The Literature Review Seminar

Steps of the process

- Understand the generic steps of the review process
- Appreciate the critical methodological choices in the search, screen, and analysis

ork/

Generic steps: Examples

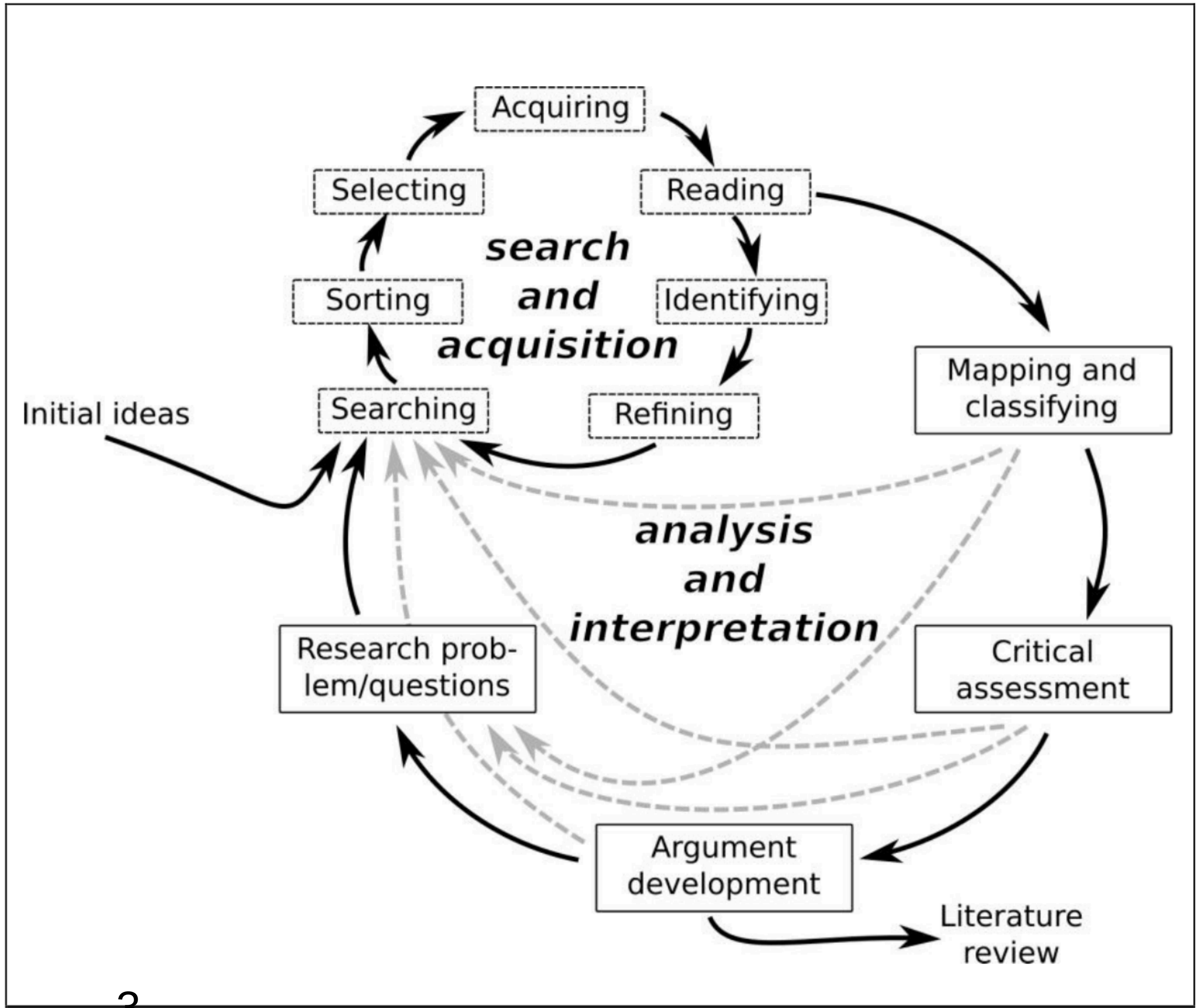


Figure 1. A hermeneutic framework for the literature review

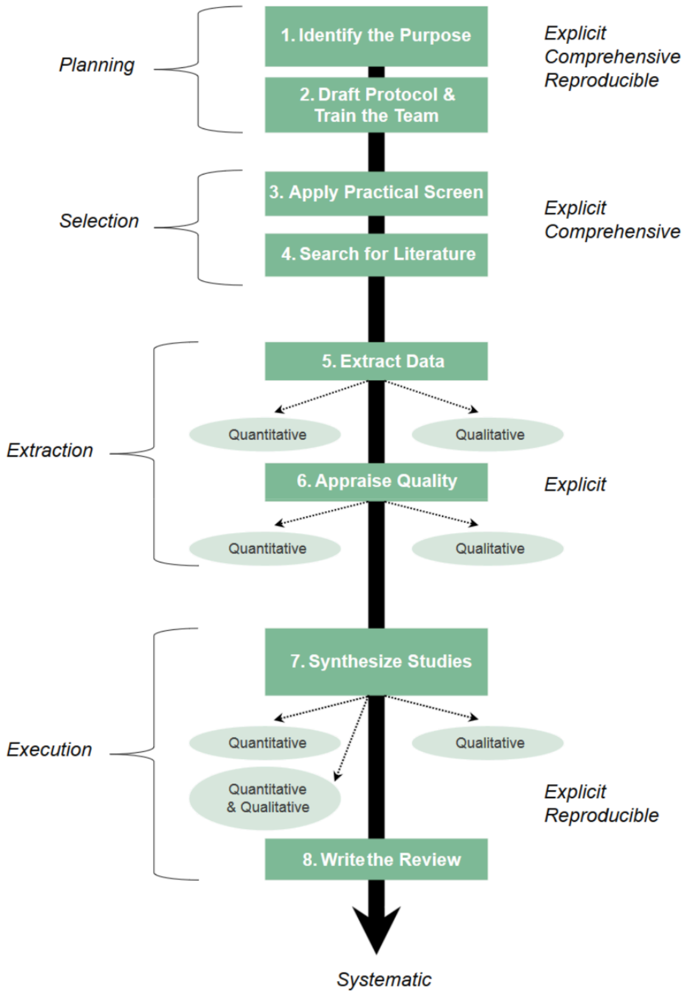


Figure 1. A Systematic Guide to Literature Review Development

Table 8. Frequency of reporting items per review type.

	Narrative ($n = 25$)	Descriptive ($n = 22$)	Scoping ($n = 9$)	Critical ($n = 16$)	Meta-analysis ($n = 12$)	Qualitative systematic ($n = 6$)	Theory development ($n = 52$)
Step 1: Problem formulation							
Primary goals or research questions	100	100	100	100	100	100	100
Key concepts or theories being investigated	84	91	89	94	100	100	100
Step 2: Literature search							
How the literature search is performed	44	95	89	69	100	100	25
Multiple search strategies			44	19	100	33	13
Multiple publication types			22	13	92	33	17
Comprehensiveness of search & restrictions if applicable		86	78	50	100	83	21
How reputation of the sources is considered	28			63			13
Strategies used to minimise publication bias					25	0	
Step 3: Screening for inclusion							
How primary studies are screened or selected	20	91	67		67	67	21
Results of parallel independent study		5	11		8	0	

Table 8. Frequency of reporting items per review type.

	Narrative ($n = 25$)	Descriptive ($n = 22$)	Scoping ($n = 9$)	Critical ($n = 16$)	Meta-analysis ($n = 12$)	Qualitative systematic ($n = 6$)	Theory development ($n = 52$)
Step 4: Quality assessment							
How quality assessment is performed					8	0	
Results of parallel independent assessment					8	0	
Step 5: Data extraction							
Data extraction plan		95	56		92	100	
Tools or methods used to extract data		77	67		83	83	10
Results of parallel independent coding process		41	33		67	67	
Step 6: Data analysis and interpretation							
How data analysis is performed			56		100	83	19
How study quality is considered in interpretation of findings					0	0	
Profile of the included studies		91	67		75	67	
Justification of data			22		100	67	

ork/

Generic steps

Summary

- The **nature of steps varies**, including their labels, their characteristics, and how they are arranged
- The steps **depend on the review type**
- Some steps are more **generic**, while others are more **specific** and only apply to selected types of reviews

In the following, we focus on the steps summarized by Templier and Paré (2018):

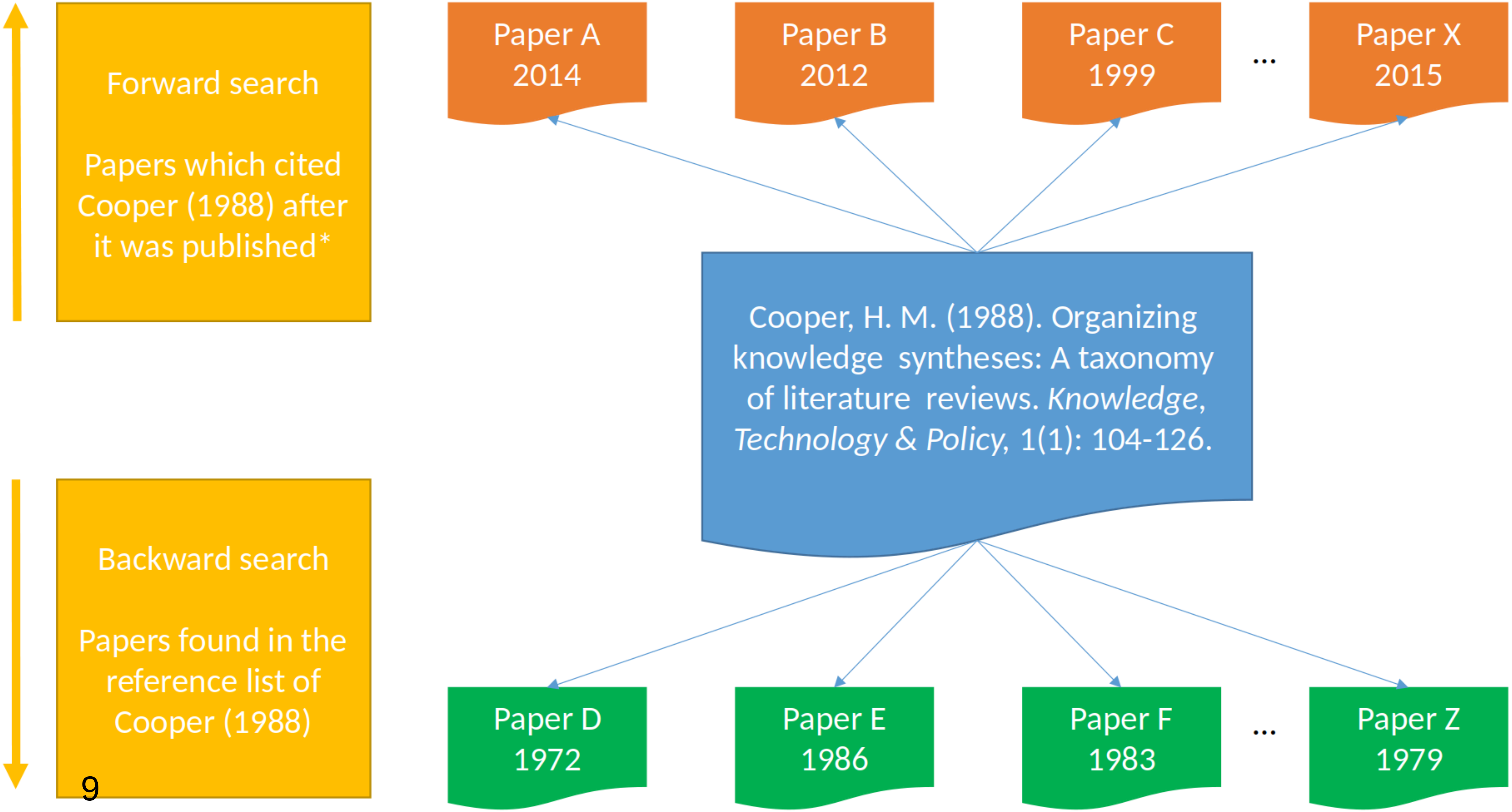
1. Problem formulation
2. Literature search
3. Screening for inclusion
4. Quality assessment
5. Data extraction
6. Data analysis and interpretation

Problem formulation

- Rationale for the review, including an overview of related review papers
- Gap-spotting or problematization (Alvesson and Sandberg 2011):
  - Gap-spotting is seen as (too) common, and may only signify a contribution if the authors can make a convincing argument that the gap is important
  - Problematization, as an approach that challenges existing theory and the underlying assumptions, can lead to more interesting and noteworthy contributions
- Research question or objectives

Literature search: Foundations

- Search types: *Lookup* vs. *exploratory* vs. **systematic search** (Gusenbauer and Haddaway 2021)
- Search scope: time, journals, and academic vs. gray literature
- Search techniques (with associated search sources):
  - Database search (keyword-based)
  - Backward search, i.e., search reference sections to go back in time (aka. snowballing, pearl-growing)
  - Forward search, i.e., using citation indices to go forward in time
  - Table-of-content search (whole journals)
  - Sampling from prior review papers
  - Consulting with peers (e.g., through direct contact or mailing lists)


Forward search
Papers which cited Cooper (1988) after it was published*
Backward search
Papers found in the reference list of Cooper (1988)
Paper A 2014
Paper B 2012
Paper C 1999
...
Paper X 2015
Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge, Technology & Policy,* 1(1): 104-126.
Paper D 1972
Paper E 1986
Paper F 1983
...
Paper Z 1979

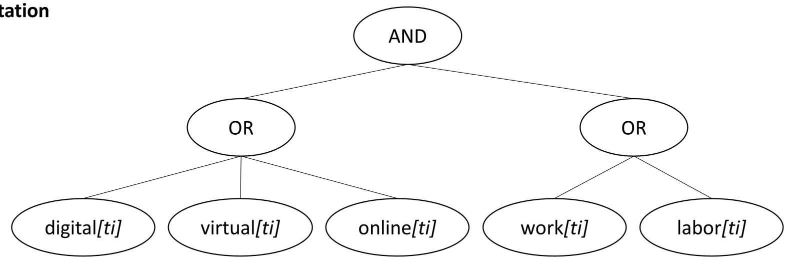
Literature search: The database search

- Most common search strategy in the management disciplines (58% according to Hiebl, 2023)
- Common databases: Web of Science, EBSCO Host, ABI Informs, AIS eLibrary, ACM Digital Library, IEEEXplore, etc. (Knackstedt and Winkelmann 2006, Hiebl 2023)
- Effective search strategies for database searches combine search terms with Boolean operators

String representation

(digital OR virtual OR online) AND (work OR labor) *[in title field]*

Tree representation



List representation

1	digital *[ti]*
2	virtual *[ti]*
3	online *[ti]*
4	1 OR 2 OR 3
5	work *[ti]*
6	labor *[ti]*
7	5 OR 6
8	4 AND 7

For a comprehensive overview, see searchsmart.org.

Literature search: The "building-blocks" approach

- RQ: What factors do influence physicians' acceptance of telemedicine?

Synonyms	Concept 1	AND	Concept 2	AND	Concept 3
	Telemedicine		Physician		Acceptance
OR	Telehealth		Doctor		Adoption
OR	Teleconsultation		Clinician		Resistance
OR	Tele-expertise				
OR					
OR					
OR					

Resulting search string: `(telemedicine OR telehealth OR ...) AND (physician OR doctor OR ...) AND (adoption OR`

Literature search: Application

- Draft a search strategy for your topic, following the building-blocks approach.

Concepts →

Synonyms ↓

- Does the building block approach provide a good fit with your context?

Exercise: Reviewing a search strategy

Imagine you serve as a reviewer for a conference. You review a paper on algorithmic decision-making, along with Table 2.

Task: Evaluate the proposed search strategy critically, taking into account the building-block approach. Make a recommendation to accept, revise, or reject.

Table 2
Search terms.

Themes of search term: algorithm, artificial intelligence, and machine learning; decision, advice, recommendation, and decision aid

ID	Search syntax	Total hits	Ultimately retained**
1	Algorithm* Aversion	840	16
2	Algorithm* Appreciation	788	3
3	(AI OR "Artificial Intelligence") AND Aversion	162	2
4	(AI OR "Artificial Intelligence") AND Appreciation	197	–
5	"AI recommendation" OR "Artificial intelligence recommendation" OR "Algorithm* recommendations" OR "Machine Learning recommendation" OR "ML recommendation"	249	1
6	"AI decision*" OR "Artificial intelligence decision*" OR "Algorithm* decision*" OR "Machine Learning decision*" OR "ML decision*"	2009	3
7	"AI Advice" OR "Artificial intelligence Advice" OR "Algorithm* advice" OR "Machine Learning advice" OR "ML advice"	15	3
8	("AI" OR "Artificial intelligence" OR "Algorithm*" OR "Machine Learning" OR "ML") AND "Decision aid"	697	5

* Takes the place of one or more characters in the search term.
** Figures represent the number of studies after completing selection process.

* Note: Example taken from Mahmud, H., Islam, A. N., Ahmed, S. I., & Smolander, K. (2022). What influences algorithmic decision-making? A systematic literature review on algorithm aversion. *Technological Forecasting and Social Change*, 175, 121390.

The search-query package supports the validation (linting) of search queries to identify syntactical errors and suggest improvements.

Literature search: Strengths and shortcomings of database searches

Strengths:

- Relatively efficient (see Wagner et al. 2021, Appendix A3)
- Transparent and reproducible

Shortcomings:

- Keyword searches rely on exact matches *
- Need to be familiar with the vocabulary (check keywords or taxonomies like MeSH etc.)
- Assumption of controlled scientific vocabulary although disciplines like Information Systems have abandoned corresponding efforts decades ago (Weber 2003)
- Some literature reviews report the intended coverage (e.g., comprehensive), but none report to which degree it was accomplished (using evidence and metrics)

> * This is why the health sciences strictly enforce the use of descriptive titles and standard terminology in primary studies.

Literature search: Search metrics

The common objective is to identify all relevant papers. Literature searches retrieve documents:

	Relevant	Not relevant	
Retrieved	True positive	False positive	
Not retrieved	False negative	True negative	unknown

Three key metrics are particularly relevant in the context of literature searches (Cooper et al. 2018):

1. **Sensitivity** aka. recall: $TP/(TP+FN)$. *How many of the relevant papers do we find?* ❓

Literature search: Assessing searches

- **Precision** is the only metric that can be measured in a typical literature review
- A **highly precise search strategy should be suspicious** because the search may not be comprehensive enough
- Based on the SYNERGY dataset, average precision is 2% - 4% in medicine, chemistry, and computer science

Question: Would you expect more precise searches in disciplines like Information Systems, Management, or the Social Sciences?

Datasets and variables

The SYNERGY dataset comprises the study selection of 26 systematic reviews. The dataset contains 169,288 records of which 2,834 records are manually labeled as inclusion by the authors of the systematic review. The list of systematic review and basic properties:

Nr	Dataset	Topic(s)	Records	Included	%
1	Appenzeller-Herzog_2019	Medicine	2873	26	0.9
2	Bos_2018	Medicine	4878	10	0.2
3	Brouwer_2019	Psychology, Medicine	38114	62	0.2
4	Chou_2003	Medicine	1908	15	0.8
5	Chou_2004	Medicine	1630	9	0.6
6	Donners_2021	Medicine	258	15	5.8
7	Hall_2012	Computer science	8793	104	1.2
8	Jeyaraman_2020	Medicine	1175	96	8.2
9	Leenaars_2019	Psychology, Chemistry, Medicine	5812	17	0.3

Terminating the search

No formal "stopping rule" exists — reviewers look for **completeness, transparency**, and **justifiability**.

Often guided by the **criterion of saturation**:

- Saturation in understanding *(Boell & Cecez-Kecmanovic, 2014)*
- "New articles only introduce familiar arguments, methodologies, findings, authors, and studies" *(Levy and Ellis, 2006)*
- Theoretical saturation *(Wolfswinkel et al., 2013)*

Credibility depends on:

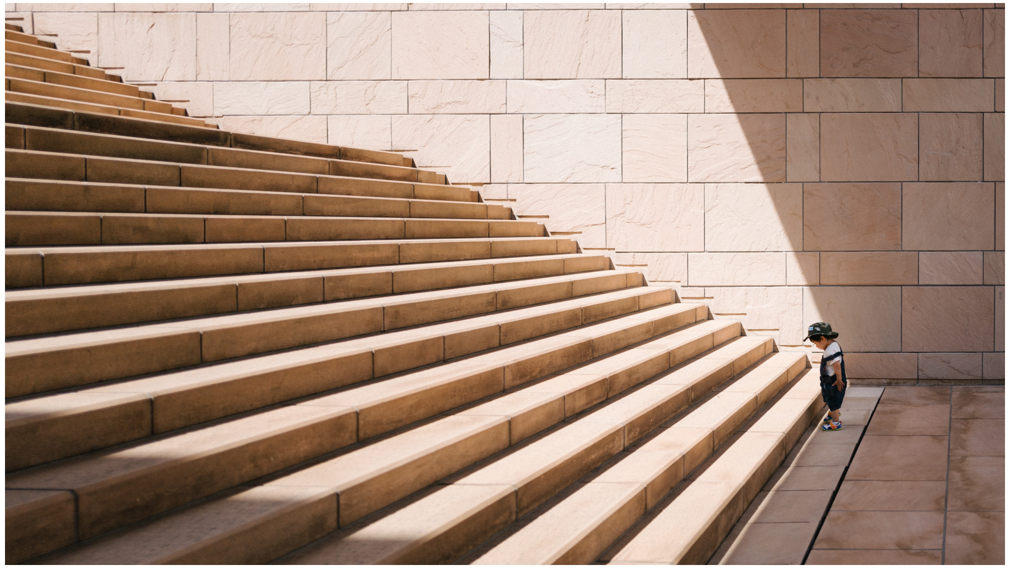
- **Comprehensiveness** of the applied **search techniques**
- Alignment with the topic's **epistemological context**:
  - **High-paradigm disciplines**: database search may suffice
  - **Low-paradigm or emergent fields**: require iterative, **citation-based strategies** (e.g., forward/backward search)

> "Saturation is a matter of judgment, not of formula."
> — *Boell & Cecez-Kecmanovic (2014)*

Literature search: Outlook

Open challenge:

- How can we iterate efficiently?
- How do we justify the decision to terminate a search?
- How can we use evidence to search effectively?
- How can we make progress without database providers?



Screen

- The screen is typically completed in two parts:
  - A pre-screen based on metadata (*"include if in doubt"*)
  - A screen based on full-text documents, resulting in the final sample
- The screen is often based on explicit inclusion and exclusion criteria

Study Selection

Studies were included if (1) a randomized controlled trial (RCT) design was used, (2) the intervention involved using a Fitbit device to improve PA and/or other health-related outcomes (eg, weight loss), and (3) the study reported outcomes related to healthy lifestyle measures (eg, steps, MVPA, weight, and BMI). Only peer-reviewed journal and conference papers were considered.

Articles were screened in a two-step process. First, all titles and abstracts were examined by one author (MR). Any citations that clearly did not meet the inclusion criteria were excluded. Second, all abstracts and full-text articles were examined independently by two authors (MR and GW). Any disagreements in the selection process were resolved through discussion with a third author (GP or SK).

Screening reliability

Screening tasks are often split among the review team to complete the process **more quickly**, and to ensure **reliable decisions**.

Process:

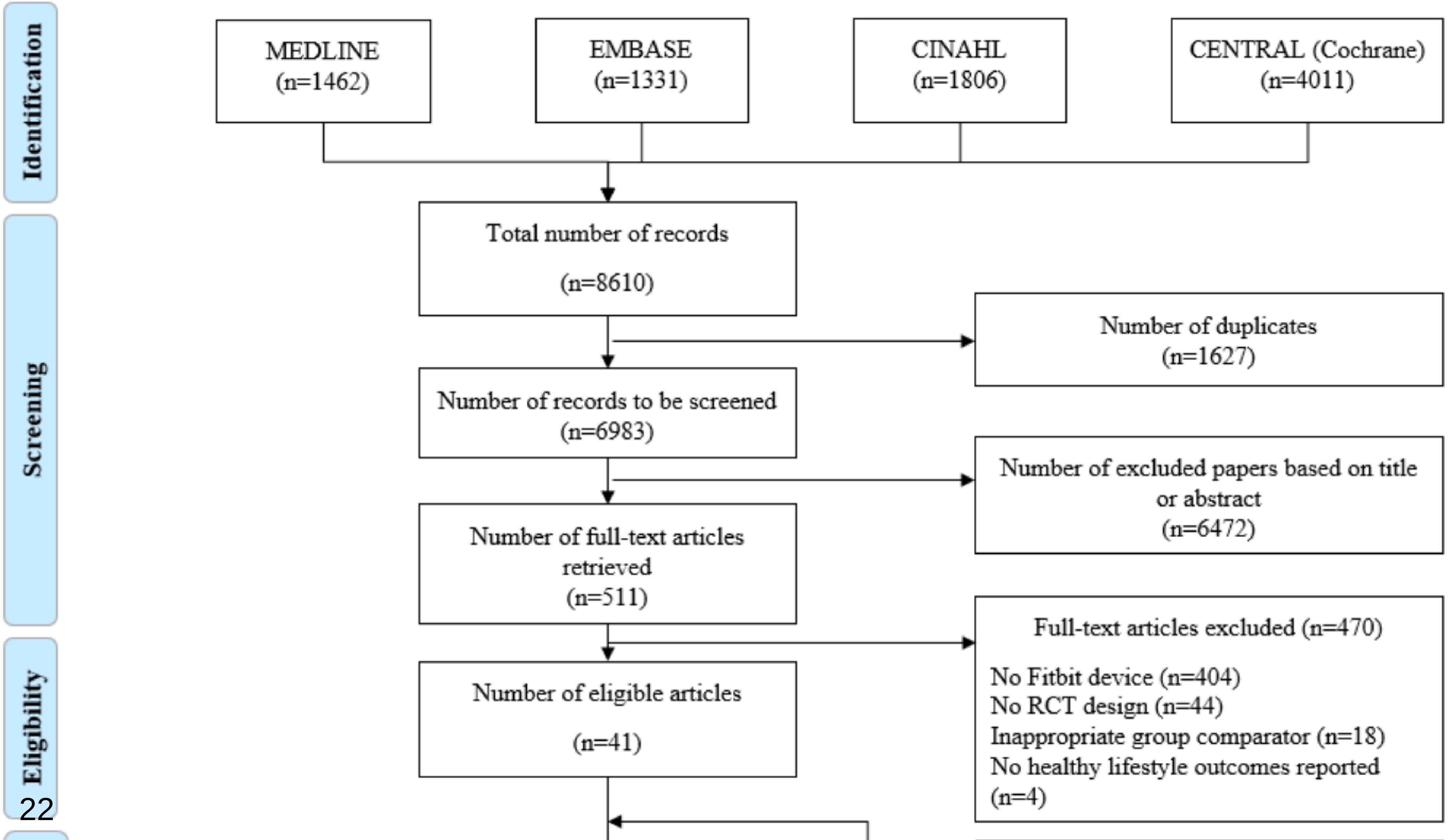
1. Definition of criteria, training, and pilot test
2. Parallel-independent screen (partially or fully overlapping sample)
3. Independent screen of the remaining papers (if any)
4. Reconciliation: in case of disagreements, final decisions are made by selected team members (often more senior researchers)
5. Calculate inter-rater agreement (e.g., Cohen's Kappa) and report the process

Interpretation of Kappa Values

Kappa Value Range	Interpretation
≤ 0	No agreement
0.01 – 0.20	None to slight
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Substantial
0.81 – 1.00	Almost perfect agreement

Note: When data is skewed—meaning one category is much more common than others—the Kappa statistic can be artificially low even if there is a high level of agreement. This occurs because Kappa adjusts for the level of agreement that would be expected purely by chance. In skewed distributions, the expected chance agreement tends to be high, which lowers the Kappa score. Essentially, in skewed distributions, even a relatively high observed agreement may not lead to a high Kappa, as the metric accounts for the imbalance.

Figure 1. Flow diagram. RCT: randomized controlled trial.



Break

Reading strategies

The reading activities can be organized strategically at two levels:

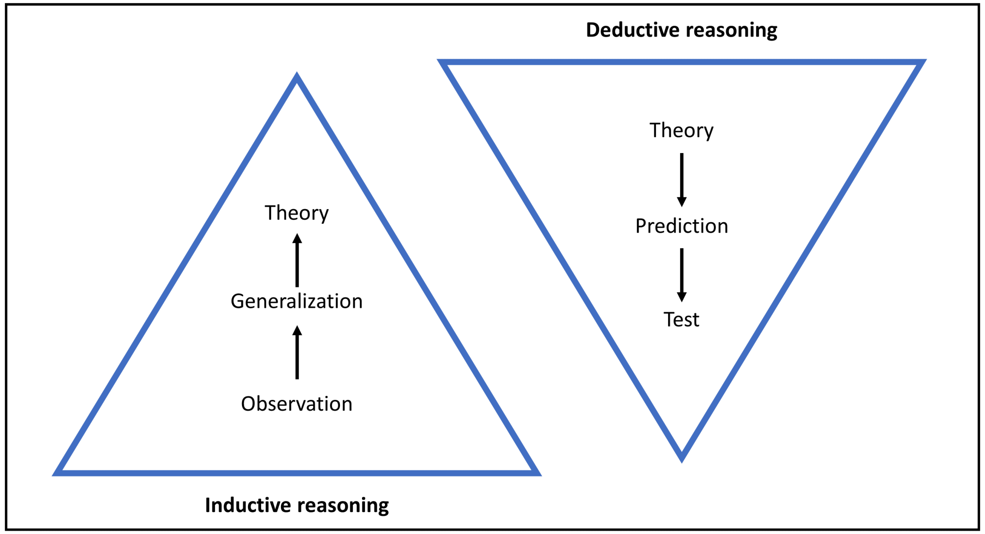
- The overall corpus level: In which order should papers be read or skimmed?
- The individual paper level: How should the different parts of a paper be read?

Question: Assume you have 300 papers to cover, how would you organize the reading activities?

Data analysis

Key differences with regard to data extraction and analysis:

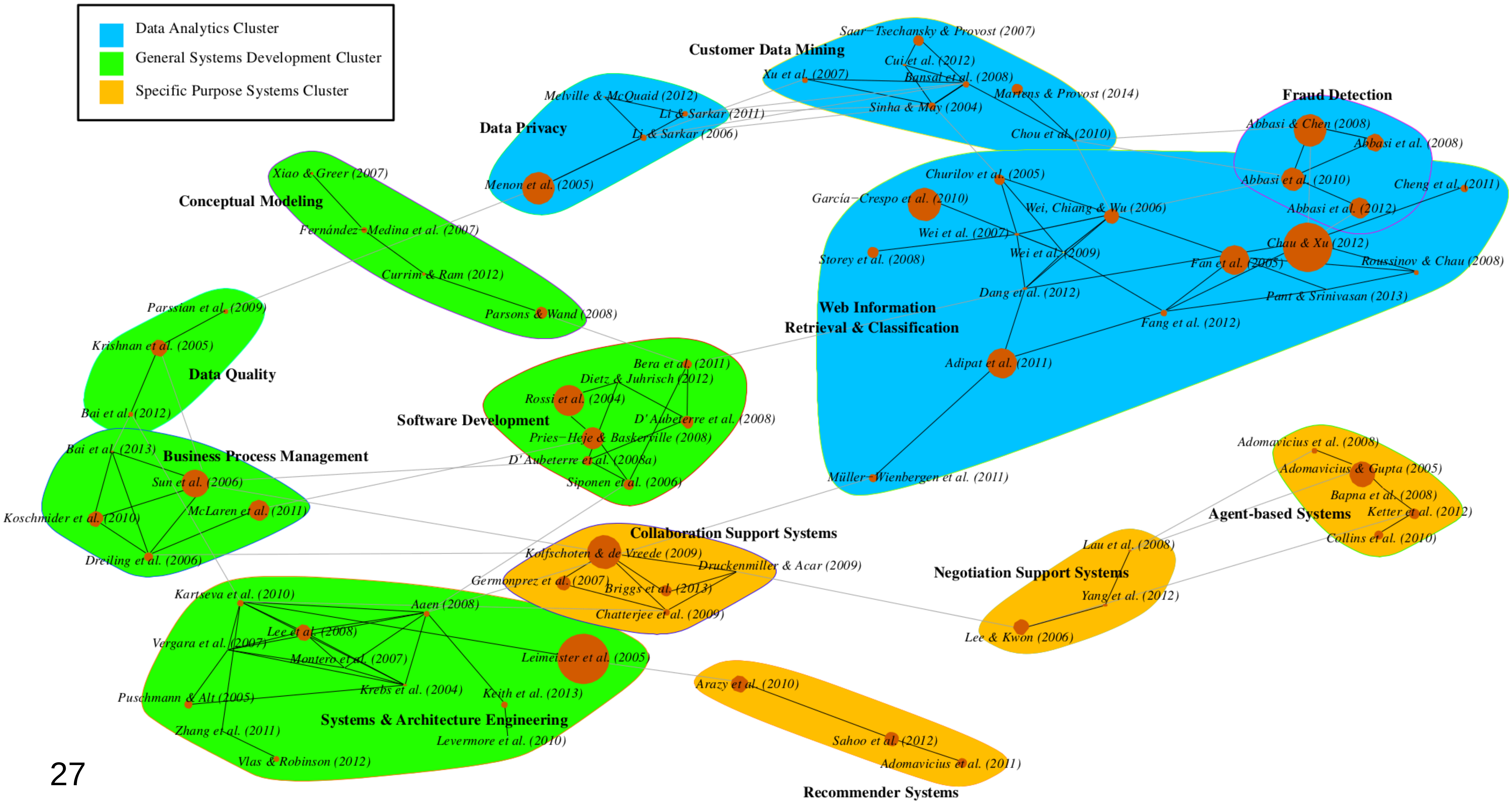
- Focus on metadata vs content
- Inductive vs deductive reasoning



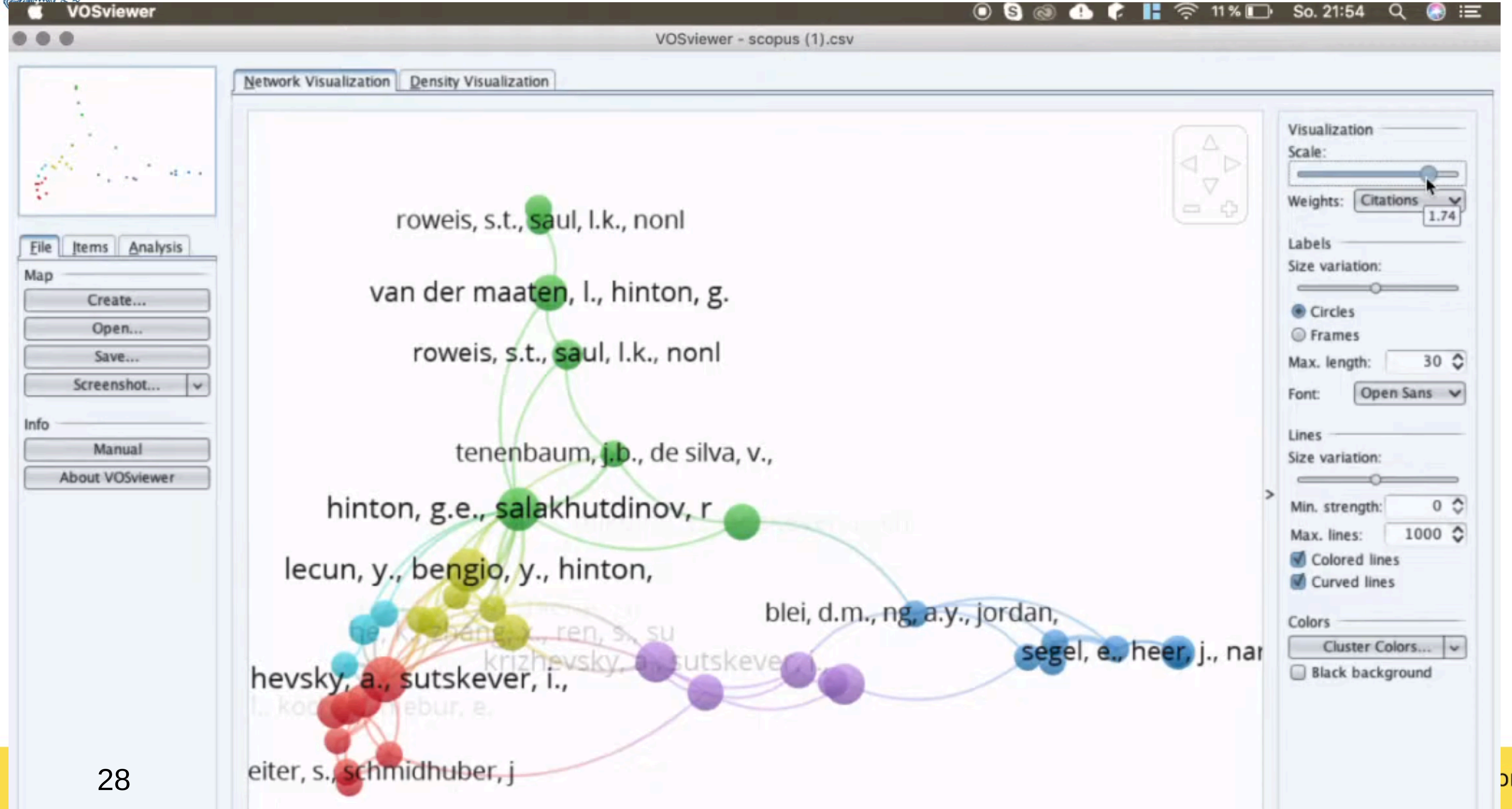
Note: The distinction between inductive and deductive modes of reasoning has critical implications. For instance, it would be incoherent to present an inductive analysis with inter-coder reliability assessment, or a deductive analysis without a pre-defined coding schema.

Table A.2
Frequency table.

Journals and conferences	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	Sum
Decision Support Systems	–	1	–	–	1	–	–	–	1	–	–	–	–	3
e-Service Journal	–	1	–	–	–	–	–	–	–	–	–	–	–	1
Electronic Markets	–	–	–	–	–	–	–	–	–	–	–	1	–	1
Information & Management	–	–	–	–	–	–	–	1	–	–	–	–	–	1
Information Systems Frontiers	–	–	–	–	–	–	–	1	–	1	–	–	1	3
Information Systems Journal	–	–	–	–	–	–	–	–	–	–	–	2	–	2
Information Systems Research	–	–	–	–	–	–	–	–	1	1	1	–	1	4
International Journal of Electronic Commerce	–	–	–	–	–	–	–	–	–	–	–	–	1	1
Journal of Management Information Systems	–	–	–	–	–	–	–	–	–	–	1	–	–	1
MIS Quarterly	1	–	–	–	–	–	–	–	–	–	–	–	–	1
MIS Quarterly Executive	–	–	–	–	–	–	–	–	–	–	1	–	–	1
The Journal of Strategic Information Systems	–	–	–	–	–	–	–	–	–	1	1	1	–	3
Americas Conference on Information Systems	–	–	–	1	–	–	–	1	–	–	1	1	–	4
European Conference on Information Systems	–	–	–	1	–	–	–	–	–	–	–	1	–	2
Hawaii Int. Conference on System Sciences	–	–	–	–	–	–	–	–	–	1	3	–	–	4
International Conference on Information Systems	–	–	–	–	1	–	2	2	1	2	4	2	*	14
Pacific Asia Conference on Information Systems	–	–	–	–	–	–	–	–	–	1	1	–	1	3



Data Analytics Cluster
General Systems Development Cluster
Specific Purpose Systems Cluster
Customer Data Mining
Saar−Tsechansky & Provost (2007)
Cui et al. (2012)
Xu et al. (2007)
Bansal et al. (2008)
Martens & Provost (2014)
Sinha & May (2004)
Chou et al. (2010)
Data Privacy
Melville & McQuaid (2012)
Li & Sarkar (2011)
Li & Sarkar (2006)
Menon et al. (2005)
Fraud Detection
Abbasi & Chen (2008)
Abbasi et al. (2008)
Abbasi et al. (2010)
Cheng et al. (2011)
Abbasi et al. (2012)
Chau & Xu (2012)
Roussinov & Chau (2008)
Conceptual Modeling
Xiao & Greer (2007)
Fernández−Medina et al. (2007)
Currim & Ram (2012)
Parsons & Wand (2008)
Churilov et al. (2005)
García−Crespo et al. (2010)
Wei, Chiang & Wu (2006)
Wei et al. (2007)
Storey et al. (2008)
Wei et al. (2009)
Fan et al. (2005)
Dang et al. (2012)
Pant & Srinivasan (2013)
Fang et al. (2012)
Web Information Retrieval & Classification
Adipat et al. (2011)
Müller−Wienbergen et al. (2011)
Data Quality
Parssian et al. (2009)
Krishnan et al. (2005)
Bai et al. (2012)
Software Development
Bera et al. (2011)
Dietz & Juhrisch (2012)
Rossi et al. (2004)
D'Aubeterre et al. (2008)
Pries−Heje & Baskerville (2008)
D'Aubeterre et al. (2008a)
Siponen et al. (2006)
Business Process Management
Bai et al. (2013)
Sun et al. (2006)
Koschmider et al. (2010)
McLaren et al. (2011)
Dreiling et al. (2006)
Agent-based Systems
Adomavicius et al. (2008)
Adomavicius & Gupta (2005)
Bapna et al. (2008)
Ketter et al. (2012)
Collins et al. (2010)
Collaboration Support Systems
Kolfschoten & de Vreede (2009)
Druckenmiller & Acar (2009)
Germonprez et al. (2007)
Briggs et al. (2013)
Chatterjee et al. (2009)
Negotiation Support Systems
Lau et al. (2008)
Yang et al. (2012)
Lee & Kwon (2006)
Systems & Architecture Engineering
Kartseva et al. (2010)
Aaen (2008)
Lee et al. (2008)
Vergara et al. (2007)
Montero et al. (2007)
Leimeister et al. (2005)
Krebs et al. (2004)
Keith et al. (2013)
Puschmann & Alt (2005)
Zhang et al. (2011)
Levermore et al. (2010)
Vlas & Robinson (2012)
Recommender Systems
Arazy et al. (2010)
Sahoo et al. (2012)
Adomavicius et al. (2011)


VOSviewer
So. 21:54
VOSviewer - scopus (1).csv
Network Visualization
Density Visualization
File
Items
Analysis
Map
Create...
Open...
Save...
Screenshot...
Info
Manual
About VOSviewer
roweis, s.t., saul, l.k., nonl
van der maaten, l., hinton, g.
roweis, s.t., saul, l.k., nonl
tenenbaum, j.b., de silva, v.,
hinton, g.e., salakhutdinov, r
lecun, y., bengio, y., hinton,
blei, d.m., ng, a.y., jordan,
segel, e., heer, j., nar
hevsky, a., sutskever, i.,
eiter, s., schmidhuber, j
Visualization
Scale:
Weights:
Citations
1.74
Labels
Size variation:
Circles
Frames
Max. length:
30
Font:
Open Sans
Lines
Size variation:
Min. strength:
0
Max. lines:
1000
Colored lines
Curved lines
Colors
Cluster Colors...
Black background

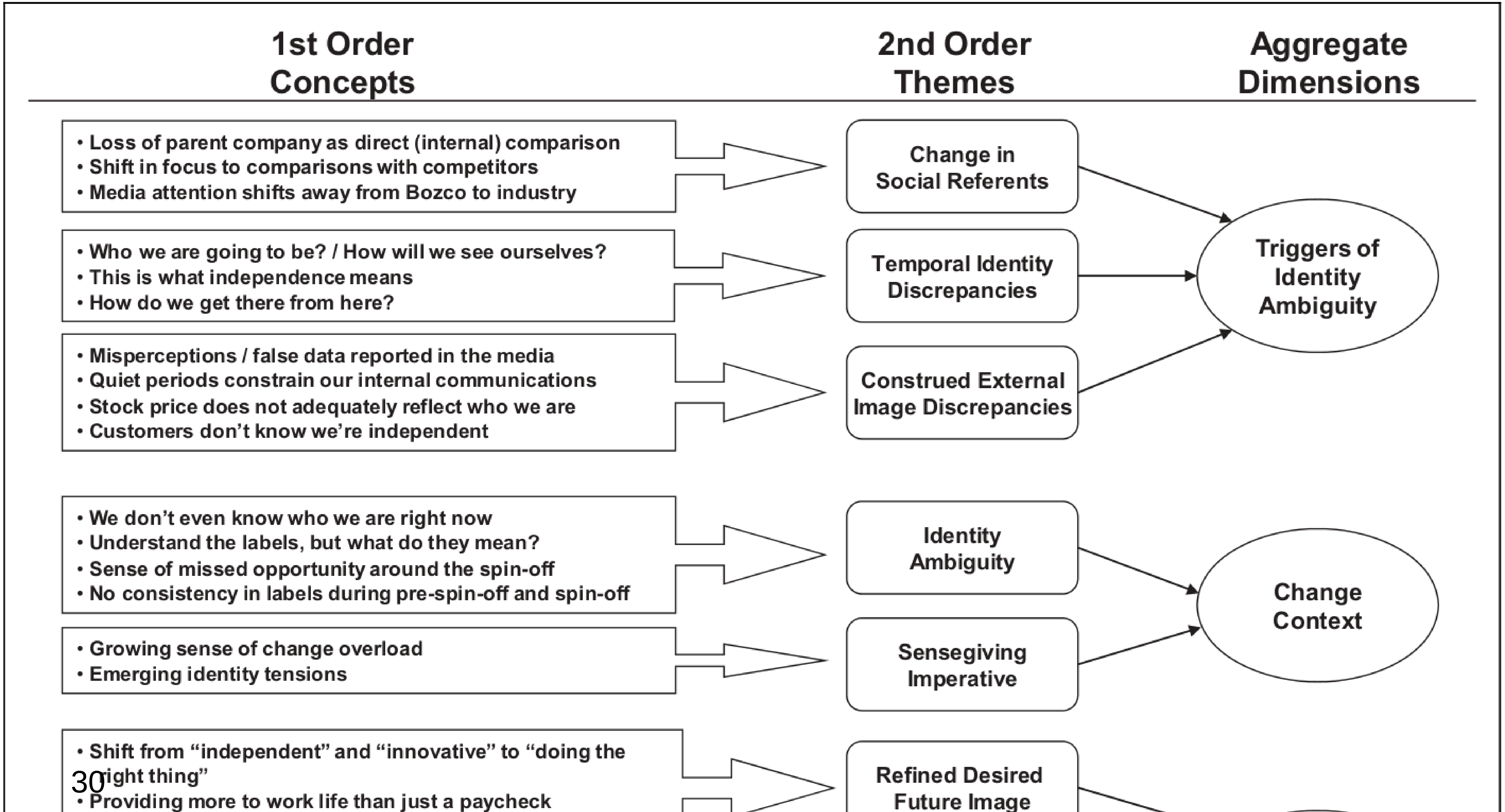
ork/

Data analysis example: Inductive coding

Grounded theory is an inductive method commonly used in literature reviews (Wolfswinkel et al. 2013)

In the data analysis phase, the three coding techniques are central:

- **Open coding** generates higher-abstraction level type categories from sets of concepts/variables
- **Axial coding** develops categories and relates them to their possible sub-categories
- **Selective coding** integrates and refines the categories


1st Order Concepts
2nd Order Themes
Aggregate Dimensions
• Loss of parent company as direct (internal) comparison
• Shift in focus to comparisons with competitors
• Media attention shifts away from Bozco to industry
Change in Social Referents
• Who we are going to be? / How will we see ourselves?
• This is what independence means
• How do we get there from here?
Temporal Identity Discrepancies
• Misperceptions / false data reported in the media
• Quiet periods constrain our internal communications
• Stock price does not adequately reflect who we are
• Customers don't know we're independent
Construed External Image Discrepancies
Triggers of Identity Ambiguity
• We don't even know who we are right now
• Understand the labels, but what do they mean?
• Sense of missed opportunity around the spin-off
• No consistency in labels during pre-spin-off and spin-off
Identity Ambiguity
• Growing sense of change overload
• Emerging identity tensions
Sensegiving Imperative
Change Context
• Shift from "independent" and "innovative" to "doing the right thing"
• Providing more to work life than just a paycheck
Refined Desired Future Image

Data analysis: Example for inductive analysis

Context:

- Scope: Digital platforms for knowledge-intensive services, such as Upwork, Fiverr, or TopCoder
- Sample: 50 papers, mostly published in the Information Systems discipline
- Data: Text fragments and figures have been pre-selected (see worksheet)

Task: Analyze extant research and inductively develop a process model

Data analysis example: Aggregating evidence (I)

Vote counting is one technique to aggregate the evidence from prior empirical studies

- Key variables are extracted and compiled in a list of master codes
- Effects between independent and dependent variables are coded:
    - *+1* for a positive significant effect
    - *0* for no-significant effects
    - *-1* for negative significant effects

Example: Lacity et al. (2011)

MOTIVATIONS TO OUTSOURCE

- Cost Reduction (+ +)
- Focus on Core Capabilities (+ +)
- Access to Skills/Expertise (+ +)
- Business/Process Improvements (+ +)
- Technical Reasons (+ +)
- Political Reasons (+)
- Concern for Security (-)
- Fear of Losing Control (- -)

TRANSACTION ATTRIBUTES

- Uncertainty (-)
- Critical Role of IS – Transaction (-)
- Transaction Costs (- -)
- Business Risk (-)

ITO DECISIONS

INFLUENCE SOURCES

- Mimetic (+ +)

CLIENT FIRM CHARACTERISTICS

- Prior IS Department Performance (-)

LEGEND:

(+ +) more than 80 % of the evidence is positive and significant

ork/

Data analysis example: Aggregating evidence (II)

Strength of vote counting:

- Aggregates evidence from **quantitative and qualitative studies**

Shortcoming of vote counting:

- Risk of bias is not assessed
- Effect sizes are not determined

Meta-analysis techniques address these shortcomings.

Quality appraisal / Risk of bias assessment (I)

- Example: Ringeval et al. (2020): "Fitbit-Based Interventions for Healthy Lifestyle Outcomes: Systematic Review and Meta-Analysis"
- The Cochrane risk-of-bias tool for randomized trials (RoB 2) covers seven **domains of bias**, as illustrated in the table

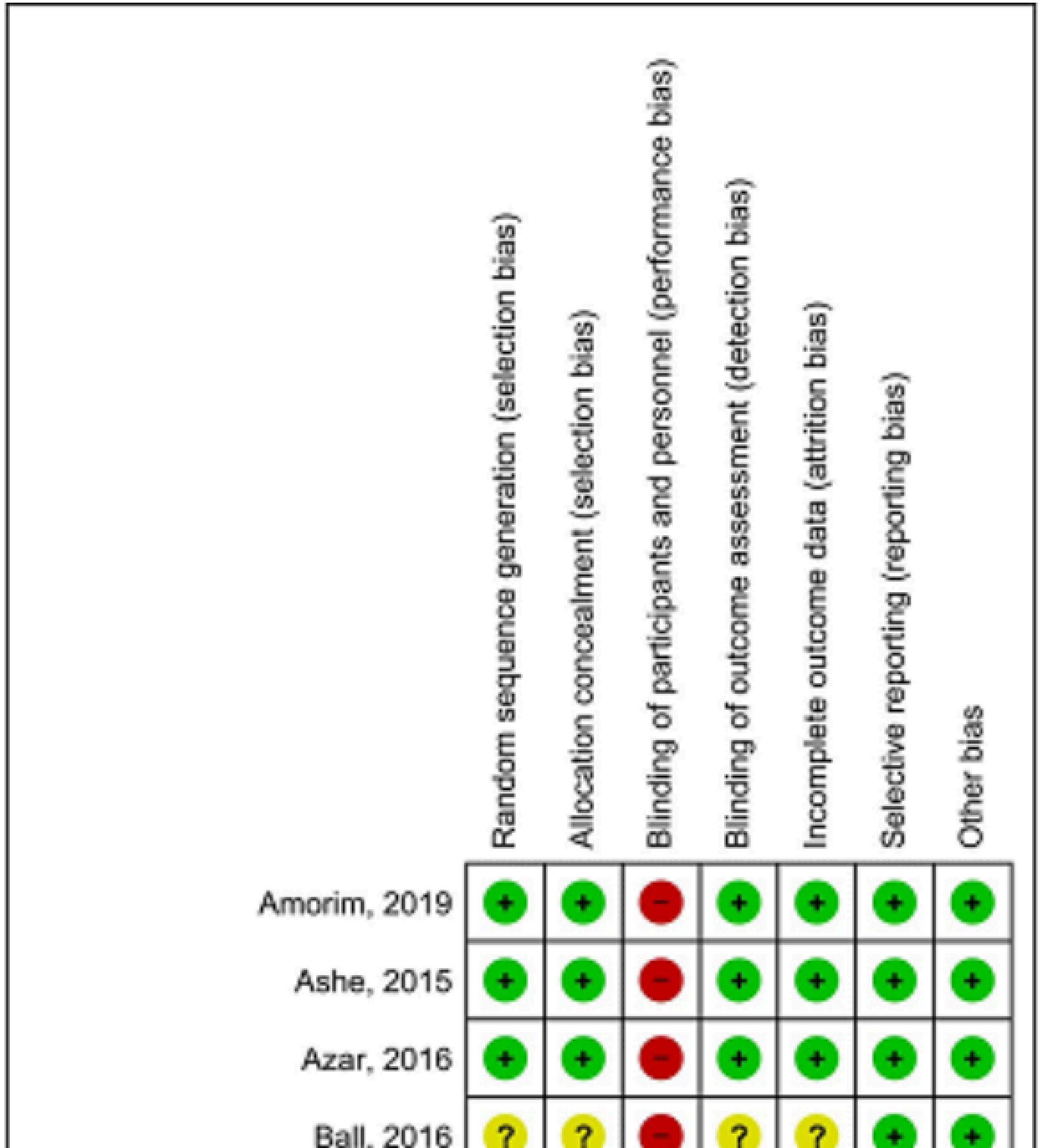
Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"[…] using a computer-generated random number schedule of 10 permuted blocks of 6 and the final block of 8." (p. 3)
Allocation concealment (selection bias)	Low risk	"To ensure allocation concealment, randomization to groups was undertaken by a blinded remote investigator (MS) not involved in recruitment […]" (p.3). It is a central allocation.
Blinding of participants and personnel (performance bias)	High risk	Due to the nature of the intervention and control conditions make blinding impossible.
Blinding of outcome assessment (detection bias)	Low risk	"We conducted a pilot randomized controlled trial with blinded outcome assessment." (p. 2) "Study investigators conducting data collection were blinded to group allocation" (p. 3)
Incomplete outcome data (attrition bias)	Low risk	"Overall, there were 20% of missing data at the 6-month questionnaire follow-up and 16% of missing data across the 6-month weekly surveys." (p. 7). The reasons for missing data are not related to true outcome (p. 7) but they just mentioned they analyzed data by "intention to treat" (p. 6)
Selective reporting (reporting bias)	Low risk	The study protocol is available and all of the study's pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way.
Other bias	Low risk	The study appears to be free of other sources of bias.

Note: For non-experimental studies, other domains of bias may apply (such as the use of fixed-effects for years as a control for omitted time-varying confounders/endogeneity).

Note: It is good practice to analyze whether results differ between high and low quality studies (e.g., through subgroup analyses) instead of excluding low-quality studies.

Quality appraisal / Risk



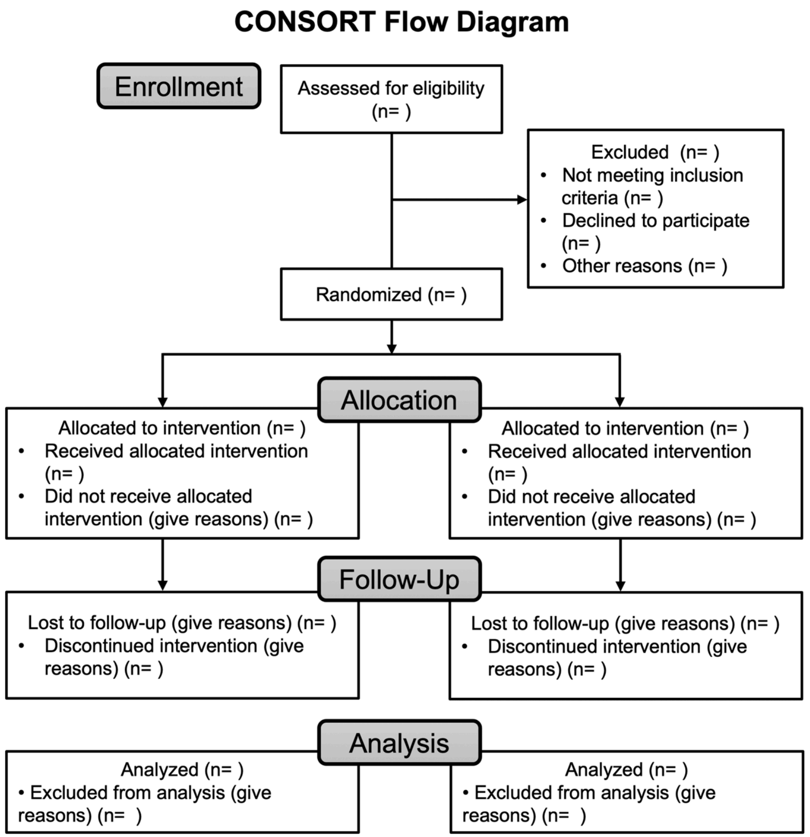
Data analysis: Data extraction

Research objective: "to assess the effects of Fitbit-based interventions, compared with non-wearable control groups, on healthy lifestyle outcomes." (Ringeval et al. 2020)

Type of primary studies: Randomized clinical trials (RCTs), as illustrated in the CONSORT flow diagram

Outcome of interest (at follow-up):

- **Steps per day** (our focus)
- Moderate-to-vigorous physical activity (MVPA)
- Weight loss
- Sedentary behavior (self-reported)



Data analysis

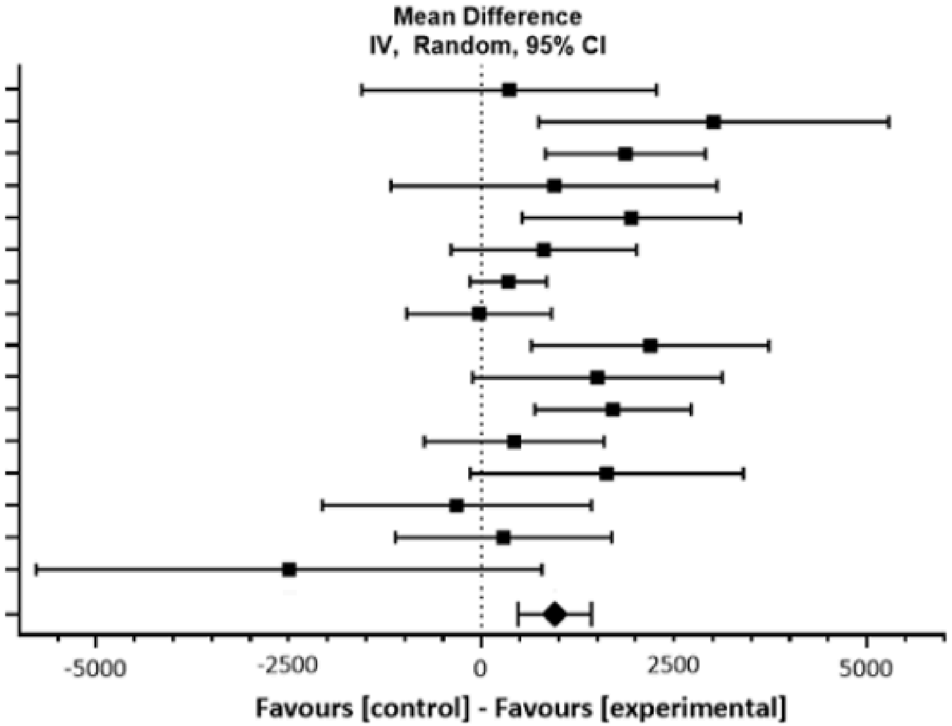
Task: Extract the data from two randomized controlled trials: Thorndike et al. 2014, van Blarigan et al. 2019 based on the following coding sheet:

	Intervention group			Control group		
Study	Steps per day	SD	n	Steps per day	SD	n
Thorndike et al. 2014						
van Blarigan et al. 2019						

Data analysis: Forest plot of standardized mean differences

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
Amorim, 2019	7,379.86	3,627	31	7,020.14	3,554.71	24	4.3%	359.72 [-1551.48, 2270.92]
Ashe, 2015	7,606	3,917	12	4,593	663	7	3.3%	3013.00 [743.02, 5282.98]
Cadmus-Bertram, 2019	1,470	1,881	24	-398	1,751	23	8.5%	1868.00 [829.54, 2906.46]
Cheung, 2019	6,608.29	4,169.86	26	5,665.43	2,338.42	11	3.7%	942.86 [-1173.42, 3059.14]
Christiansen, 2019	6,114	1,989	14	4,169	1,890	15	6.3%	1945.00 [530.67, 3359.33]
Duscha, 2018	411	1,836	16	-398	1,225	9	7.4%	809.00 [-395.09, 2013.09]
Finkelstein, 2016	-130	2,601.31	203	-480	2,516.41	201	12.3%	350.00 [-149.07, 849.07]
Hornikx, 2015	984	1,208	12	1,013	1,275	15	9.1%	-29.00 [-968.93, 910.93]
Katz, 2018	1,441	2,829	31	-747	3,064	26	5.7%	2188.00 [645.66, 3730.34]
Li, 2018	8,217.4	3,095.5	30	6,713.6	3,354.3	31	5.4%	1503.80 [-115.22, 3122.82]
Miragall, 2018	7,958	2,005	22	6,251	1,484	26	8.6%	1707.00 [693.43, 2720.57]
Oliveira, 2019	7,010	3,163	46	6,584	2,612	50	7.6%	426.00 [-740.04, 1592.04]
Paxton, 2018	6,917	3,445	22	5,291	2,298	19	4.8%	1626.00 [-146.00, 3398.00]
Simons, 2018	7,741	4,553.55	55	8,061	5,111.59	63	4.9%	-320.00 [-2063.96, 1423.96]
Thorndike, 2014	7,886	3,622	50	7,600	3,492	49	6.3%	286.00 [-1115.39, 1687.39]
Van Blarigan, 2019	10,047	4,461	20	12,541	5,535	17	1.8%	-2494.00 [-5771.98, 783.98]
Total (95% CI)			**614**			**586**	**100.0%**	**950.54 [475.89, 1425.18]**

Heterogeneity: $Tau^2 = 413106.08$; $Chi^2 = 30.69$, $df = 15$ ($P = 0.010$); $I^2 = 51\%$
Test for overall effect: $Z = 3.93$ ($P < 0.0001$)



Data analysis: Meta-analysis

We extract or calculate **Standardized Mean Differences (SMD)**:

$$d = \frac{\bar{X}_{\text{intervention}} - \bar{X}_{\text{control}}}{SD_{\text{pooled}}}$$

Pooled standard deviation:

$$SD_{\text{pooled}} = \sqrt{\frac{(n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2}{n_1 + n_2 - 2}}$$

> SMD is also known as *Cohen's d*. For small sample sizes, the corrections of *Hedge's g* should be used.
> Note: For research models, we will typically rely on correlations as effect sizes (beta coefficients depend on the other variables of the model).

Random Effects Meta-Analysis

We assume the true effect size varies between studies:

Weighted average of effects:

$$\hat{\mu} = \frac{\sum_{i=1}^{k} w_i \cdot d_i}{\sum_{i=1}^{k} w_i}$$

Weights (account for variance):

$$w_i = \frac{1}{SE_{d_i}^2 + \tau^2}$$

- τ^2: between-study variance
- SE_{g_i}: standard error of each SMD

> Interpretation: Larger w_i = more influence on pooled estimate. Output: Overall effect size with 95% CI shown in forest plot.
> The Doing Meta-Analysis in R book by Harrer et al. offers a good overview of meta-analysis methods.

Discussion of the data analysis section

Task: Create a quick draft for the data extraction and analysis section.

- Would you follow an inductive or deductive approach (why)?
- What outcomes would you expect ideally?

Take-home exercise

For the next session, please sign up for a GitHub account if you don't have one already.

Task: Select an exemplary review and fill out the PRISMA checklist to assess the transparency of reporting.

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	

Summary

- Literature reviews vary in **steps and structure** — tailored to review type and disciplinary context
- We covered the following six steps (in line with Templier & Paré, 2018):

  i. **Problem formulation** – define contribution via gap-spotting or problematization
  ii. **Literature search** – combine database and citation strategies; justify stopping with saturation
  iii. **Screening** – apply transparent inclusion/exclusion criteria; ensure reliability
  iv. **Quality assessment** – assess risk of bias; consider study quality in the interpretation
  v. **Data extraction** – adopt inductive or deductive approaches in line with the review type
  vi. **Data analysis** – choose between techniques like thematic analysis, vote counting, or meta-analysis

Transparency, methodological rigor, and fit with the topic's epistemology are key to credibility.

References

Generic steps

Okoli, C. (2015). A guide to conducting a standalone systematic literature review. *Communications of the Association for Information Systems*, 37. doi:10.17705/1CAIS.03743

Boell, S. K., & Cecez-Kecmanovic, D. (2014). A hermeneutic approach for conducting literature reviews and literature searches. *Communications of the Association for information Systems*, 34, 12. doi:10.17705/1CAIS.03412

Templier, M., & Pare, G. (2018). Transparency in literature reviews: an assessment of reporting practices across review types and genres in top IS journals. *European Journal of Information Systems*, 27(5), 503-550. doi:10.1080/0960085X.2017.1398880

Problem formulation

Alvesson, M., & Sandberg, J. (2011). Generating research questions through problematization. *Academy of Management Review*, 36(2), 247-271. doi:10.5465/amr.2009.0188

Search

Gusenbauer, M., & Haddaway, N. R. (2021). What every researcher should know about searching–clarified concepts, search advice, and an agenda to improve finding in academia. *Research Synthesis Methods*, 12(2), 136-147. doi:10.1002/jrsm.1457

Hiebl, M. R. (2023). Sample selection in systematic literature reviews of management research. *Organizational Research MNethods*, 26(2), 229-261. doi:10.1177/109442812098685

Knackstedt, R., & Winkelmann, A. (2006). Online-Literaturdatenbanken im Bereich der Wirtschaftsinformatik: Bereitstellung wissenschaftlicher Literatur und Analyse von Interaktionen der Wissensteilung. *Wirtschaftsinformatik*, 1(48), 47-59. doi:10.1007/s11576-006-0006-1

Wagner, G., Prester, J., & Paré, G. (2021). Exploring the boundaries and processes of digital platforms for knowledge work: A review of information systems research. *The Journal of Strategic Information Systems*, 30(4), 101694. doi:10.1016/j.jsis.2021.101694

Screen

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., ... & Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of Internal Medicine*, 169(7), 467-473. doi:10.7326/M18-0850

Data analysis

Wolfswinkel, J. F., Furtmueller, E., & Wilderom, C. P. (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, 22(1), 45-55. doi:10.1057/ejis.2011.51

Higgins J, Savovic J, Page MJ, Elbers RG, Sterne JA. Chapter 8: Assessing risk of bias in a randomized trial. In: *Cochrane Handbook for Systematic Reviews of Interventions*. London: Cochrane; 2019. link

Lacity, M. C., Solomon, S., Yan, A., & Willcocks, L. P. (2011). Business process outsourcing studies: a critical review and research directions. *Journal of Information Technology*, 26, 221-258. doi:10.1057/jit.2011.25

Ringeval, M., Wagner, G., Denford, J., Paré, G., & Kitsiou, S. (2020). Fitbit-based interventions for healthy lifestyle outcomes: systematic review and meta-analysis. *Journal of Medical Internet Research*, 22(10), e23954. doi:10.2196/23954