severe with presence of high-frequency, theoretically unimportant categories

The naive approach: percent agreement

Example: dichotomous coding 0/1, 2 coders

	1	2	3	4	5	6	7
Lars	0	0	1	0	0	0	0
Solrun	0	0	0	0	1	0	0

Percent agreement $= \frac{5}{7} = 71$ percent

- problem: does not correct for chance
- most severe with presence of high-frequency, theoretically unimportant categories

The naive approach: percent agreement

Example: dichotomous coding 0/1, 2 coders

	1	2	3	4	5	6	7
Lars	0	0	1	0	0	0	0
Solrun	0	0	0	0	1	0	0

Percent agreement $= \frac{5}{7} = 71$ percent

- problem: does not correct for chance
- most severe with presence of high-frequency, theoretically unimportant categories

The naive approach: percent agreement

Example: dichotomous coding 0/1, 2 coders

	1	2	3	4	5	6	7
Lars	0	0	1	0	0	0	0
Solrun	0	0	0	0	1	0	0

Percent agreement $= \frac{5}{7} = 71$ percent

- problem: does not correct for chance
- most severe with presence of high-frequency, theoretically unimportant categories

The naive approach: percent agreement

Example: dichotomous coding 0/1, 2 coders

	1	2	3	4	5	6	7
Lars	0	0	1	0	0	0	0
Solrun	0	0	0	0	1	0	0

Percent agreement $= \frac{5}{7} = 71$ percent

- problem: does not correct for chance
- most severe with presence of high-frequency, theoretically unimportant categories

The naive approach: percent agreement

Example: dichotomous coding 0/1, 2 coders

	1	2	3	4	5	6	7
Lars	0	0	1	0	0	0	0
Solrun	0	0	0	0	1	0	0

Percent agreement $= \frac{5}{7} = 71$ percent

- problem: does not correct for chance
- most severe with presence of high-frequency, theoretically unimportant categories

The naive approach: percent agreement

Example: dichotomous coding 0/1, 2 coders

	1	2	3	4	5	6	7
Lars	0	0	1	0	0	0	0
Solrun	0	0	0	0	1	0	0

Percent agreement $= \frac{5}{7} = 71$ percent

- problem: does not correct for chance
- most severe with presence of high-frequency, theoretically unimportant categories

Krippendorff's α:

$$\alpha = 1 - \frac{D_{within\ units\ =\ in\ error}}{D_{within\ and\ between\ units\ =\ in\ total}} = 1 - \frac{D_o}{D_e} \tag{1}$$

where

$$\alpha_{metric} = 1 - \frac{D_o}{D_e} = 1 - \frac{\frac{1}{n}\sum_c \sum_k o_{ck}\delta_{metric}(c,k)}{\frac{1}{n(n-1)}\sum_c \sum_k n_c n_k\ \delta_{metric}(c,k)} \tag{2}$$

Krippendorff's α:

$$\alpha = 1 - \frac{D_{within\ units\ =\ in\ error}}{D_{within\ and\ between\ units\ =\ in\ total}} = 1 - \frac{D_o}{D_e} \tag{1}$$

where

$$\alpha_{metric} = 1 - \frac{D_o}{D_e} = 1 - \frac{\frac{1}{n}\sum_c \sum_k o_{ck}\delta_{metric}(c,k)}{\frac{1}{n(n-1)}\sum_c \sum_k n_c n_k\ \delta_{metric}(c,k)} \tag{2}$$

Krippendorff's α:

$$\alpha = 1 - \frac{D_{within\ units\ =\ in\ error}}{D_{within\ and\ between\ units\ =\ in\ total}} = 1 - \frac{D_o}{D_e} \tag{1}$$

where

$$\alpha_{metric} = 1 - \frac{D_o}{D_e} = 1 - \frac{\frac{1}{n}\sum_c \sum_k o_{ck}\delta_{metric}(c,k)}{\frac{1}{n(n-1)}\sum_c \sum_k n_c n_k\ \delta_{metric}(c,k)} \tag{2}$$

$$\alpha_{metric} = 1 - \frac{D_o}{D_e} = 1 - \frac{\frac{1}{n}\sum_c \sum_k o_{ck}\delta_{metric}(c,k)}{\frac{1}{n(n-1)}\sum_c \sum_k n_c n_k \ \delta_{metric}(c,k)}$$

o_{ck} and n_c/n_k are cell counts and row/column sums in a *coincidence matrix*:

	1	.	k	.	.
1	o_{11}	.	o_{1k}	.	n_1
.	.	.	.	.	.
c	o_{c1}	.	o_{ck}	.	n_c
.	.	.	.	.	.
	n_1	.	n_k	.	n

o_{ck}: count of codings assigned values c and k

$$\alpha_{metric} = 1 - \frac{D_o}{D_e} = 1 - \frac{\frac{1}{n}\sum_c \sum_k o_{ck}\delta_{metric}(c,k)}{\frac{1}{n(n-1)}\sum_c \sum_k n_c n_k \ \delta_{metric}(c,k)}$$

o_{ck} and n_c/n_k are cell counts and row/column sums in a *coincidence matrix*:

	1	.	k	.	.
1	o_{11}	.	o_{1k}	.	n_1
.	.	.	.	.	.
c	o_{c1}	.	o_{ck}	.	n_c
.	.	.	.	.	.
	n_1	.	n_k	.	n

o_{ck}: count of codings assigned values c and k

$$\alpha_{metric} = 1 - \frac{D_o}{D_e} = 1 - \frac{\frac{1}{n}\sum_c \sum_k o_{ck}\delta_{metric}(c,k)}{\frac{1}{n(n-1)}\sum_c \sum_k n_c n_k \ \delta_{metric}(c,k)}$$

o_{ck} and n_c/n_k are cell counts and row/column sums in a *coincidence matrix*:

	1	.	k	.	.
1	o_{11}	.	o_{1k}	.	n_1
.	.	.	.	.	.
c	o_{c1}	.	o_{ck}	.	n_c
.	.	.	.	.	.
	n_1	.	n_k	.	n

o_{ck}: count of codings assigned values c and k

$$\alpha_{metric} = 1 - \frac{D_o}{D_e} = 1 - \frac{\frac{1}{n}\sum_c \sum_k o_{ck}\delta_{metric}(c,k)}{\frac{1}{n(n-1)}\sum_c \sum_k n_c n_k \ \delta_{metric}(c,k)}$$

$\delta_{metric}(c,k)$ depends on level of measurement:

- Nominal data: $\delta_{nominal}(c,k) = \begin{cases} 0 & \text{iff } c = k \\ 1 & \text{iff } c \neq k \end{cases}$
- Ordinal data: $\delta_{ordinal}(c,k) = \left(\sum_{g=c}^{g=k} n_g - \frac{n_c + n_k}{2}\right)^2$
- Interval data: $\delta_{interval}(c,k) = (c-k)^2$

Note: for nominal data, $\sum_c \sum_k o_{ck}\delta_{metric}(c,k)$ reduces to sum of off-diagonal cells!

$$\alpha_{metric} = 1 - \frac{D_o}{D_e} = 1 - \frac{\frac{1}{n}\sum_c \sum_k o_{ck}\delta_{metric}(c,k)}{\frac{1}{n(n-1)}\sum_c \sum_k n_c n_k \ \delta_{metric}(c,k)}$$

$\delta_{metric}(c,k)$ depends on level of measurement:

- Nominal data: $\delta_{nominal}(c,k) = \begin{cases} 0 & \text{iff } c = k \\ 1 & \text{iff } c \neq k \end{cases}$
- Ordinal data: $\delta_{ordinal}(c,k) = \left(\sum_{g=c}^{g=k} n_g - \frac{n_c + n_k}{2}\right)^2$
- Interval data: $\delta_{interval}(c,k) = (c-k)^2$

Note: for nominal data, $\sum_c \sum_k o_{ck}\delta_{metric}(c,k)$ reduces to sum of off-diagonal cells!

$$\alpha_{metric} = 1 - \frac{D_o}{D_e} = 1 - \frac{\frac{1}{n}\sum_c \sum_k o_{ck}\delta_{metric}(c,k)}{\frac{1}{n(n-1)}\sum_c \sum_k n_c n_k \ \delta_{metric}(c,k)}$$

$\delta_{metric}(c,k)$ depends on level of measurement:

- Nominal data: $\delta_{nominal}(c,k) = \begin{cases} 0 & \text{iff } c = k \\ 1 & \text{iff } c \neq k \end{cases}$
- Ordinal data: $\delta_{ordinal}(c,k) = \left(\sum_{g=c}^{g=k} n_g - \frac{n_c + n_k}{2}\right)^2$
- Interval data: $\delta_{interval}(c,k) = (c-k)^2$

Note: for nominal data, $\sum_c \sum_k o_{ck}\delta_{metric}(c,k)$ reduces to sum of off-diagonal cells!

$$\alpha_{metric} = 1 - \frac{D_o}{D_e} = 1 - \frac{\frac{1}{n}\sum_c \sum_k o_{ck}\delta_{metric}(c,k)}{\frac{1}{n(n-1)}\sum_c \sum_k n_c n_k \ \delta_{metric}(c,k)}$$

$\delta_{metric}(c,k)$ depends on level of measurement:

- Nominal data: $\delta_{nominal}(c,k) = \begin{cases} 0 & \text{iff } c = k \\ 1 & \text{iff } c \neq k \end{cases}$
- Ordinal data: $\delta_{ordinal}(c,k) = \left(\sum_{g=c}^{g=k} n_g - \frac{n_c + n_k}{2}\right)^2$
- Interval data: $\delta_{interval}(c,k) = (c-k)^2$

Note: for nominal data, $\sum_c \sum_k o_{ck}\delta_{metric}(c,k)$ reduces to sum of off-diagonal cells!

$$\alpha_{metric} = 1 - \frac{D_o}{D_e} = 1 - \frac{\frac{1}{n}\sum_c \sum_k o_{ck}\delta_{metric}(c,k)}{\frac{1}{n(n-1)}\sum_c \sum_k n_c n_k \ \delta_{metric}(c,k)}$$

$\delta_{metric}(c,k)$ depends on level of measurement:

- Nominal data: $\delta_{nominal}(c,k) = \begin{cases} 0 & \text{iff } c = k \\ 1 & \text{iff } c \neq k \end{cases}$
- Ordinal data: $\delta_{ordinal}(c,k) = \left(\sum_{g=c}^{g=k} n_g - \frac{n_c + n_k}{2}\right)^2$
- Interval data: $\delta_{interval}(c,k) = (c-k)^2$

Note: for nominal data, $\sum_c \sum_k o_{ck}\delta_{metric}(c,k)$ reduces to sum of off-diagonal cells!

$$\alpha_{metric} = 1 - \frac{D_o}{D_e} = 1 - \frac{\frac{1}{n}\sum_c \sum_k o_{ck}\delta_{metric}(c,k)}{\frac{1}{n(n-1)}\sum_c \sum_k n_c n_k \ \delta_{metric}(c,k)}$$

$\delta_{metric}(c,k)$ depends on level of measurement:

- Nominal data: $\delta_{nominal}(c,k) = \begin{cases} 0 & \text{iff } c = k \\ 1 & \text{iff } c \neq k \end{cases}$
- Ordinal data: $\delta_{ordinal}(c,k) = \left(\sum_{g=c}^{g=k} n_g - \frac{n_c + n_k}{2}\right)^2$
- Interval data: $\delta_{interval}(c,k) = (c-k)^2$

Note: for nominal data, $\sum_c \sum_k o_{ck}\delta_{metric}(c,k)$ reduces to sum of off-diagonal cells!

	1	2	3	4	5	6	7
Lars	0	0	1	0	0	0	0
Solrun	0	0	0	0	1	0	0

$$\alpha_{nominal} = 1 - \frac{D_o}{D_e} = 1 - \frac{13}{12} = -.083$$

	1	2	3	4	5	6	7
Lars	0	0	1	0	0	0	0
Solrun	0	0	0	0	1	0	0

$$\alpha_{nominal} = 1 - \frac{D_o}{D_e} = 1 - \frac{13}{12} = -.083$$

Coding on 1-7 scale, 2 coders

	1	2	3	4	5	6	7
Margrethe	1	2	3	5	7	3	2
Henrik	1	2	2	5	6	4	3

Coincidence matrix:

	1	2	3	4	5	6	7
1	2						
2		2	2				
3		2		1			
4			1				
5					2		
6							1
7						1	

Coding on 1-7 scale, 2 coders

	1	2	3	4	5	6	7
Margrethe	1	2	3	5	7	3	2
Henrik	1	2	2	5	6	4	3

Coincidence matrix:

	1	2	3	4	5	6	7
1	2						
2		2	2				
3		2		1			
4			1				
5					2		
6							1
7						1	

Coding on 1-7 scale, 2 coders

	1	2	3	4	5	6	7
Margrethe	1	2	3	5	7	3	2
Henrik	1	2	2	5	6	4	3

Coincidence matrix:

	1	2	3	4	5	6	7
1	2						
2		2	2				
3		2		1			
4			1				
5					2		
6							1
7						1	

Coding on 1-7 scale, 2 coders

	1	2	3	4	5	6	7
Margrethe	1	2	3	5	7	3	2
Henrik	1	2	2	5	6	4	3

Coincidence matrix:

	1	2	3	4	5	6	7
1	2						
2		2	2				
3		2		1			
4			1				
5					2		
6							1
7						1	

Coding on 1-7 scale, 2 coders

	1	2	3	4	5	6	7
Margrethe	1	2	3	5	7	3	2
Henrik	1	2	2	5	6	4	3

Coincidence matrix:

	1	2	3	4	5	6	7
1	2						
2		2	2				
3		2		1			
4			1				
5					2		
6							1
7						1	

In R, package irr function kripp.alpha:

```
> kripp.alpha(comat,method="nominal")
Krippendorff's alpha
Subjects = 7
Raters = 2
alpha = 0.35
```

In R, package irr function kripp.alpha:

```
> kripp.alpha(comat,method="nominal")
Krippendorff's alpha
Subjects = 7
Raters = 2
alpha = 0.35
```

In R, package irr function kripp.alpha:

```
> kripp.alpha(comat,method="nominal")
Krippendorff's alpha
Subjects = 7
Raters = 2
alpha = 0.35
```

In R, package irr function kripp.alpha:

```
> kripp.alpha(comat,method="nominal")
Krippendorff's alpha
Subjects = 7
Raters = 2
alpha = 0.35
```

When assigning higher levels of measurement:

```
> kripp.alpha(comat,method="ordinal")
Krippendorff's alpha
Subjects = 7
Raters = 2
alpha = 0.875
> kripp.alpha(comat,method="interval")
Krippendorff's alpha
Subjects = 7
Raters = 2
alpha = 0.917
```

When assigning higher levels of measurement:

```
> kripp.alpha(comat,method="ordinal")
Krippendorff's alpha
Subjects = 7
Raters = 2
alpha = 0.875
> kripp.alpha(comat,method="interval")
Krippendorff's alpha
Subjects = 7
Raters = 2
alpha = 0.917
```

When assigning higher levels of measurement:

```
> kripp.alpha(comat,method="ordinal")
Krippendorff's alpha
Subjects = 7
Raters = 2
alpha = 0.875
> kripp.alpha(comat,method="interval")
Krippendorff's alpha
Subjects = 7
Raters = 2
alpha = 0.917
```

When assigning higher levels of measurement:

```
> kripp.alpha(comat,method="ordinal")
Krippendorff's alpha
Subjects = 7
Raters = 2
alpha = 0.875
> kripp.alpha(comat,method="interval")
Krippendorff's alpha
Subjects = 7
Raters = 2
alpha = 0.917
```

When assigning higher levels of measurement:

```
> kripp.alpha(comat,method="ordinal")
Krippendorff's alpha
Subjects = 7
Raters = 2
alpha = 0.875
> kripp.alpha(comat,method="interval")
Krippendorff's alpha
Subjects = 7
Raters = 2
alpha = 0.917
```

Exercise 4

Assume reasonable reliability for the data coded in Ex. 3. How would you analyze it? What would be your testable hypothesis?

Exercise 4

Assume reasonable reliability for the data coded in Ex. 3. How would you analyze it? What would be your testable hypothesis?

Four principles for automated text analysis, from Grimmer & Stewart (2013)

1. All quantitative models of language are wrong—but some are useful.
2. Quantitative methods augment humans, they do not replace them.
3. There is no globally best method for automated text analysis.
4. Validate, validate, validate.

Four principles for automated text analysis, from Grimmer & Stewart (2013)

1. All quantitative models of language are wrong—but some are useful.
2. Quantitative methods augment humans, they do not replace them.
3. There is no globally best method for automated text analysis.
4. Validate, validate, validate.

Four principles for automated text analysis, from Grimmer & Stewart (2013)

1. All quantitative models of language are wrong—but some are useful.
2. Quantitative methods augment humans, they do not replace them.
3. There is no globally best method for automated text analysis.
4. Validate, validate, validate.

Four principles for automated text analysis, from Grimmer & Stewart (2013)

1. All quantitative models of language are wrong—but some are useful.
2. Quantitative methods augment humans, they do not replace them.
3. There is no globally best method for automated text analysis.
4. Validate, validate, validate.

Four principles for automated text analysis, from Grimmer & Stewart (2013)

1. All quantitative models of language are wrong—but some are useful.
2. Quantitative methods augment humans, they do not replace them.
3. There is no globally best method for automated text analysis.
4. Validate, validate, validate.

Typical tasks:

- Assign texts positions on a (left-right) scale: Wordscores/Wordfish
- Assign texts values on a variable: dictionary approaches (e.g., Lexicoder)
- Characterize the distribution of categories across texts: ReadMe
- Characterize the unknown distribution of topics across and within texts: topic models

Typical tasks:

- Assign texts positions on a (left-right) scale: Wordscores/Wordfish
- Assign texts values on a variable: dictionary approaches (e.g., Lexicoder)
- Characterize the distribution of categories across texts: ReadMe
- Characterize the unknown distribution of topics across and within texts: topic models

Typical tasks:

- Assign texts positions on a (left-right) scale: Wordscores/Wordfish
- Assign texts values on a variable: dictionary approaches (e.g., Lexicoder)
- Characterize the distribution of categories across texts: ReadMe
- Characterize the unknown distribution of topics across and within texts: topic models

Typical tasks:

- Assign texts positions on a (left-right) scale: Wordscores/Wordfish
- Assign texts values on a variable: dictionary approaches (e.g., Lexicoder)
- Characterize the distribution of categories across texts: ReadMe
- Characterize the unknown distribution of topics across and within texts: topic models

Typical tasks:

- Assign texts positions on a (left-right) scale: Wordscores/Wordfish
- Assign texts values on a variable: dictionary approaches (e.g., Lexicoder)
- Characterize the distribution of categories across texts: ReadMe
- Characterize the unknown distribution of topics across and within texts: topic models

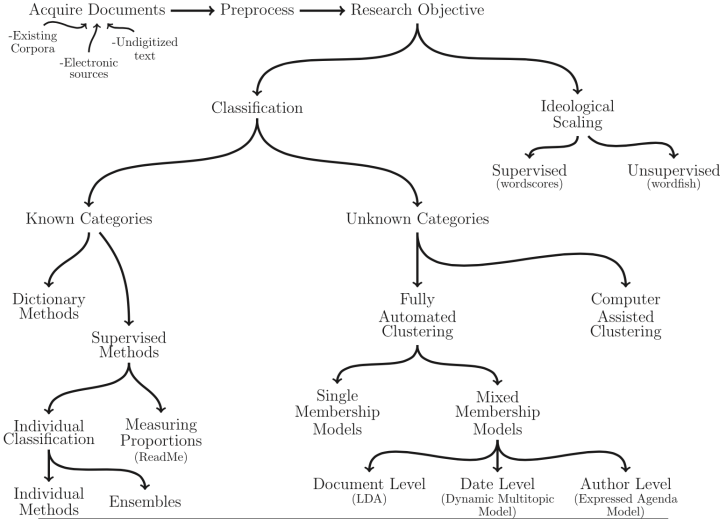
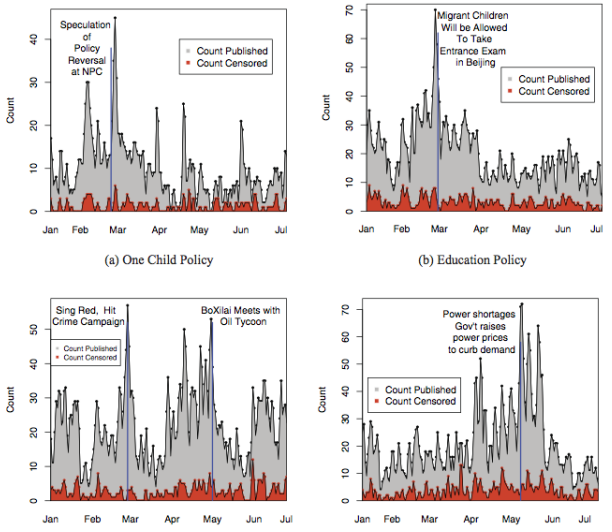

Acquire Documents
Preprocess
Research Objective
-Existing Corpora
-Electronic sources
-Undigitized text
Classification
Ideological Scaling
Supervised (wordscores)
Unsupervised (wordfish)
Known Categories
Unknown Categories
Dictionary Methods
Supervised Methods
Fully Automated Clustering
Computer Assisted Clustering
Individual Classification
Measuring Proportions (ReadMe)
Single Membership Models
Mixed Membership Models
Individual Methods
Ensembles
Document Level (LDA)
Date Level (Dynamic Multitopic Model)
Author Level (Expressed Agenda Model)

Figure 1. The Fractured Structure of the Chinese Social Media Landscape

(a) Sample of Sites

(b) All Sites excluding Sina

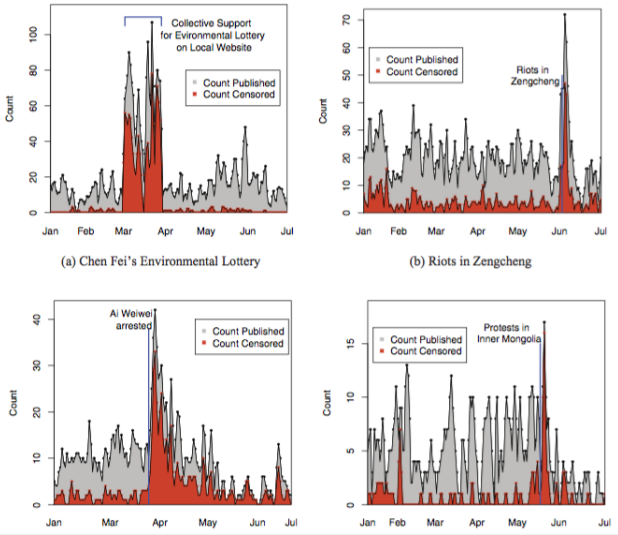
Figure 6. Low Censorship on News and Policy Events (in 2011)



(a) One Child Policy

(b) Education Policy

Figure 5. High Censorship During Collective Action Events (in 2011)



(a) Chen Fei's Environmental Lottery

(b) Riots in Zengcheng

Reliability of event coding?

»we conducted a study to verify the reliability of our event coding rules.To do this, we gave our rules above to two people familiar with Chinese politics and asked them to code each of the eighty-seven events (each associated with a volume burst) into one of the five categories.The coders worked independently and classified each of the events on their own. Decisions by the two coders agreed in 98.9% (i.e., eighty-six of eighty-seven) of the events.« - King et al. (2013)

Reliability of event coding?

»we conducted a study to verify the reliability of our event coding rules. To do this, we gave our rules above to two people familiar with Chinese politics and asked them to code each of the eighty-seven events (each associated with a volume burst) into one of the five categories. The coders worked independently and classified each of the events on their own. Decisions by the two coders agreed in 98.9% (i.e., eighty-six of eighty-seven) of the events.« - King et al. (2013)

Reliability of event coding?

»we conducted a study to verify the reliability of our event coding rules. To do this, we gave our rules above to two people familiar with Chinese politics and asked them to code each of the eighty-seven events (each associated with a volume burst) into one of the five categories. The coders worked independently and classified each of the events on their own. Decisions by the two coders agreed in 98.9% (i.e., eighty-six of eighty-seven) of the events.« - King et al. (2013)

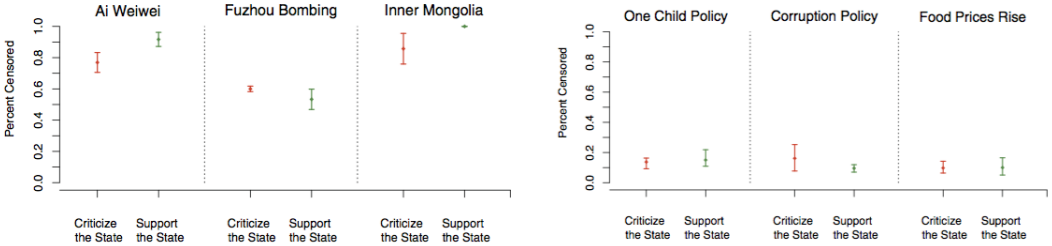
Reliability of event coding?

»we conducted a study to verify the reliability of our event coding rules. To do this, we gave our rules above to two people familiar with Chinese politics and asked them to code each of the eighty-seven events (each associated with a volume burst) into one of the five categories. The coders worked independently and classified each of the events on their own. Decisions by the two coders agreed in 98.9% (i.e., eighty-six of eighty-seven) of the events.« - King et al. (2013)

Reliability of event coding?

»we conducted a study to verify the reliability of our event coding rules. To do this, we gave our rules above to two people familiar with Chinese politics and asked them to code each of the eighty-seven events (each associated with a volume burst) into one of the five categories. The coders worked independently and classified each of the events on their own. Decisions by the two coders agreed in 98.9% (i.e., eighty-six of eighty-seven) of the events.« - King et al. (2013)

Figure 8. Content of Censored Posts by Topic Area



Thanks for now!