

Research article: Writing an introduction

Disclaimer

This slide set from the video is intended to support your learning after watching the video. Note that the slides don't contain all the explanations in the accompanying video.

Research article

—

Writing an introduction

Diane Pilkinton-Pihko, Ph.D.

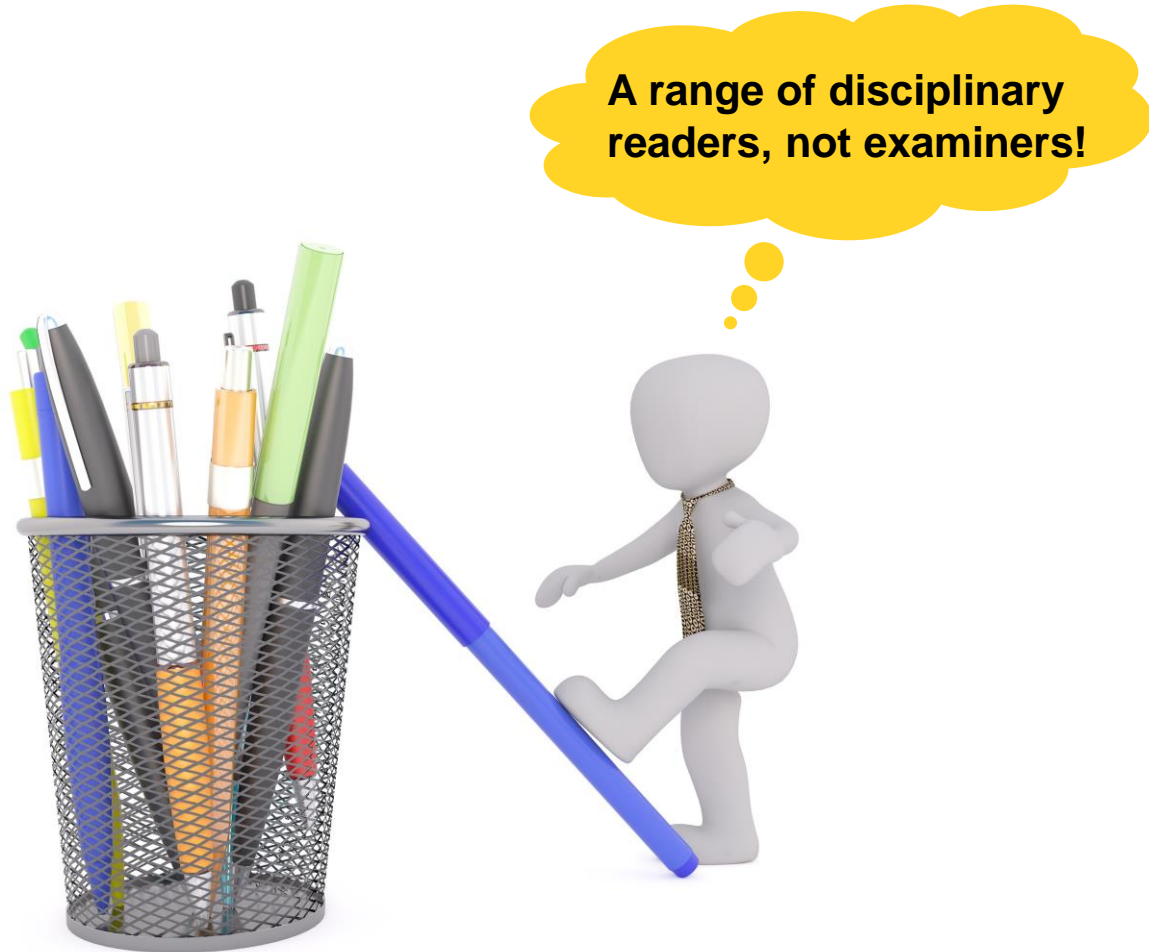


Aalto University
Language Centre



- **Capture interest**
- **Provide context**





Introduction: Five parts

Introduction: Five parts

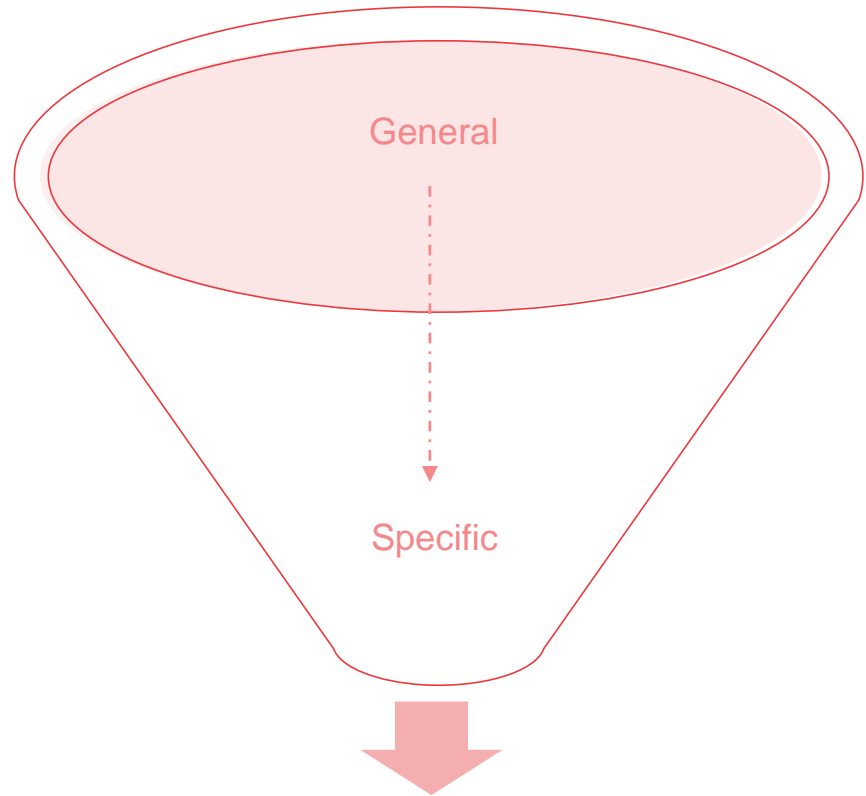
1. Background

2. Unknown / problem

3. Question / purpose

4. Experimental approach

5. Closing off



1. Background

1. Background

- **Keep it short**
- **Check journal guidelines**
- **Provide adequate context**
- **No overview sentences**



1. Background

- **General to specific**
- **Order of the ideas**
problem-solution, chronological,
most-to-least important, if-then, etc.
- **Importance**
- **Key work**
not a full literature review



Materials science is entering the data age. This transition is spearheaded by projects such as the Materials Genome Initiative,¹ the Novel Materials Discovery Laboratory² and Marvel³ that combine high-throughput screening with data storage, systematic data curation and machine learning. Such projects produce computational materials databases that contain information extracted from atomistic simulations, e.g., system geometries, details of the applied theory, electronic structures, methodology and implementation and their number is increasing rapidly.⁴⁻¹⁴

A common problem in these databases is materials classification. Often database users would like to search for specific material types, specific functions or structural classes, such as crystals, molecules, surfaces or 2D materials. To facilitate such searches, the database entries should be tagged according to a classification system. Unfortunately, classifications are not always provided when the data are uploaded to the database, and when they are, they are often based on custom or unspecified definitions. To cope with large heterogeneous datasets from atomistic calculations, automated and verifiable methods for analyzing and categorizing atomistic structures have become a necessity.

Previous work on automated classification of atomistic structures has focused on very specific areas and often required an explicit structural search pattern. For example, defect identification and detection schemes have been developed for crystals that are based on neighbourhood analysis.¹⁵⁻²⁰ In another example, a more automated workflow was applied to identify lower dimensional stable structures in crystals, such as layered solids.²¹ Conversely, many tools are available for the inverse problem that generates an atomistic representation from a given structure definition. Tool sets such as the atomic structure environment (ase)²² and pymatgen²³ include routines for automating tasks like creating a surface given a lattice, orientation and number of layers, generating crystal structures with desired symmetry properties or generating a system representing surface adsorption, given an adsorbate and an adsorbant.

In this work, we focus on structural classification and present a generic structure classification scheme that encompasses all possible structure types. We then introduce a materials structure genealogy presented as an intuitive and human-readable materials structure 'tree of life'. After this general introduction, we present an automatic and accurate classification scheme for two-dimensional structures, including surfaces and 2D materials, that requires no explicit search patterns. This classification process also returns the underlying unit cell and works even in the presence of defects, dislocations and additional atoms. By being able to identify the unit cell, these structures can be meaningfully characterized and often linked to their bulk counterpart. We also present methods that can be used to accurately identify the outlier atoms that are not part of the underlying structure. The classification does not make assumptions about the used cell or the positioning of the structure within the cell.

The NOMAD Archive² is used as a benchmark for testing the classification accuracy and the applicability of our method in a realistic database environment containing heterogeneous data. The classification tools are implemented as a python library licensed under the open-source Apache 2²⁴ license, and the source code together with installation instructions can be found from <https://github.com/SINGROUP/matid>. This library is directly compatible with the popular atomic structure manipulation library ase.

Materials science is entering the data age. This transition is spearheaded by projects such as the Materials Genome Initiative,¹ the Novel Materials Discovery Laboratory² and Marvel³ that combine high-throughput screening with data storage, systematic data curation and machine learning. Such projects produce computational materials databases that contain information extracted from atomistic simulations, e.g., system geometries, details of the applied theory, electronic structures, methodology and implementation and their number is increasing rapidly.⁴⁻¹⁴ (*Introduces the context*)

Materials science is entering the data age. This transition is spearheaded by projects such as the Materials Genome Initiative,¹ the Novel Materials Discovery Laboratory² and Marvel³ that combine high-throughput screening with data storage, systematic data curation and machine learning. Such projects produce computational materials databases that contain information extracted from atomistic simulations, e.g., system geometries, details of the applied theory, electronic structures, methodology and implementation and their number is increasing rapidly.⁴⁻¹⁴ *(Introduces the context)*

A common problem in these databases is materials classification. Often database users would like to search for specific material types, specific functions or structural classes, such as crystals, molecules, surfaces or 2D materials. [. . .] *(Further explains the problem)*

Materials science is entering the data age. This transition is spearheaded by projects such as the Materials Genome Initiative,¹ the Novel Materials Discovery Laboratory² and Marvel³ that combine high-throughput screening with data storage, systematic data curation and machine learning. Such projects produce computational materials databases that contain information extracted from atomistic simulations, e.g., system geometries, details of the applied theory, electronic structures, methodology and implementation and their number is increasing rapidly.⁴⁻¹⁴ *(Introduces the context)*

A common problem in these databases is materials classification. Often database users would like to search for specific material types, specific functions or structural classes, such as crystals, molecules, surfaces or 2D materials. [...] *(Further explains the problem)*

Previous work on automated classification of atomistic structures has focused on very specific areas and often required an explicit structural search pattern. [...] *(Gives examples)* Conversely, many tools are available for the inverse problem that generates an atomistic representation from a given structure definition. [...] *(Gives examples)*

General



Specific

Materials science is entering the data age. This transition is spearheaded by projects such as the Materials Genome Initiative,¹ the Novel Materials Discovery Laboratory² and Marvel³ that combine high-throughput screening with data storage, systematic data curation and machine learning. Such projects produce computational materials databases that contain information extracted from atomistic simulations, e.g., system geometries, details of the applied theory, electronic structures, methodology and implementation and their number is increasing rapidly.⁴⁻¹⁴ *(Introduces the context)*

A common problem in these databases is materials classification. Often database users would like to search for specific material types, specific functions or structural classes, such as crystals, molecules, surfaces or 2D materials. [...] *(Further explains the problem)*

Previous work on automated classification of atomistic structures has focused on very specific areas and often required an explicit structural search pattern. [...] *(Gives examples)* Conversely, many tools are available for the inverse problem that generates an atomistic representation from a given structure definition. [...] *(Gives examples)*

In this work, we focus on structural classification and present a generic structure classification scheme that encompasses all possible structure types. [...] *(Outlines their solution)*

The NOMAD Archive² is used as a benchmark for testing the classification accuracy and the applicability of our method in a realistic database environment containing heterogeneous data. [...] *(Outlines experimental approach)*

Materials science is entering the data age. [. . .] *(Introduces the context)*

A common problem in these databases is materials classification. Often database users would like to search for specific material types, specific functions or structural classes, such as crystals, molecules, surfaces or 2D materials. [. . .] *(Further explains the problem)*

Previous work on automated classification of atomistic structures has focused on very specific areas and often required an explicit structural search pattern. [. . .] *(Gives examples)* Conversely, many tools are available for the inverse problem that generates an atomistic representation from a given structure definition. [. . .] *(Gives examples)*

Problem

Materials science is entering the data age. [...] *(Introduces the context)*

Problem

A common problem in these databases is materials classification. Often database users would like to search for specific material types, specific functions or structural classes, such as crystals, molecules, surfaces or 2D materials. [...] *(Further explains the problem)*

Previous work on automated classification of atomistic structures has focused on very specific areas and often required an explicit structural search pattern. [...] *(Gives examples)* Conversely, many tools are available for the inverse problem that generates an atomistic representation from a given structure definition. [...] *(Gives examples)*

Solution

In this work, we focus on structural classification and present a generic structure classification scheme that encompasses all possible structure types. We then introduce a materials structure genealogy presented as an intuitive and human-readable materials structure 'tree of life'. After this general introduction, we present an automatic and accurate classification scheme for two-dimensional structures, including surfaces and 2D materials, that requires no explicit search patterns. This classification process also returns the underlying unit cell and works even in the presence of defects, dislocations and additional atoms. By being able to identify the unit cell, these structures can be meaningfully characterized and often linked to their bulk counterpart. We also present methods that can be used to accurately identify the outlier atoms that are not part of the underlying structure. The classification does not make assumptions about the used cell or the positioning of the structure within the cell.

Materials science is entering the data age. This transition is spearheaded by projects such as the Materials Genome Initiative,¹ the Novel Materials Discovery Laboratory² and Marvel³ that combine high-throughput screening with data storage, systematic data curation and machine learning. Such projects produce computational materials databases that contain information extracted from atomistic simulations, e.g., system geometries, details of the applied theory, electronic structures, methodology and implementation and their number is increasing rapidly.⁴⁻¹⁴ *(Introduces the context)*

A common problem in these databases is materials classification. Often database users would like to search for specific material types, specific functions or structural classes, such as crystals, molecules, surfaces or 2D materials. [. . .] *(Further explains the problem)* To cope with large heterogeneous datasets from atomistic calculations, automated and verifiable methods for analyzing and categorizing atomistic structures have become a necessity.

[cut two paragraphs]

The NOMAD Archive² is used as a benchmark [. . .]. The classification tools are implemented as a python library licensed under the open-source Apache 2²⁴ license, and the source code together with installation instructions can be found from <https://github.com/SINGROUP/matid>. This library is directly compatible with the popular atomic structure manipulation library ase. *(Summaries experimental approach)*

Materials science is entering the data age. This transition is spearheaded by projects such as the Materials Genome Initiative,¹ the Novel Materials Discovery Laboratory² and Marvel³ that combine high-throughput screening with data storage, systematic data curation and machine learning. Such projects produce computational materials databases that contain information extracted from atomistic simulations, e.g., system geometries, details of the applied theory, electronic structures, methodology and implementation and their number is increasing rapidly.⁴⁻¹⁴ (*Introduces the context*)

A common problem in these databases is materials classification. Often database users would like to search for specific material types, specific functions or structural classes, such as crystals, molecules, surfaces or 2D materials. [. . .] (*Further explains the problem*) To cope with large heterogeneous datasets from atomistic calculations, automated and verifiable methods for analyzing and categorizing atomistic structures have become a necessity.

[cut two paragraphs]

The NOMAD Archive² is used as a benchmark [. . .]. The classification tools are implemented as a python library licensed under the open-source Apache 2²⁴ license, and the source code together with installation instructions can be found from <https://github.com/SINGROUP/matid>. This library is directly compatible with the popular atomic structure manipulation library ase. (*Summaries experimental approach*)

Design of high-efficiency and high-performance electrical machine requires accurate prediction of resistive winding losses. These losses can be divided into skin- and proximity effects, and circulating currents. The latter can be especially significant and dominate in random-wound machines with stranded windings and a high-supply frequency, such as high-speed machines and multipole permanent magnet machines. Indeed, resistive loss increases of several tens of per cent have been observed [1, 2].

Due to manufacture reasons, the positions of the strands in a random-wound winding cannot be exactly known or controlled. In other words, their positions can be regarded as uncertain. Furthermore, they can vary significantly from slot to slot and machine to machine. Correspondingly, the circulating current losses also often exhibit significant variance, even between nominally identical machines [3]. As such, the losses are also stochastic in nature.

This paper extends the authors' previous work in [3] for uncertainty quantification of the circulating current losses. First, an equivalent circuit approach for modelling stranded windings is briefly recounted. Next, a phenomenological sampling algorithm is proposed for modelling the uncertainty in the winding. This algorithm is then coupled to the circuit model. Finally, the statistical properties of the resistive losses are then estimated with Monte Carlo analysis and compared to measurement data.

Design of high-efficiency and high-performance electrical machine requires accurate prediction of resistive winding losses. These losses can be divided into skin- and proximity effects, and circulating currents. The latter can be especially significant and dominate in random-wound machines with stranded windings and a high-supply frequency, such as high-speed machines and multipole permanent magnet machines. Indeed, resistive loss increases of several tens of per cent have been observed [1, 2].

Design of high-efficiency and high-performance electrical machine requires accurate prediction of resistive winding losses. These losses can be divided into skin- and proximity effects, and circulating currents. The latter can be especially significant and dominate in random-wound machines with stranded windings and a high-supply frequency, such as high-speed machines and multipole permanent magnet machines. Indeed, resistive loss increases of several tens of per cent have been observed [1, 2].

Due to manufacture reasons, the positions of the strands in a random-wound winding cannot be exactly known or controlled. In other words, their positions can be regarded as uncertain. Furthermore, they can vary significantly from slot to slot and machine to machine. Correspondingly, the circulating current losses also often exhibit significant variance, even between nominally identical machines [3]. As such, the losses are also stochastic in nature.

Design of high-efficiency and high-performance electrical machine requires accurate prediction of resistive winding losses. These losses can be divided into skin- and proximity effects, and circulating currents. The latter can be especially significant and dominate in random-wound machines with stranded windings and a high-supply frequency, such as high-speed machines and multipole permanent magnet machines. Indeed, resistive loss increases of several tens of per cent have been observed [1, 2].

Due to manufacture reasons, the positions of the strands in a random-wound winding cannot be exactly known or controlled. In other words, their positions can be regarded as uncertain. Furthermore, they can vary significantly from slot to slot and machine to machine. Correspondingly, the circulating current losses also often exhibit significant variance, even between nominally identical machines [3]. As such, the losses are also stochastic in nature.

This paper extends the authors' previous work in [3] for uncertainty quantification of the circulating current losses. First, an equivalent circuit approach for modelling stranded windings is briefly recounted. Next, a phenomenological sampling algorithm is proposed for modelling the uncertainty in the winding. This algorithm is then coupled to the circuit model. Finally, the statistical properties of the resistive losses are then estimated with Monte Carlo analysis and compared to measurement data.

General

Specific

Situation

Design of high-efficiency and high-performance electrical machine requires accurate prediction of resistive winding losses. These losses can be divided into skin- and proximity effects, and circulating currents. The latter can be especially significant and dominate in random-wound machines with stranded windings and a high-supply frequency, such as high-speed machines and multipole permanent magnet machines. Indeed, resistive loss increases of several tens of per cent have been observed [1, 2].

Problem

Due to manufacture reasons, the positions of the strands in a random-wound winding cannot be exactly known or controlled. In other words, their positions can be regarded as uncertain. Furthermore, they can vary significantly from slot to slot and machine to machine. Correspondingly, the circulating current losses also often exhibit significant variance, even between nominally identical machines [3]. As such, the losses are also stochastic in nature.

Situation	<p>Design of high-efficiency and high-performance electrical machine requires accurate prediction of resistive winding losses. These losses can be divided into skin- and proximity effects, and circulating currents. The latter can be especially significant and dominate in random-wound machines with stranded windings and a high-supply frequency, such as high-speed machines and multipole permanent magnet machines. Indeed, resistive loss increases of several tens of per cent have been observed [1, 2].</p>
Problem	<p>Due to manufacture reasons, the positions of the strands in a random-wound winding cannot be exactly known or controlled. In other words, their positions can be regarded as uncertain. Furthermore, they can vary significantly from slot to slot and machine to machine. Correspondingly, the circulating current losses also often exhibit significant variance, even between nominally identical machines [3]. As such, the losses are also stochastic in nature.</p>
Solution	<p>This paper extends the authors' previous work in [3] for uncertainty quantification of the circulating current losses. First, an equivalent circuit approach for modelling stranded windings is briefly recounted. Next, a phenomenological sampling algorithm is proposed for modelling the uncertainty in the winding. This algorithm is then coupled to the circuit model. Finally, the statistical properties of the resistive losses are then estimated with Monte Carlo analysis and compared to measurement data.</p>

Design of high-efficiency and high-performance electrical machine requires accurate prediction of resistive winding losses. These losses can be divided into skin- and proximity effects, and circulating currents. The latter can be especially significant and dominate in random-wound machines with stranded windings and a high-supply frequency, such as high-speed machines and multipole permanent magnet machines. Indeed, resistive loss **increases** of several tens of per cent have been observed [1, 2].

Due to manufacture reasons, the positions of the strands in a random-wound winding cannot be exactly known or controlled. In other words, their positions can be regarded as **uncertain**. Furthermore, they can vary **significantly** from slot to slot and machine to machine. Correspondingly, also the circulating current losses often exhibit **significant** variance, even between nominally identical machines [3]. As such, the losses are also stochastic in nature.

This paper extends the authors' previous work in [3] for uncertainty quantification of the circulating current losses. First, an equivalent circuit approach for modelling stranded windings is briefly recounted. Next, a phenomenological sampling algorithm is proposed for modelling the uncertainty in the winding. This algorithm is then coupled to the circuit model. Finally, the statistical properties of the resistive losses are then estimated with Monte Carlo analysis and compared to measurement data.

Design of high-efficiency and high-performance electrical machine requires accurate prediction of resistive winding losses. These losses can be divided into skin- and proximity effects, and circulating currents. The latter can be especially significant and dominate in random-wound machines with stranded windings and a high-supply frequency, such as high-speed machines and multipole permanent magnet machines. Indeed, resistive loss increases of several tens of per cent have been observed [1, 2].

Due to manufacture reasons, the positions of the strands in a random-wound winding cannot be exactly known or controlled. In other words, their positions can be regarded as uncertain. Furthermore, they can vary significantly from slot to slot and machine to machine. Correspondingly, also the circulating current losses often exhibit significant variance, even between nominally identical machines [3]. As such, the losses are also stochastic in nature.

This paper extends the authors' previous work in [3] for uncertainty quantification of the circulating current losses. First, an equivalent circuit approach for modelling stranded windings is briefly recounted. Next, a phenomenological sampling algorithm is proposed for modelling the uncertainty in the winding. This algorithm is then coupled to the circuit model. Finally, the statistical properties of the resistive losses are then estimated with Monte Carlo analysis and compared to measurement data.

2. Unknown or problem

2. Unknown or problem

- **State the unknown/problem (gap)**
- **Use negative language to signal it**
- **Keep tone respectful and objective**
- **Use standard academic phraseology**



2. Unknown or problem

Little is known about X

None of these studies explain ...

Previous studies have not yet dealt with ...

Previously published studies on X are not consistent



See

<http://www.phrasebank.manchester.ac.uk/introducing-work/>

2. Unknown or problem

Materials science is entering the data age. This transition is spearheaded by projects such as the Materials Genome Initiative,¹ the Novel Materials Discovery Laboratory² and Marvel³ that combine high-throughput screening with data storage, systematic data curation and machine learning. Such projects produce computational materials databases that contain information extracted from atomistic simulations, e.g., system geometries, details of the applied theory, electronic structures, methodology and implementation and their number is increasing rapidly.⁴⁻¹⁴

A common problem in these databases is materials classification. Often database users would like to search for specific material types, specific functions or structural classes, such as crystals, molecules, surfaces or 2D materials. To facilitate such searches, the database entries should be tagged according to a classification system. To facilitate such searches, the database entries should be tagged according to a classification system. Unfortunately, classifications are not always provided when the data are uploaded to the database, and when they are, they are often based on custom or unspecified definitions. To cope with large heterogeneous datasets from atomistic calculations, automated and verifiable methods for analyzing and categorizing atomistic structures have become a necessity.

Problem signaled but with attitudinal language

2. Unknown or problem

Materials science is entering the data age. This transition is spearheaded by projects such as the Materials Genome Initiative,¹ the Novel Materials Discovery Laboratory² and Marvel³ that combine high-throughput screening with data storage, systematic data curation and machine learning. Such projects produce computational materials databases that contain information extracted from atomistic simulations, e.g., system geometries, details of the applied theory, electronic structures, methodology and implementation and their number is increasing rapidly.⁴⁻¹⁴

A common problem in these databases is materials classification. Often database users would like to search for specific material types, specific functions or structural classes, such as crystals, molecules, surfaces or 2D materials. To facilitate such searches, the database entries should be tagged according to a classification system. ~~To facilitate such searches, the database entries should be tagged according to a classification system. Unfortunately,~~ **However,** classifications are not always provided when the data are uploaded to the database, and when they are, they are often based on custom or unspecified definitions. To cope with large heterogeneous datasets from atomistic calculations, automated and verifiable methods for analyzing and categorizing atomistic structures have become a necessity.

Problem signaled but with attitudinal language

2. Unknown or problem

Design of high-efficiency and high-performance electrical machine requires accurate prediction of resistive winding losses. These losses can be divided into skin- and proximity effects, and circulating currents. The latter can be especially significant and dominate in random-wound machines with stranded windings and a high-supply frequency, such as high-speed machines and multipole permanent magnet machines. Indeed, resistive loss increases of several tens of per cent have been observed [1, 2].

Due to manufacture reasons, the positions of the strands in a random-wound winding **cannot be exactly known or controlled**. In other words, their positions can be regarded as **uncertain**. Furthermore, they can vary significantly from slot to slot and machine to machine. Correspondingly, also the circulating current losses often exhibit significant variance, even between nominally identical machines [3]. As such, the losses are also stochastic in nature.

This paper extends the authors' previous work in [3] for uncertainty quantification of the circulating current losses. First, an equivalent circuit approach for modelling stranded windings is briefly recounted. Next, a phenomenological sampling algorithm is proposed for modelling the uncertainty in the winding. This algorithm is then coupled to the circuit model. Finally, the statistical properties of the resistive losses are then estimated with Monte Carlo analysis and compared to measurement data.

The aluminum reduction cell, hereafter referred to as “the cell,” is a complex multi-variable system, which is characterized by energy balance and mass balance coupling. The electrolyte temperature (ET) can be reduced by decreasing the liquidus temperature based on aluminum fluoride (AlF₃) addition, and thus reducing the loss of molten aluminum [1,2]. A well-shaped hearth can be obtained with a precise AlF₃ feeding amount (AFA) to a certain degree [3]. Some research indicates that a well-shaped cell hearth will result in high current efficiency [4,5]. However, an inaccurate AFA may cause a large fluctuation of the side ledge (SL), which will prevent the ideal energy equilibrium from being achieved. As a result of the inherent complexity of the reduction process, making the decision on the amount of AlF₃ addition (MDAAA) mainly relies on technicians and experts. However, it is difficult for inexperienced technicians to perform this task. Because experienced experts may not always be available, circumstances of excess or insufficient AFA frequently occur. Therefore, it is desirable for an accurate AlF₃ addition to be determined using a scientific strategy.

These problems have attracted the attention of researchers. There are three types of research on MDAAA, all of which mainly focus on controlling the AlF₃ concentration (CAIF₃). The first type of research takes an empirical approach that depends on understanding the dynamic of AlF₃. CAIF₃ is monitored by analyzing electrolyte samples, which is done very sporadically. This method has revealed a very strong correlation between CAIF₃ and temperature [6]. Temperature and electrolyte sample analysis with a time lag (TL) are used in CAIF₃ adjusting strategies in the control feedback loop; building a logic rule base is the core method for these strategies [7–9]. The second type of research considers AFA as a function of deviation from a target CAIF₃ and/or temperature. In practice, CAIF₃ was found to change with the SL thickness, and some linear regression models for MDAAA were proposed [10–12]. In the third type of research, strategies are proposed based on the AlF₃ mass balance and/or energy balance. MDAAA models have been built by analyzing AlF₃ evolution from cells, and CAIF₃ control strategies were developed based on estimation and decoupling techniques with detailed process and plant knowledge [13–17]. The methods in the first type of research always rely on human experience, and it is easy for human subjectivity to influence knowledge model construction. Because of the complexity of making a decision about the amount of AlF₃, it is difficult for methods of the second type to capture all of the complex features of AlF₃ addition. Due to the detection of dead zones in the aluminum reduction cell, it is difficult to implement refined AlF₃ addition using methods of the third type, which are based on AlF₃ mass balance and/or energy balance.

Existing research on MDAAA mainly focuses on data-driven or knowledge-driven methods alone. However, data-driven methods may fail to cover the complex characteristics of the cell, and knowledge-driven methods may be overly subjective. Therefore, it is desirable to develop a model that combines historical data with the experience of experts. To address this challenge, modeling with fuzzy cognitive maps (FCMs) seems practical, as it is characterized by intuition and the simplicity of causal representations [18]. FCMs have been widely used in decision analysis, control, modeling, and prediction [19–22]. *(Due to space limitations, the remainder of this introduction was omitted.)*

2. Unknown or problem

The aluminum reduction cell, hereafter referred to as “the cell,” is a complex multi-variable system, which is characterized by energy balance and mass balance coupling. The electrolyte temperature (ET) can be reduced by decreasing the liquidus temperature based on aluminum fluoride (AlF₃) addition, and thus reducing the loss of molten aluminum [1,2]. A well-shaped hearth can be obtained with a precise AlF₃ feeding amount (AFA) to a certain degree [3]. Some research indicates that a well-shaped cell hearth will result in high current efficiency [4,5]. However, an inaccurate AFA may cause a large fluctuation of the side ledge (SL), which will prevent the ideal energy equilibrium from being achieved. As a result of the inherent complexity of the reduction process, making the decision on the amount of AlF₃ addition (MDAAA) mainly relies on technicians and experts. However, it is difficult for inexperienced technicians to perform this task. Because experienced experts may not always be available, circumstances of excess or insufficient AFA frequently occur. Therefore, it is desirable for an accurate AlF₃ addition to be determined using a scientific strategy.

These problems have attracted the attention of researchers. There are three types of research on MDAAA, all of which mainly focus on controlling the AlF₃ concentration. [. . .] (*Explains the pluses and minuses of the three types of research*)

Existing research on MDAAA mainly focuses on data-driven or knowledge-driven methods alone. **However**, data-driven methods **may fail** to cover the complex characteristics of the cell, and knowledge-driven methods **may be overly subjective**. Therefore, it is desirable to develop a model that combines historical data with the experience of experts. To address this challenge, [. . .]

2. Unknown or problem

Signaling the gap > Any indication of

- Contrast? However / nevertheless / despite
- Negative elements? Little is know about / None of these studies . . .
- Extending previous knowledge? Studies . . . are still needed

Contrast	Verbs		Quantity		Adjectives
however	fail	neglect	few	scare	ineffective
nevertheless	ignore	overlook	less	limited	inconclusive
despite	lack	question	little	restricted	uncertain
although	prevent	challenge	no	difficult	unclear
but	hinder	deter	not	inefficient	unreliable
		limit	none	controversial	unsatisfactory

3. Question/Purpose

3. Question/Purpose

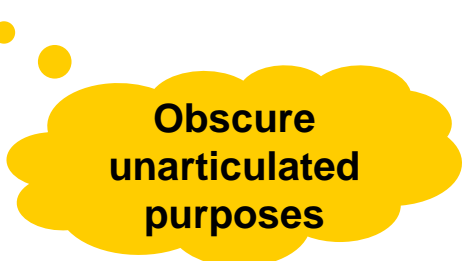
- **State the central question or purpose**
- **No question marks!**
- **Relate every part of the paper to the question or purpose**



Avoid vague verbs

This paper **focuses on** describing the process for identifying service value and transferring this knowledge into concept design.

This paper **studies** the process for identifying service value and transferring this knowledge into concept design.



**Obscure
unarticulated
purposes**

3. Question/Purpose

Three typical purpose statements in engineering and related descriptive verbs



Three types of purpose statements

1. Comprehending a phenomenon

Descriptive verbs

analyze, compare, examine,
investigate, define, determine,
monitor, understand,
experiment with

3. Question/Purpose

Three types of purpose statements

1. Comprehending a phenomenon
2. Designing a solution

Descriptive verbs

build, construct, develop,
model, integrate, propose

Three types of purpose statements

1. Comprehending a phenomenon
2. Designing a solution
3. Applying verified knowledge of a product or service in practice

Descriptive verbs

confirm, compare,
corroborate, experiment with,
evaluate, measure, monitor,
prove, simulate, test, validate,
verify

3. Question/Purpose

Hm, what verb did I use in my purpose statement?



3. Question or purpose

Materials science is entering the data age. [. . .] *(Introduces the context)*

A common problem in these databases is materials classification. Often database users would like to search for specific material types, specific functions or structural classes, such as crystals, molecules, surfaces or 2D materials. [. . .] *(Further explains the problem)*

Previous work on automated classification of atomistic structures has focused on very specific areas and often required an explicit structural search pattern. [. . .] *(Provides examples)* Conversely, many tools are available for the inverse problem that generates an atomistic representation from a given structure definition. [. . .] *(Provides examples)*

In this work, we **focus on** structural classification and **present** a generic structure classification scheme that encompasses all possible structure types. We then **introduce** a materials structure genealogy presented as an intuitive and human-readable materials structure 'tree of life'. After this general introduction, we **present** an automatic and accurate classification scheme for two-dimensional structures, including surfaces and 2D materials, that requires no explicit search patterns. This classification process also returns the underlying unit cell and works even in the presence of defects, dislocations and additional atoms. By being able to identify the unit cell, these structures can be meaningfully characterized and often linked to their bulk counterpart. We also **present** methods that can be used to accurately identify the outlier atoms that are not part of the underlying structure. The classification does not make assumptions about the used cell or the positioning of the structure within the cell.

Use
descriptive
verbs

3. Question or purpose

Materials science is entering the data age. [. . .] *(Introduces the context)*

A common problem in these databases is materials classification. Often database users would like to search for specific material types, specific functions or structural classes, such as crystals, molecules, surfaces or 2D materials. [. . .] *(Further explains the problem)*

Previous work on automated classification of atomistic structures has focused on very specific areas and often required an explicit structural search pattern. [. . .] *(Provides examples)* Conversely, many tools are available for the inverse problem that generates an atomistic representation from a given structure definition. [. . .] *(Provides examples)*

To overcome these existing problems with structural classification, we present

~~In this work, we focus on structural classification and~~ present a generic structure classification scheme that encompasses all possible structure types. We then **introduce** a materials structure genealogy presented as an intuitive and human-readable materials structure 'tree of life'. After this general introduction, we **present** an automatic and accurate classification scheme for two-dimensional structures, including surfaces and 2D materials, that requires no explicit search patterns. This classification process also returns the underlying unit cell and works even in the presence of defects, dislocations and additional atoms. By being able to identify the unit cell, these structures can be meaningfully characterized and often linked to their bulk counterpart. We also **present** methods that can be used to accurately identify the outlier atoms that are not part of the underlying structure. The classification does not make assumptions about the used cell or the positioning of the structure within the cell.

3. Question or purpose

The aluminum reduction cell, hereafter referred to as "the cell," is a complex multi-variable system, which [. . .] (*Outlines the problem*). Therefore, it is desirable for an accurate AlF₃ addition to be determined using a scientific strategy.

These problems have attracted the attention of researchers. There are three types of research on MDAAA, all of which mainly focus on controlling the AlF₃ concentration. [. . .] (*Explains the pluses and minuses of the three types of research*.)

Existing research on MDAAA mainly focuses on data-driven or knowledge-driven methods alone. However, data-driven methods may fail to cover the complex characteristics of the cell, and knowledge-driven methods may be overly subjective. Therefore, it is desirable to develop a model that combines historical data with the experience of experts. To address this challenge, modeling with fuzzy cognitive maps (FCMs) seems practical, as it is characterized by intuition and the simplicity of causal representations [18]. FCMs have been widely used in decision analysis, control, modeling, and prediction [19–22]. [. . .] (*Gives examples*)

In this study, a data and knowledge collaboration strategy for MDAAA **is proposed**, combined with experiential knowledge from experts and data from the aluminum reduction process. The available data is used to extract meaningful fuzzy rules based on fuzzy decision trees and the clustering method, and is also used to detect the edge strength using the state transition algorithm (STA). [. . .] (*Further explains the experimental procedure*) To the best of our knowledge, this is the first time that a collaboration model that simultaneously integrates expert knowledge with production data is used for MDAAA based on augmented FCMs. In this study, the validity of the proposed strategy is verified.

Purpose:
Signaled
with
passive
voice

3. Question or purpose

Design of high-efficiency and high-performance electrical machine requires accurate prediction of resistive winding losses. These losses can be divided into skin- and proximity effects, and circulating currents. The latter can be especially significant and dominate in random-wound machines with stranded windings and a high-supply frequency, such as high-speed machines and multipole permanent magnet machines. Indeed, resistive loss increases of several tens of per cent have been observed [1, 2].

Due to manufacture reasons, the positions of the strands in a random-wound winding cannot be exactly known or controlled. In other words, their positions can be regarded as uncertain. Furthermore, they can vary significantly from slot to slot and machine to machine. Correspondingly, also the circulating current losses often exhibit significant variance, even between nominally identical machines [3]. As such, the losses are also stochastic in nature.

This paper extends the authors' previous work in [3] for uncertainty quantification of the circulating current losses. First, an equivalent circuit approach for modelling stranded windings is briefly recounted. Next, a phenomenological sampling algorithm is proposed for modelling the uncertainty in the winding. This algorithm is then coupled to the circuit model. Finally, the statistical properties of the resistive losses are then estimated with Monte Carlo analysis and compared to measurement data.

3. Question or purpose

Our daily social life is filled with joint actions, during which we adjust our movements according to the ongoing actions of others to fit the demands of various tasks. Behavioral studies have shown [...] *(Summarizes the main point)* However, the neural basis of such between-individuals mutual adaptation is still unclear.

In studies of social cognition, increasing attention is currently being paid to [...] *(States the key point and cites related literature)*

Hyperscanning studies have provided insight into [...] *(States the key point and cites related literature)* However, these studies have not linked [...] *(States the research gap)*

A number of brain-imaging studies have demonstrated that limb kinematics parameters, [...] *(Summarizes the related literature)* Coupling between limb kinematics and brain activity thus seems a useful measure to study the neural underpinnings of one's own movements.

The present study aimed to clarify **how social interaction modulates movement parameters and the brain activity related to hand kinematics.** **For this purpose,** we adopted a joint hand-movement task in which one subject of a dyad either followed or led the movements of their partner. [...] *(Summarizes the experimental procedure)*

4. Experimental approach

4. Experimental approach

- **Include a few sentences about the experimental approach**
- **Optionally: Results and significance**



4. Experimental approach

Materials science is entering the data age. [. . .] (*Introduces the context*)

A common problem in these databases is materials classification. Often database users would like to search for specific material types, specific functions or structural classes, such as crystals, molecules, surfaces or 2D materials. [. . .] (*Further explains the problem*)

Previous work on automated classification of atomistic structures has focused on very specific areas and often required an explicit structural search pattern. [. . .] (*Gives examples*) Conversely, many tools are available for the inverse problem that generates an atomistic representation from a given structure definition. [. . .] (*Gives examples*)

In this work, we focus on structural classification and present a generic structure classification scheme that encompasses all possible structure types. We then **introduce** a materials structure genealogy presented as an intuitive and human-readable materials structure 'tree of life'. After this general introduction, we **present** an automatic and accurate classification scheme for two-dimensional structures, including surfaces and 2D materials, that requires no explicit search patterns. This classification process also **returns** the underlying unit cell **and works** even in the presence of defects, dislocations and additional atoms. By being able to identify the unit cell, these structures can be meaningfully characterized and often linked to their bulk counterpart. We also **present** methods that can be used to accurately identify the outlier atoms that are not part of the underlying structure. The classification **does not make** assumptions about the used cell or the positioning of the structure within the cell.

Summarizes
experimental
approach
using mainly
active voice

4. Experimental approach

The aluminum reduction cell, hereafter referred to as "the cell," is a complex multi-variable system, which [...] *(Outlines the problem)* Therefore, it is desirable for an accurate AlF₃ addition to be determined using a scientific strategy.

These problems have attracted the attention of researchers. There are three types of research on MDAAA, all of which mainly focus on controlling the AlF₃ concentration. [...] *(Explains the pluses and minuses of the three types of research)*

Existing research on MDAAA mainly focuses on data-driven or knowledge-driven methods alone. However, data-driven methods may fail to cover the complex characteristics of the cell, and knowledge-driven methods may be overly subjective. [...] *(Explains the practical solution)*

In this study, a data and knowledge collaboration strategy for MDAAA is proposed, combined with experiential knowledge from experts and data from the aluminum reduction process. The available data is used to extract meaningful fuzzy rules based on fuzzy decision trees and the clustering method, and is also used to detect the edge strength using the state transition algorithm (STA). The initial structure of MDAAA provided by experts is amended using the above fuzzy rules. The problem of having to rely on authoritative experts for FCMs modeling can then be alleviated. The accuracy of MDAAA modeling based on FCMs is sensitive to the edge strength [29], which can be relaxed by detecting strength using the STA. Based on the augmented FCMs, the AFA can be obtained by removing the normalization of the concepts. To the best of our knowledge, this is the first time that a collaboration model that simultaneously integrates expert knowledge with production data is used for MDAAA based on augmented FCMs. In this study, the validity of the proposed strategy is verified.

Summarizes experimental approach using mainly passive voice

4. Experimental approach

The aluminum reduction cell, hereafter referred to as "the cell," is a complex multi-variable system, which [...] *(Outlines the problem)* Therefore, it is desirable for an accurate AlF₃ addition to be determined using a scientific strategy.

These problems have attracted the attention of researchers. There are three types of research on MDAAA, all of which mainly focus on controlling the AlF₃ concentration. [...] *(Explains the pluses and minuses of the three types of research)*

Existing research on MDAAA mainly focuses on data-driven or knowledge-driven methods alone. However, data-driven methods may fail to cover the complex characteristics of the cell, and knowledge-driven methods may be overly subjective. [...] *(Explains the practical solution)*

In this study, a data and knowledge collaboration strategy for MDAAA is proposed, combined with experiential knowledge from experts and data from the aluminum reduction process. The available data is used to extract meaningful fuzzy rules based on fuzzy decision trees and the clustering method, and is also used to detect the edge strength using the state transition algorithm (STA). The initial structure of MDAAA provided by experts is amended using the above fuzzy rules. The problem of having to rely on authoritative experts for FCMs modeling can then be alleviated. The accuracy of MDAAA modeling based on FCMs is sensitive to the edge strength [29], which can be relaxed by detecting strength using the STA. Based on the augmented FCMs, the AFA can be obtained by removing the normalization of the concepts. **To the best of our knowledge, this is the first time** that a collaboration model that simultaneously integrates expert knowledge with production data is used for MDAAA based on augmented FCMs. In this study, the validity of the proposed strategy is verified.

5. Closing off

5. Closing off

Variations

- Outlining the rest of the paper
- Outlining purposes
- Announcing principal findings
- Stating the value of the present research
- Listing research questions and hypotheses



5. Closing off

In this study, a data and knowledge collaboration strategy for MDAAA is proposed, combined with experiential knowledge from experts and data from the aluminum reduction process. The available data is used to extract meaningful fuzzy rules based on fuzzy decision trees and the clustering method, and is also used to detect the edge strength using the state transition algorithm (STA). The initial structure of MDAAA provided by experts is amended using the above fuzzy rules. The problem of having to rely on authoritative experts for FCMs modeling can then be alleviated. The accuracy of MDAAA modeling based on FCMs is sensitive to the edge strength [29], which can be relaxed by detecting strength using the STA. Based on the augmented FCMs, the AFA can be obtained by removing the normalization of the concepts. To the best of our knowledge, this is the first time that a collaboration model that simultaneously integrates expert knowledge with production data is used for MDAAA based on augmented FCMs. In this study, the validity of the proposed strategy is verified.

The outline of this paper is as follows. Section 2 analyzes the role and evolution of AlF₃, and describes the difficulties of and solutions to MDAAA. Section 3 provides the details of fuzzy decision trees and extended fuzzy -means (EFKM), which are used to extract fuzzy rules. The STA is then introduced to detect strength. Section 4 describes the initial structure design and the learning problem. Section 5 models the MDAAA based on augmented FCMs, verifies the effectiveness of the proposed strategy, and provides the discussion. The last section gives the conclusions.

Outlines the rest of the paper, but it lacks sentence variety

5. Closing off

In this study, a data and knowledge collaboration strategy for MDAAA is proposed, combined with experiential knowledge from experts and data from the aluminum reduction process. The available data is used to extract meaningful fuzzy rules based on fuzzy decision trees and the clustering method, and is also used to detect the edge strength using the state transition algorithm (STA). The initial structure of MDAAA provided by experts is amended using the above fuzzy rules. The problem of having to rely on authoritative experts for FCMs modeling can then be alleviated. The accuracy of MDAAA modeling based on FCMs is sensitive to the edge strength [29], which can be relaxed by detecting strength using the STA. Based on the augmented FCMs, the AFA can be obtained by removing the normalization of the concepts. To the best of our knowledge, this is the first time that a collaboration model that simultaneously integrates expert knowledge with production data is used for MDAAA based on augmented FCMs. In this study, the validity of the proposed strategy is verified.

The outline of this paper is as follows. Section 2 analyzes the role and evolution of AlF₃, and describes the difficulties of and solutions to MDAAA. *In Section 3, we provide* Section 3 provides the details of fuzzy decision trees and extended fuzzy -means (EFKM), which are used to extract fuzzy rules. The STA is then introduced to detect strength. Section 4 describes the initial structure design and the learning problem-*, after which Section 5 models FCMs. This section also verifies and discusses* Section 5 models the MDAAA based on augmented FCMs; ~~verifies~~ the effectiveness of the proposed strategy, and ~~provides the discussion.~~ *draws* The last section ~~gives~~ the conclusions.

5. Closing off

Design of high-efficiency and high-performance electrical machine requires accurate prediction of resistive winding losses. These losses can be divided into skin- and proximity effects, and circulating currents. The latter can be especially significant and dominate in random-wound machines with stranded windings and a high-supply frequency, such as high-speed machines and multipole permanent magnet machines. Indeed, resistive loss increases of several tens of per cent have been observed [1, 2].

Due to manufacture reasons, the positions of the strands in a random-wound winding cannot be exactly known or controlled. In other words, their positions can be regarded as uncertain. Furthermore, they can vary significantly from slot to slot and machine to machine. Correspondingly, also the circulating current losses often exhibit significant variance, even between nominally identical machines [3]. As such, the losses are also stochastic in nature.

This paper extends the authors' previous work in [3] for uncertainty quantification of the circulating current losses. **First**, an equivalent circuit approach for modelling stranded windings is briefly recounted. **Next**, a phenomenological sampling algorithm is proposed for modelling the uncertainty in the winding. This algorithm is **then** coupled to the circuit model. **Finally**, the statistical properties of the resistive losses are then estimated with Monte Carlo analysis and compared to measurement data.

5. Closing off

Here, we address the fundamental principles that organize the human feeling states (SI Appendix, Fig. S1). We asked (i) how humans organize their feeling states, (ii) what kind of mental experiences and bodily sensations would best explain the representational structure of the feelings, and (iii) whether the mental experiences and bodily sensations are associated with distinct neural activation patterns. We focused on mapping the basic dimensions (Experiment 1), ontology (Experiment 2), as well as bodily (Experiment 3) and neural (meta-analysis and synthesis of Experiments 1–3) basis of a broad array of feeling states (SI Appendix, Fig. S2 and Table S1). We first quantified the relative intensities of four hypothesized core subjective dimensions (intensity of bodily sensations, saliency of mental experience, emotional valence, and agency) of 100 common subjective feelings spanning from homeostatic (e.g., hunger) and emotional (e.g., pleasure) states to cognitive functions (e.g., recalling). We also measured the relative frequency of experiencing each feeling as the lapse since the last remembered occurrence of each feeling. Next, we measured the experienced similarity of these subjective feelings and mapped the topography of bodily sensations associated with each feeling. Neural activation patterns associated with each state were derived using large-scale meta-analysis of fMRI data. We quantified the spatial representations of these states and linked the representational organization of the subjective states with their bodily and neural activation patterns. **We show that subjective mental states are embodied and emotionally valenced, and that there is a clear correspondence between the mental experiences and their bodily basis that also pertains to the underlying neural activation patterns in the bodily domain.**

Announces
principal
findings

5. Closing off

**Observe what other writers do
in your target journal**



What else?

What else?

Avoid common problems

Excessive length

Introduction is too long

Context/background too narrow

Readers not inspired to read

Overview sentences

Sentences that don't contribute to the meaning

Missing parts

Gap (unknown) – Most commonly missing; implied > not stated

Obscured parts

Parts not properly signaled with language that guides the reader

What else?

Make the five parts clear to the reader



What else?

Get help with signaling at
<http://www.phrasebank.manchester.ac.uk/introducing-work>



Checklist

	Keep it short
	Apply a general to specific pattern
	Include five parts: background, unknown, purpose, approach, closing
	Use language to signal the five parts (see the academic phrasebank)
	Cite the key literature (not a complete literature review)
	Eliminate overview sentences (that don't contribute to the meaning)
	Write for a range of disciplinary readers, not for examiners!
	Check journal guidelines
	Signal the importance of the topic
	Use negative (standard) phraseology to signal the gap (the unknown)



Aalto University
Language Centre

Thank you!



aalto.fi

Special thanks: This material was produced by Diane Pilkinton-Pihko as part of an Aalto University Online Learning (A!OLE) project. It is licensed under Creative Commons Attribution-NonCommercial 4.0 International



Film credits

Text sources

Himanen, L., Rinke, P. & Foster, A.S. Materials structure genealogy and high-throughput topological classification of surfaces and 2D materials. *npj Computational Materials* 4, 52 (2018). Open access under CC-BY-4.0

Lappalainen, P. *Journal of Academic Writing*, Vol. 6 No 1 Autumn 2016, pages 108-121

Lehikoinen, A., Chiodetto, N., Arkkio, A., & Belahcen, A. (2019). Improved sampling algorithm for stochastic modelling of random-wound electrical machines. *The Journal of Engineering*, (17), 3976-3980. <https://doi.org/10.1049/joe.2018.8093> CC-BY-4.0

Nummenmaa L, Hari R, Hietanen JK, & Glerean E. Maps of subjective feelings. *Proc Natl Acad Sci U S A*. 2018; 115(37): 9198-9203. doi:10.1073/pnas.1807390115 CC-BY-4.0

Weichao Yue, Weihua Gui, Xiaofang Chen, Zhaohui Zeng, Yongfang Xie. A Data and Knowledge Collaboration Strategy for Decision-Making on the Amount of Aluminum Fluoride Addition Based on Augmented Fuzzy Cognitive Maps[J]. *Engineering*, 2019, 5(6):1060-1076. Open access under CC-BY-NC-ND-4.0

Zhou G, Bourguignon M, Parkkonen L, Hari R. Neural signatures of hand kinematics in leaders vs. followers: A dual-MEG study. *Neuroimage*. 2016;125:731-738. doi:10.1016/neuroimage.2015.11.002 Open access under CC-BY-NC-ND-4.0

Film credits

Creative commons images under CC-BY-0

3D white male looking for a pen

3D white male amongst spilled pens and canister

3D white male walking up a pen

3D white male with right arm extended

3D white male with a question mark