

Minggu #2

The General Framework of a Research

Step 1 : BACKGROUND

- Why are we doing this research? What are the problems and facts? Fishbone diagram mapping.
- What are the research questions? What are the output of the research?
- What are the benefits of this research? What are the scope and limitations?

Step 2 : LITERATURE STUDY

- What are the relevant theories? What are the relevant previous research?
→ concise, compare, criticize, contrast, summarize → develop theoretical framework

Step 3 : DEVELOP RESEARCH METHODOLOGY

- What are the steps for doing this research? What are the output for each steps?
- Identify the data collection methods, analysis methods, decision making methods.
- Qualitative vs quantitative, empirical/positivistic vs interpretive, etc.

Step 4 : DATA COLLECTION

- Primary data (quantitative) : survey, observations.
- Primary data (qualitative) : depth interview, focus group discussions.
- Secondary data : from any formal, credible, and relevant documents.

Step 5 : DATA ANALYSIS, DEVELOP CONCLUSIONS AND RECOMMENDATIONS

- Quantitative analysis : statistics, operation research, AHP/ANP, etc.
- Qualitative analysis : text analysis (hermeneutics).
- Develop specific conclusions and recommendations (based on research, not normative).

Tahapan Awal Penelitian yang Penting

- Mengidentifikasi atribut dari suatu penelitian yang dikatakan “baik”
- Mendapatkan ide untuk topik yang sesuai.
- Menjadikan semua ide menjadi pertanyaan penelitian, tujuan dan scope penelitian yang jelas.

Atribut Penelitian yang Baik

Capability: is it feasible?

- Are you fascinated by the topic?
- Do you have the necessary research skills?
- Can you complete the project in the time available?
- Will the research still be current when you finish?
- Do you have sufficient financial and other resources?
- Will you be able to gain access to data?

Appropriateness: is it worthwhile?

- Will the examining institute's standards be met?
- Does the topic contain issues with clear links to theory?
- Are the research questions and objectives clearly stated?
- Will the proposed research provide fresh insights into the topic?
- Are the findings likely to be symmetrical?
- Does the research topic match your career goals?

And -(if relevant)

- Does the topic relate clearly to an idea you were given -
- possibly by your organisation ?

Problem Formulation

Problem Formulation

**Identification Research
Problem**



**Identification Scope of
Research**



**Identification Research
Questions**



**Identification Research
Objectives**

1. Statistical Hypothesis
2. Hypothetical Statement

Problem Identification

- Determine a well define general problem
 - Sources: research articles, scientific discussion, experience, etc.
 - Point of interest: phenomena, theories, previous research results, etc.
- Exercise your thought to express the problems
- Write them as quickly as possible
- Be ready, that some (may be all) of your general problems are not researchable
 - Conduct literature search
 - Consult experts

Apa yang Disebut dengan “Permasalahan”?

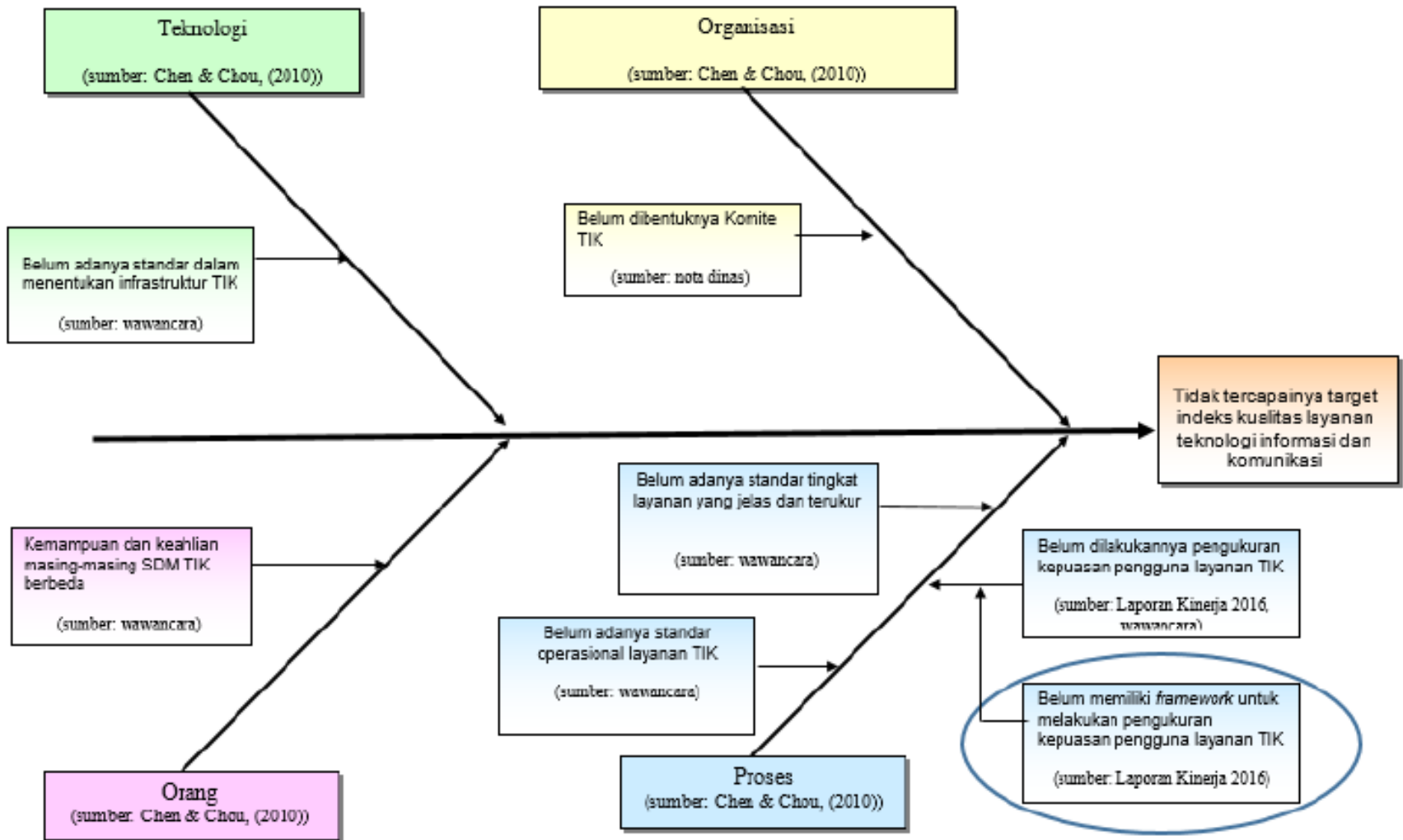
- Adanya perbedaan antara Realitas dan Ekspektasi (Applied Research)
- Contoh:
 - Sebuah organisasi biasanya memiliki ekspektasi yang formal seperti kebijakan (policy), SOP, struktur, etika, peraturan, dan lain-lain.
 - Bandingkan dengan kenyataan
 - Lakukan analisis akar masalah
 - Tentukan RQ dan tujuan penelitian

Identifikasi Kesenjangan (GAP)

Tabel 1. 4 Analisis Kesenjangan

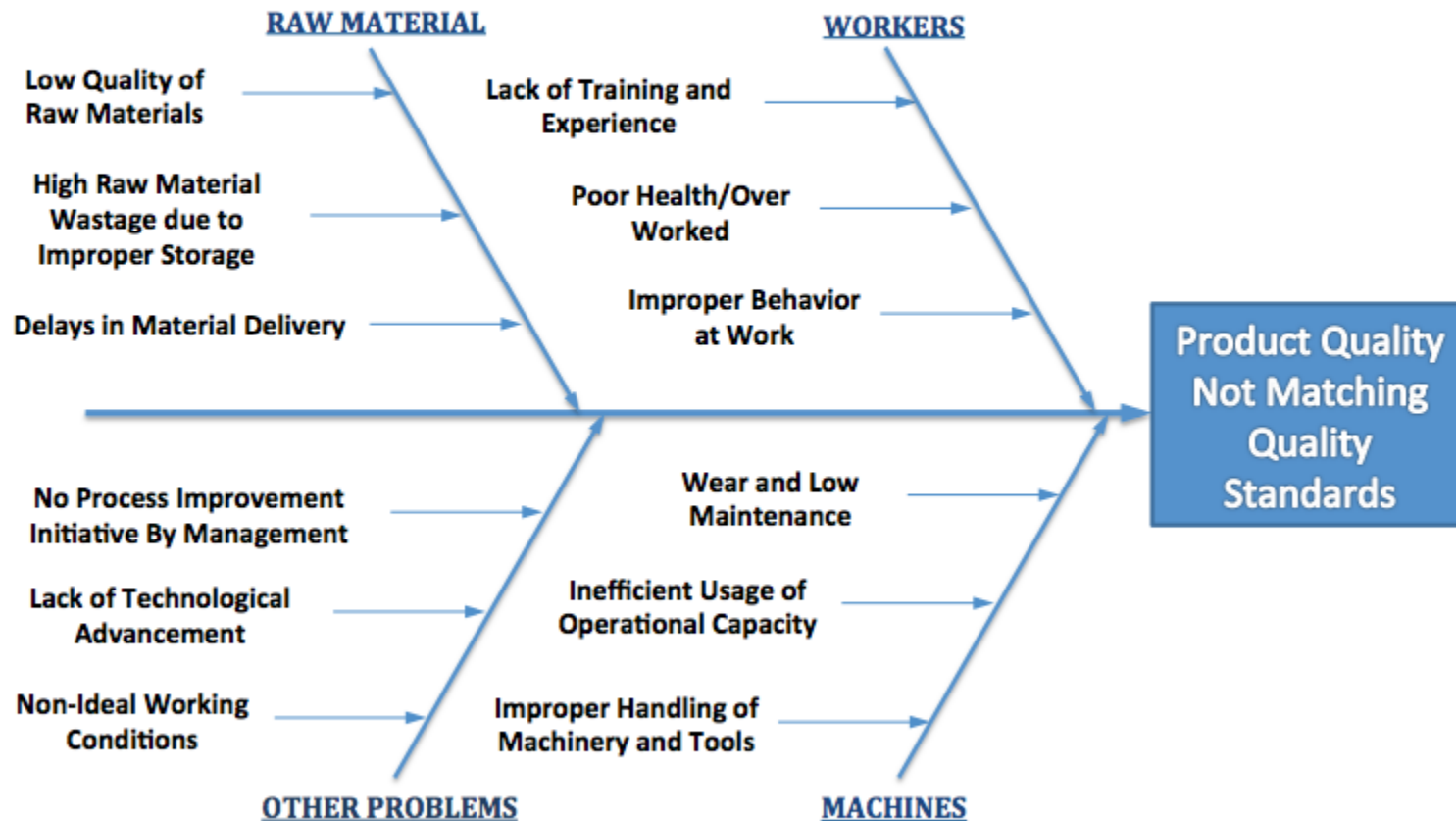
Ekspektasi	Kenyataan	Masalah	Dampak
Target indeks kualitas layanan teknologi informasi dan komunikasi	Capaian indeks kualitas layanan teknologi informasi dan komunikasi tahun 2016 sebesar 3.66	Target indeks kualitas layanan teknologi informasi dan komunikasi tidak tercapai	<ul style="list-style-type: none">- Mempengaruhi capaian kinerja Pustekinfokom KP dan kinerja Kemlu- Mengurangi penilaian publik kepada Kemlu

Identifikasi Akar Masalah



Problems Analysis Tools (1) : Fishbone

CAUSES OF POOR PRODUCT QUALITY



Fishbone Analysis

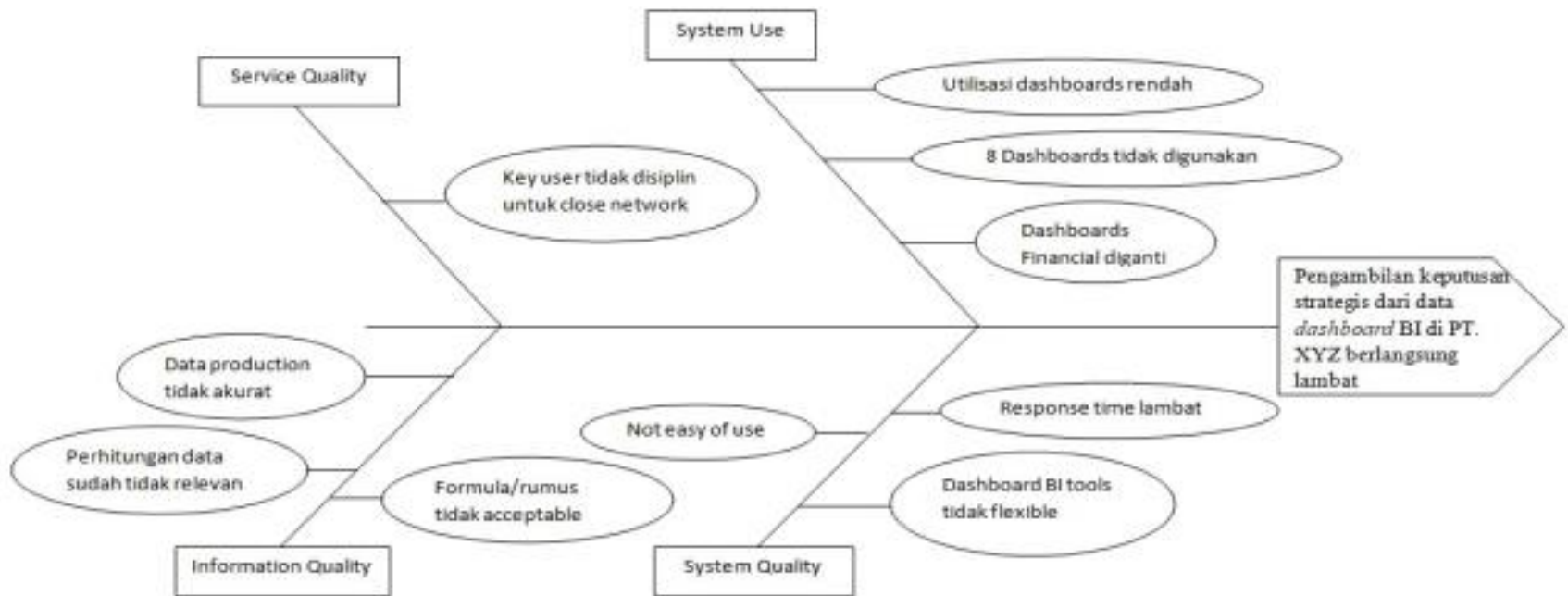
- Permasalahan Organisasi menjadi kepala ikan
- Sirip tulang ikan – Faktor-faktor
- Untuk Bidang Komputer diambil Faktor Teknologi
- Cari akar masalahnya lagi yang akan menjadi topik penelitian anda

PERBAIKAN RANCANGAN DASBOR INTELIJEN BISNIS
MENGUNAKAN METODE HEURISTIC EVALUATION:
STUDI KASUS PT XYZ



Gambar 1.1 Analisis Gap antara Ekspektasi dan Realita Masalah PT. XYZ

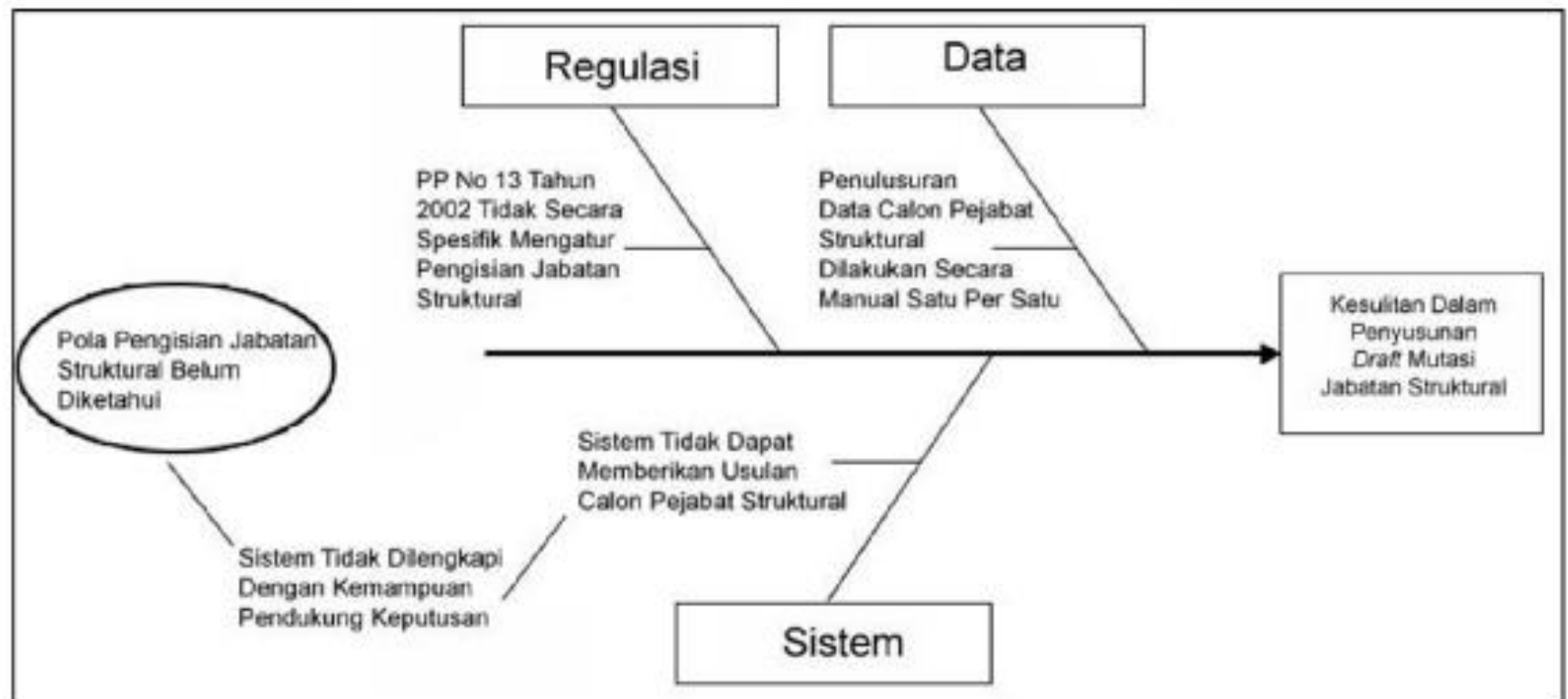
Mencari akar masalah



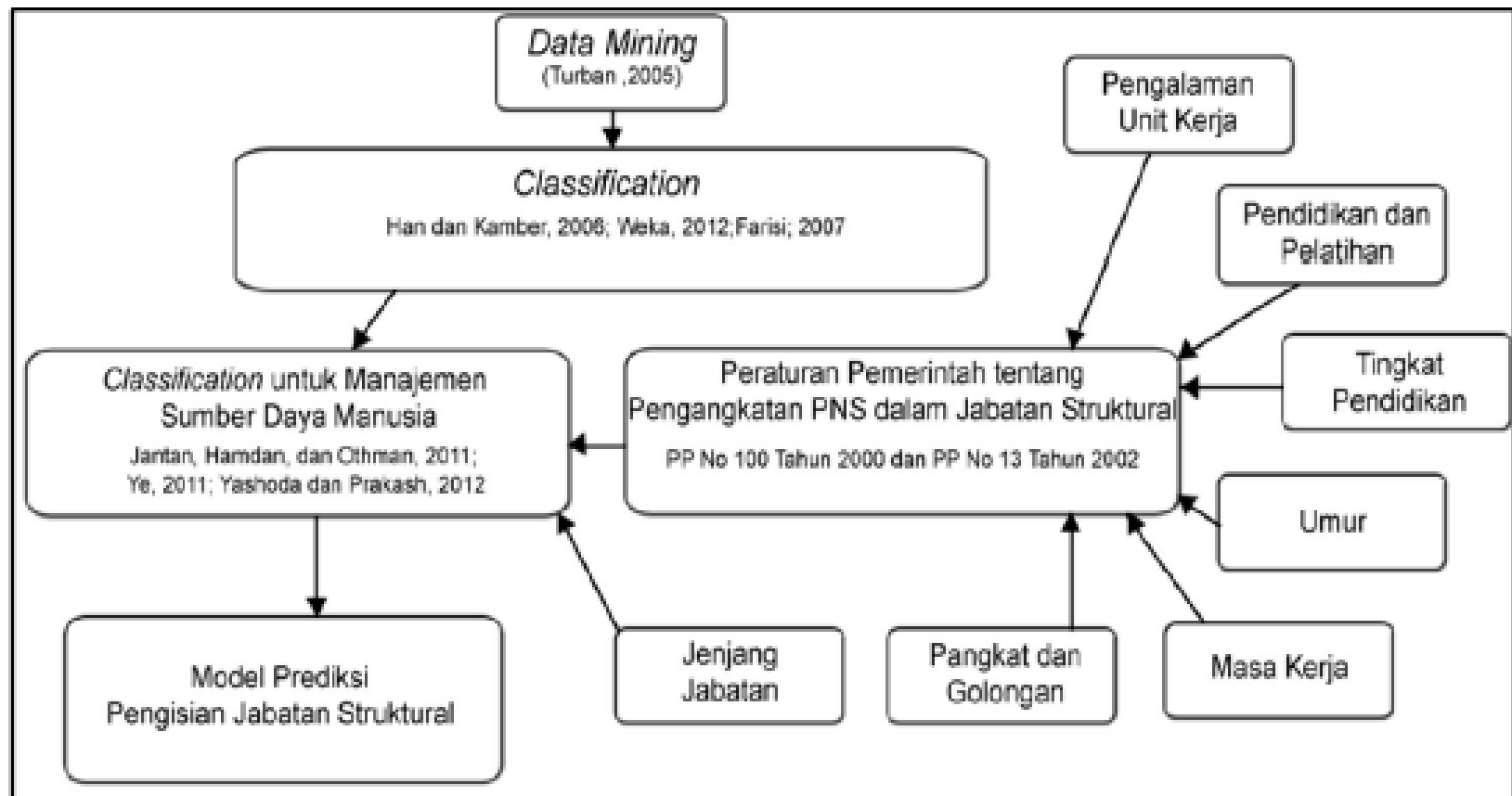
Gambar 1.4 Akar Permasalahan Penelitian

- RQ: Bagaimana rancangan dasbor intelijen bisnis dengan menggunakan metode heuristic evaluation?
- RO: Untuk menyusun perbaikan rancangan dasbor intelijen bisnis dengan menggunakan metode heuristic evaluation.

PEMANFAATAN TEKNIK DATA MINING CLASSIFICATION
UNTUK POLA PENGISIAN JABATAN STRUKTURAL:
STUDI KASUS PEMERINTAH KOTA BOGOR



- “Apakah pola yang sesuai untuk melakukan pengisian jabatan struktural disetiap tingkat eselon pada Pemerintah Kota Bogor?”



Gambar 2.5 Theoretical Framework Penelitian

Melakukan Perbandingan

- Identifikasi permasalahan yang ada di lapangan
 - Gap antara realitas dan harapan
- Menentukan akar masalah
- Memilih permasalahan penelitian
- Mengajukan solusi pemecahan penelitian
- Memilih metode-metode yang terbaik untuk solusi tersebut
- Melakukan perbandingan metode-metode yang terbaik.
- Mengembangkan system dengan menggunakan metode yang terpilih.

Theoritical Research

- Menentukan obyek penelitian.
- Dari kajian literature, identifikasi metode-metode yang ada.
- Analisis kekurangan dan kelebihan metode.
- Pilih salah satu metode yang dipilih dijadikan topik penelitian
- Ajukan solusi perbaikan metode
- Rumuskan RQ (pertanyaan penelitian) dan RO (tujuan penelitian)

Identify the research gap

Tahapan Memulai Penelitian untuk Masalah

Research Problems (RP)		Literature Supports
RP1	Poor accuracy of classifiers due to the noisy attribute predictors of data sets	There are noisy data points in the software defect data sets that can not be confidently assumed to be erroneous using such simple method (Gray et al. 2011)
		Noisy and irrelevant features on software defect prediction results in inefficient outcome of the model (Gayatri et al. 2010)
		The performances of software defect prediction improved when irrelevant and redundant attributes are removed (Wang et al. 2011)
		The software defect prediction performance decreases significantly because the dataset contains noisy attributes (Kim et al. 2011)
RP2	Low reliability of classifiers due to the class imbalance of data sets	Software defect datasets have an imbalanced nature with very few defective modules compared to defect-free ones (Tosun et al. 2010)
		Imbalance can lead to a model that is not practical in software defect prediction , because most instances will be predicted as non-defect prone (Khoshgoftaar et al. 2010)
		Class imbalance will reduce or boost classifier performance (Gray et al. 2011)
		Learning from imbalanced datasets is difficult . The insufficient information that is associated with the minority class impedes making a clear understanding of the inherent structure of the dataset (Liu et al. 2011)
		The balance of on which models are trained and tested is acknowledged by a few studies as fundamental to the reliability of models (Hall et al. 2011)

Defining RQ and RO

Tahapan Memulai Penelitian untuk Maha

Research Problems (RP)		Research Questions (RQ)		Research Objectives (RO)	
RP1	Poor accuracy of classifiers due to the noisy attribute predictors of data sets	RQ1	Which attribute weighting method perform best when used in improving the accuracy of software defect prediction?	RO1	To identify the best attribute weighting method for improving the accuracy of software defect prediction
		RQ2	How does proposed attribute weighting method affect the accuracy of software defect prediction?	RO2	To develop a proposed attribute weighting method for improving the accuracy of software defect prediction
RP2	Low reliability of classifiers due to the class imbalance of data sets	RQ3	Which boosting method perform best when used in improving the reliability of software defect prediction?	RO3	To identify the best boosting method for improving the reliability of software defect prediction
		RQ4	How does proposed boosting method affect the reliability of software defect prediction?	RO4	To develop a proposed boosting method for improving the reliability of software defect prediction

Contoh:

Evaluating Weightless Neural Networks for Bias Identification on News

Rafael Dutra Cavalcanti*, Priscila M.V. Lima*, Massimo De Gregorio[†] and Daniel Sadoc Menasche*

*Universidade Federal do Rio de Janeiro - PPGI

Rio de Janeiro - Brazil

[†]Istituto di Scienze Applicate e Sistemi Intelligenti - CNR

Pozzuoli (NA) - Italy

Mengevaluasi WNN dengan cara membandingkan
dengan metode yang lain

RQ

- Is it feasible to automatically classify the sources of articles on politics?
- What are the advantages and disadvantages of the classification tools, with respect to accuracy and efficiency/ performance?

TABLE I
ACCURACY(ACC) AND STANDARD DEVIATION(SD)

Model	Unigrams		Bigrams		Trigrams	
	Acc	SD	Acc	SD	Acc	SD
Logistic Regression	88%	4%	95%	2%	94%	3%
WiSARD 4 bits with bleaching	88	5	93%	3%	95%	2%
WiSARD 8 bits with bleaching	88	5	93%	3%	94%	3%
SVM (kernel linear)	87	4	93%	4%	92	4
WiSARD 16 bits with bleaching	89	6	93%	6%	93	4
WiSARD 16 bits	89	4	92%	5%	93	3
WiSARD 8 bits	90	6	91%	3%	94	3
WiSARD 4 bits	91	4	91%	4%	94	3
GB(estimators=150, learning rate=0.01, depth=5)	83	8	86%	4%	85	5
Bernoulli Naive Bayes	80	5	85%	5%	86	5
WiSARD 32 bits with bleaching	83	6	84%	4%	93	3
WiSARD 32 bits	82	6	82%	6%	93	4

Experimental
Research

Contoh:

Research on Product Quality Evaluation Based on Big Data Analysis

Abstract—In order to evaluate product quality from nonnumerical data, we propose the product quality evaluation model based on big data analysis including data collecting, data preprocessing, quality feature extraction, vector quantization and quality classification. Quality feature word extension algorithm, reviews quantization algorithm and machine learning algorithm are applied. We finally obtain the qualified rate(88.94%) and 7 features that most concerned by consumers through the analysis of 184,967 effective product reviews of wooden toys. In the end, we compare the SVM machine learning algorithm with decision tree and naive bayes, and discuss the credibility of the results. Our research on product quality evaluation extends the application of big data analysis, and also presents a new method to evaluate product quality in the field of manufacture.

Keywords-quality evaluation; online reviews; big data analysis; machine learning

In this paper, we propose the method to evaluate product quality from the perspective of consumers with data mining technology. And online reviews about wooden toys are data sources. The evaluation method based on nonnumerical data not only can make up the shortfalls of traditional methods, at the same time can also find problems concerned by consumers and supplement standard of product quality

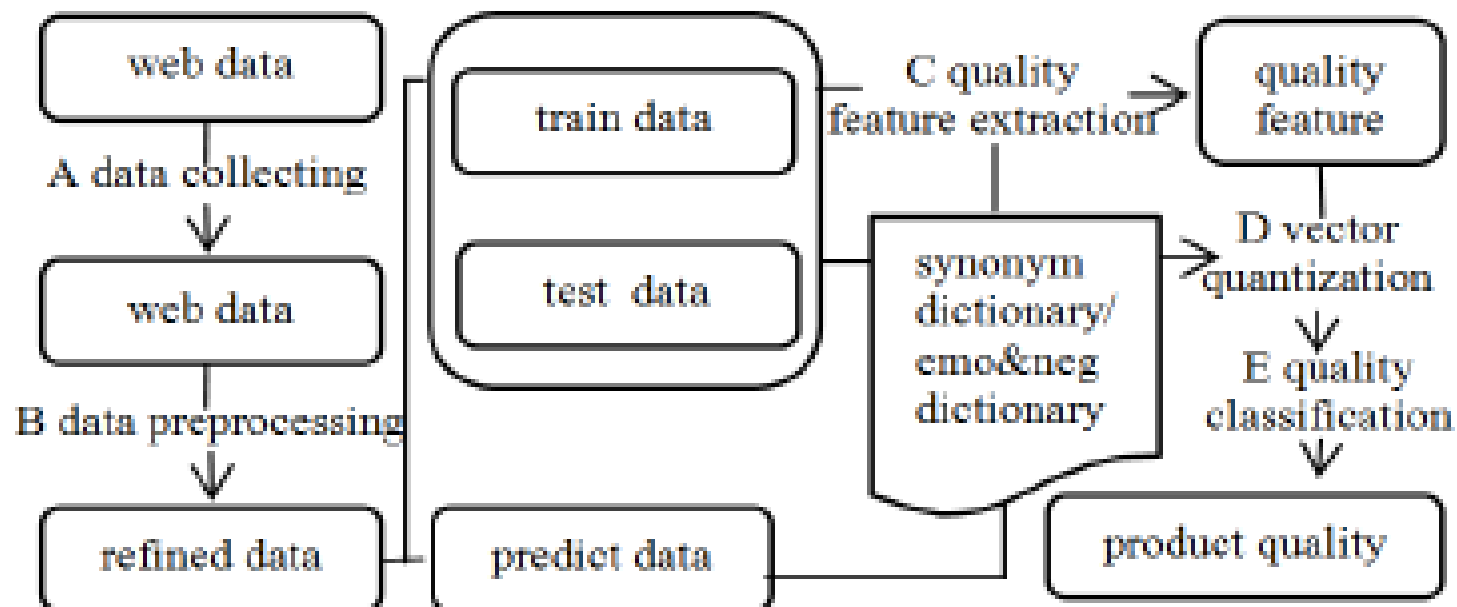


Figure 1. Overview of product quality evaluation based on big data analysis.

Writing Research Questions

- **Write research questions that are**
 - Relevant to the problems
 - Consistent with expected standards
 - Able to produce clear conclusions
 - At the right level (not too difficult)
 - Not too descriptive

Research Objectives

- **OUTPUT**
 - Direct results of the research.
 - Example : a description, a correlational analysis, a design, an evaluation, etc.
- **OUTCOME :**
 - The impact to organization or community.
 - Example : to improve the quality, to give certain directions, to rise the efficiency, etc.