

Guidelines for Image Description

Document attribute	Information
Document owner	Ulrica Lejbro and Eva Bengtsson
Date	2024-11-29
Reference Number	MTM [2022/510]

Table of Content

Introduction.....	3
1 Background and Purpose.....	5
1.1 Prerequisites for those who design image descriptions	5
1.2 Reference Material.....	5
2 Selection and Considerations.....	6
3 Guidelines for All Image Descriptions	8
3.1 When is an Image Description Needed and What Needs to be Described?	9
3.2 For Whom Should the Image be Described?.....	9
3.3 How Should the Image be Described?	10
3.3.1 Describe the image from a general perspective to a specific perspective	10
3.3.2 Be Concise and Objective	11
3.3.3 Adapt Language and Style in Image Descriptions	11
3.4 Form or Content.....	12
3.5 Quick Guide: Guiding Questions During the Image Description Process	13
4 Guidelines for Specific Types of Images	14
Naturalistic and Artistic Images	14
4.1.1 Photographs	14
4.1.2 Art Images	15
4.1.3 Children's Book Illustrations	16
4.1.4 Other Image Production Techniques.....	18
4.2 Schematic and Factual Images	19
4.2.1 Diagrams	19
4.2.2 Tables.....	20
4.2.3 Maps.....	21
4.2.4 Flowchart.....	22
4.2.5 Tree Diagrams and Family Trees.....	22
4.2.6 Timeline	23
4.2.7 Mathematical Expressions	24
4.2.8 Musical Notation.....	25
5 AI-Generated Image Descriptions.....	26

Introduction

Today's printed and digital texts are not only linguistic but also complex and multimodal. This means that, in addition to the linguistic text, they also contain images and various graphic elements. Images play a significant role in how we think, learn, and experience the world. In this complex text world, it is especially important to make images accessible so that users with visual impairments and blindness do not miss important information and have the same opportunity to absorb the content as sighted people. The best way we can make images accessible is through image description.

Although people with visual impairments and blindness are the primary group of users, image descriptions can also help people with other needs, such as those with dyslexia, ADHD, attention disorders, and second language learners.

MTM's various books are used by people with reading disabilities such as dyslexia and visual impairments. The basic rule is to design image descriptions so that they are meaningful for non-sighted people but also so that sighted people still perceive an informational value in the descriptions.

There are different types of images and different traditions in how images and their functions are classified. An important distinction is whether images are *significant* or *decorative*. Some researchers distinguish between *artistic* and *instrumental* images. Another classification is based on the level of abstraction of the image, that is, how stripped down the representation is in relation to reality. According to this classification, for example, photos and portraits have the highest degree of similarity to reality, drawings and sketches have a lesser degree of similarity, and diagrams and graphs have the least.

We will in this document distinguish between *naturalistic* images (photos, drawings, children's book illustrations) and *schematic* images (diagrams, tables, maps, timelines, flowcharts).

- *Naturalistic* images resemble reality and are easy to understand.
- *Schematic* images are abstract and stripped down, and one must learn rules and conventions to understand and use them.

Describing the content of an image in words is a qualified task that requires several competences. People who create image descriptions must assess which images should be made accessible, what information from the image should be described, find the right concepts and linguistic formulations to convey the information, and structure it in a pedagogical way. They must also consider how the image description is received: it should contribute to understanding and experience without cognitively burdening the user.

These guidelines contain a number of principles and recommendations that help the image describer with the task of describing the image in the best way. The guidelines are the result of collective experience and competence in the field and thus ensure the quality of the image descriptions produced. The guidelines are needed to produce more uniform and effective image descriptions.

Important sources for formulating the guidelines have been three national and four international guidelines for image description (DIAGRAM, UKAAF, NWEA, ABS), studies on the evaluation of image descriptions (Handu 2009, DIAGRAM 2013), feedback from key persons responsible for image description practice, and relevant research in the field. The work on the guidelines and the collection of examples was carried out in collaboration with team of specialists from MTM. I would like to thank the team for excellent cooperation and interesting discussions about the collected material.

Lund 2024-07-02

Jana Holsanova,
Associate Professor of Cognitive Science
at Lund University

1 Background and Purpose

This document presents guidelines for how image descriptions should be performed. Image descriptions are used in MTM's media, such as talking books, braille books, and e-textbooks. The guidelines apply to both written and oral image descriptions. The guidelines may be revised during the ongoing contract period and adapted to new circumstances and new research.

Following this, the guidelines for image description are presented. Section 3 contains guidelines that apply to all image descriptions. Section 4 contains guidelines that apply to specific types of images. The guidelines have an attachment - a collection of examples where different types of commonly occurring images are shown in their context along with surrounding text, headings, information about the book, and the subject. The images in the example collection are provided with a recommended image description. Due to copyright reasons, the example collection is only distributed to a limited extent.

The purpose of these guidelines is for MTM to be able to produce image descriptions that contribute to reading comprehension and improve the reading experience for users. The purpose is also for MTM to be able to produce uniform image descriptions in an efficient manner.

1.1 Prerequisites for those who design image descriptions

To design good image descriptions, one should have sufficient knowledge of the book's subject. This is a prerequisite for being able to assess the relationship of the images to the text content. For several subjects within higher education literature such as linguistics, medicine, chemistry, and physics, this means that one should have studied at the university level or have comparable knowledge in the relevant subject. It is also important to have knowledge of correct technical terminology.

1.2 Reference Material

It is not always easy to know what is the characteristic of a specific flower, building, insect, etc., especially if the image is unclear. To be sure that one has correctly described, for example, a bird, it is appropriate to use descriptions from technical literature. Most handbooks and encyclopedias contain, in addition to illustrations, a descriptive text of the plant, bird, or animal.

By using descriptions from technical literature, one gets good guidance on what distinguishes birds, plants, buildings, etc. The book itself, exercise booklets, answer keys, etc., can also be used to gather facts.

However, it is important to use the original image as the basis for the description. Do not use images from your reference material and do not add information that is not visible in the image.

2 Selection and Considerations

Before performing an image description, one needs to consider various questions and make several assessments. One needs to decide which selection of images should be made accessible with an image description and which aspects of an image are relevant. The following questions can be helpful to consider:

- Which images should be described, and which should not?
- What is the purpose of the image?
- How does the image interact with the text?
- Does the image provide important information that is not in the text or caption?
- What image information is relevant in the context and for the task?
- How will the image be used in the concrete reading situation?
- What information is important to convey?

Before performing an image description, one also needs to consider what language usage should be followed, which concepts should be used, and how extensive the image description should be. The following questions can be helpful:

- How detailed and thorough should the image description be?
- How interpretive can the image description be?
- Is the description careful, simple, and understandable?
- Is the user expected to understand the concepts used in the image description?

Finally, one needs to reflect on who the recipient of the image description is and what different purposes the recipient might have with their reading. The following questions can contribute to a better understanding of this:

- What are the needs, age, and knowledge level of the users?
- How should the language be adapted to the users and their needs and knowledge level?
- Is the image description presented in manageable portions so that the user can absorb it?

This document contains principles and recommendations that can be helpful in addressing these questions and any challenges that may arise. The document contains instructions on how to make images accessible with image descriptions. The guidelines are divided into the following order: *General Guidelines* and *Guidelines for Specific Types of Images*.

Section 3. *General Guidelines* provides an overview of important aspects to consider when making all types of images accessible. These aspects include what needs to be described, when and how something needs to be described, and for whom. These aspects also include principles that address the context in which the image is presented, the purpose of the image, and who the recipient of the image description might be. Another principle addressed in this section is the language usage and style level that should be aimed for.

Section 4. *Guidelines for Specific Types of Images* explains principles for describing specific types of images. They include both naturalistic and artistic images (photos, drawings, paintings, children's book illustrations) and schematic and factual images (diagrams, graphs, maps, tables, and timelines).

3 Guidelines for All Image Descriptions

The design, content, and scope of the image description should be shaped by three factors: *context*, *purpose*, and *user*.

Context

Images are usually not presented in isolation but as part of a larger context. They are embedded in surrounding text and shown in a certain type of book, which can be a fiction book, subject book, or technical literature. The image and the book can, in turn, occur within a certain subject or scientific field, such as biology, geology, or art history. By studying the image in its context, one can assess how the image fits into the whole and what role it plays. This is important to determine if the image should be described and how.

Purpose

It is important to know the purpose of the image and how it will be used to select relevant aspects for the description. Does the reader only need a summary of the image, or should the facts presented in the image be used to complete a task? The purpose determines how accurately the image should be described and which key elements need to be mentioned. This can involve identifying different details and components of the heart and lungs and their placement in the body. If the image is purely decorative, it does not need to be described in detail, and a short note is sufficient. If there is sufficient detailed information about the image content in the caption or surrounding text, no image description is needed.

User

The image description should facilitate understanding without being a cognitive burden for the user. To create a well-adapted image description, one must know the user and which target group they might belong to. A user's needs can vary depending on the reading technique they use. Different users can also have different needs depending on which target groups they belong to and what they wish to achieve with their reading. To understand which target group the user belongs to, one needs to reflect on the person's age, knowledge level, and prior knowledge in the relevant subject or scientific field. Formulations, word choices, and terminology must be adapted to the target group. If several people are to take part in the same image description, for example, a sighted and a non-sighted person, it may sometimes be relevant to add information about colors and similar details that are not relevant in a standard image description.

3.1 When is an Image Description Needed and What Needs to be Described?

The following principles apply to determine when an image needs to be described and how detailed it should be.

- Study the image carefully in its context to understand how it fits into a larger context. Determine what type of book the image is in.
- Examine how the surrounding text and image are content-wise related to determine if and how the image should be described.
- Use the context to determine the purpose of the image. The purpose guides which parts of the image should be described and how accurately.
- An image always needs to be described if it contains essential information that is not in the text, if the image complements and clarifies the text, if the image contains data that must be transferred or interpreted, or if the image is used as a basis for discussion.
- All types of text in an image should be described. This also applies to captions embedded in the image, as a screen reader program cannot read this.
- An image needs to be described *in detail* if it plays an essential role in teaching a concept or explaining a phenomenon.
- An image needs to be described *in detail* if it is an essential part of a task or activity. The description should then contain the specific information needed to complete the task (without revealing the answer or solving a problem). An image does not need to be described if the information from the image is already expressed in the surrounding text or if there is sufficient detailed information about the image content in the caption.
- An image does *not* need to be described in detail if it is purely decorative and does not contribute any essential or new information to the text. A short note is sufficient.
- The cover image does not need to be described.

3.2 For Whom Should the Image be Described?

Knowing the user is an important prerequisite for assessing which aspects of an image need to be explained.

- Find out who the user is, how old they are, and what subject knowledge they have.
- Analyze what needs the user might have based on their reading technique. Determine which formulations, word choices, and terminology might be appropriate for the user based on which target group the user belongs to.
- Find out which basic concepts have already been explained in the text and avoid repeating explanations in the image description.
- Use the same terminology as in the text and avoid introducing new terminology that deviates from what is in the text.

3.3 How Should the Image be Described?

The image description should be designed from a general perspective to a specific perspective. The description should be concise, objective, and the language should be adapted to the user. According to the Language Act §11, the language within public activities should be careful, simple, and understandable. In addition, the image should be described in the same style and with the same terminology as in the surrounding text. An exception to this is image descriptions aimed at an adult with visual impairment reading a picture book to a sighted child. If the surrounding and other text in the book is in English, the image description should also be in English.

3.3.1 Describe the image from a general perspective to a specific perspective

A general advice is to provide information step by step, with increasing level of detail. In cases where there is an image title, start by stating this. Then briefly describe the composition and layout of the image. Continue by giving a brief overview of the most relevant parts of the image and their function. Finally, describe the details. This helps the user build an internal mental image of the image content.

- Start with background information: state what type of image it is.
- Design the image description from a general perspective to a specific perspective, i.e., from an overview summary to relevant details.
- Continue by giving an overview: state what the image depicts, which image elements are included, what their function is, and possibly where they are located in the image.
- Describe the image with a linguistic and logical structure and sequence. Consider what is new information for the user and what is already known. For example, if a person appears in the book several times, the person only needs to be introduced and described in detail once.
- Choose relevant information! Study the text and caption. Assess if any essential information is missing that is in the image. If yes, describe that part of the image and the relevant details as accurately as possible with established terminology.
- Formulate the image content in a logical sequence, portion it into short information bits.
- Do not jump between the description of the right and left side of the image.

3.3.2 Be Concise and Objective

The amount of information needed in an image description depends on the complexity of the image, context, and how the image will be used. One must also consider how long it takes for the user to read an image description. It is desirable to create as short descriptions as possible to avoid cognitively burdening the user. A single sentence with the most important trend can be sufficient in some contexts.

An image description should be as objective, factual, and impartial as possible. One should not add additional information or values. However, this can be difficult to achieve, as all descriptions are to some extent subjective and depend on how the describer and user understand and interpret concepts in different contexts.

- Be concise. Avoid long, complicated sentences and descriptions of unimportant details.
- Read texts and any captions. Do not repeat information presented in the main text or adjacent texts.
- Avoid introducing new concepts or terms.
- Avoid personal opinions, values, and interpretations. Try to be as objective, impartial, and factual as possible. Describe what is visible.
- Do not convey your subjective opinions and associations. Let the user form their own opinions instead.
- Do not provide extra information or facts. Do not use evaluative expressions or stereotypes.
- Do not over-interpret the image. Only convey patterns and trends that are obvious.
- If a certain pattern can be discerned in, for example, a graph, diagram, or table, this should be described objectively. If there are several examples to compare, choose the example that the remaining ones should be compared with. Proceed logically and systematically.

3.3.3 Adapt Language and Style in Image Descriptions

Regarding language, the image should be described using established terminology. The description should be linguistically correct and consistent in style. It is important to consider how the image description will be read aloud by a screen reader program. Use the correct special characters in the right places to achieve the correct intonation in speech synthesis. Use the same characters as in the rest of the text in the book unless otherwise instructed in the assignment. See examples of special characters that can be confused with similar characters in the appendix *Special Characters*.

- Adapt the language to the user.
- Describe the image briefly and concisely, systematically and consistently.
- Use active verbs in the present tense.

- Use the same style and terminology as the surrounding text. Do not invent your own terms.
- Spell out abbreviations and symbols to ensure correct pronunciation when using a screen reader.
- Do not use formulations that refer to sight, such as "the image shows," "in the image you can see," or "in the image one sees." Often the heading "image description" is added in braille and reading programs, making it clear to the user that it is an image being described.
- Specify the type of image, for example, "map" instead of "image."
- Paragraph division is recommended for longer image descriptions.
- Check spelling, grammar, and punctuation.
- Proofread the image description.

3.4 Form or Content

An important assessment concerns whether the content of the image or the form of the image is important. For diagrams that summarize data, this data must be described, while for example, art images or children's book illustrations, the appearance of the image must also be described.

When making schematic images such as maps, graphs, and diagrams in textbooks and technical books accessible, describe what the image illustrates, that is, the *content*, not how the depiction itself is executed.

Example:

"On a world map, countries are assigned to four categories: Sustainable, Stable, Warning, and Alert. The five most stable countries are ... The five least stable countries are ...:"

Do not write "This is a world map, where countries are given colors as follows. Shades of blue: Sustainable, Green: Stable, Yellow: Warning, and Red: Alert." This description provides no information about which countries belong to which category. Instead, apply the color coding and describe the content as in the example above.

Form and appearance should only be described in the following cases:

- If users need to understand what a diagram is and how to use it, its appearance must be described in addition to the information the diagram conveys.
- If users need to be able to reproduce the image or formulate a subjective opinion about the image, that is, if the visual appearance is important for the task, it must be described carefully.

3.5 Quick Guide: Guiding Questions During the Image Description Process

Deciding which images should be made accessible and how to do it best is a complex process. Finally, for this section, some guiding questions are presented that need to be considered when making these decisions:

- What type of image is it?
- What information does the image convey?
- Is this information included in the surrounding text or caption?
- What is the purpose of the image? Does the image provide important information needed to understand the content/subject, or is it decorative?
- Is the image used to provide a general understanding of a concept, to show details about parts and the whole, or does the image contain data to be compared, reported, or used for a specific task?
- What needs does the user who needs the image description have? Age?
- What prior knowledge can be expected of the user?

Ask Yourself:

- What is the most important, the most relevant in the image for the user?
- How do I describe it briefly and clearly, so that the user understands, remembers, and possibly can reproduce it later?
- Read your image description aloud to someone. Can the recipient imagine the image? Is the description understandable? Test different variants and rewrite your image description if there is time and opportunity.

4 Guidelines for Specific Types of Images

Guidelines for specific types of images include both naturalistic images that resemble reality and artistic images such as photos, drawings, paintings, and children's book illustrations. The guidelines also include schematic and factual images, such as diagrams, graphs, maps, family trees, tables, and timelines.

Regarding the description of naturalistic and artistic images, such as children's book illustrations, what is visible in the image and the feeling the image conveys should be described. Regarding the description of schematic images, such as diagrams, what the image conveys in terms of content should be described, that is, the facts that the images present, not how the depiction itself is executed.

The guidelines for specific types of images are grouped under (a) naturalistic and artistic images and (b) schematic and factual images. The individual image types are briefly characterized and provided with guidelines and comments.

Naturalistic and Artistic Images

4.1.1 Photographs

Photographs are representational depictions. Photographs can be very detailed and complicated to describe, but a complete description is not always necessary. Consider whether the photograph is significant in the context or if it serves a decorative function.

- Study the photograph and the surrounding text to determine whether the photograph is decorative or significant in relation to the text and message.
- Consider how the image will be used, what is required to understand the task, and how many details are needed.
- If the photograph is significant and provides additional information that is crucial to understanding the content, it needs to be described.
- If the caption is embedded in the image, or if any other type of text is in the image, the information needs to be described.
- Avoid describing irrelevant visual information and details that may distract the user.
- If the photograph serves a decorative function, it does not need to be described in detail; a short note is sufficient.
- If the photograph adds value in the context but there are details about the image content available in the caption, it does not need to be described.
- What is visible in the image should be described without subjective interpretation.

- Depending on the context, a description of the location/environment and subject may be needed. Depending on the context, a description of the foreground, background, color, and objects may be needed.
- If the photograph is an artwork and users are expected to give an emotional response, the appearance of the photograph should be described, including the mood or feeling the image conveys.
- If the answer to a task is based on the photograph and many details are required, you can start with the following questions (if they are significant):
 - What does the image show? What objects/items are there? What are they used for? Where are the objects located? How big are they? Are they known?
 - Is the weather relevant?
 - Are there people there? Who are the people in the photo? Where are they located? What are they doing? How do they interact with each other? What are they wearing?
 - Is it a city or countryside? Are there buildings? What architecture or style?

4.1.2 Art Images

Art images are reproductions of visual artworks and include both representational and non-representational works. When describing art images, one must consider the knowledge level and context. For example, an image description in an art history textbook at the university level needs to contain different details about the subject, technique, and style than an image description intended for children. The language and terminology must also be adapted to the rest of the text in the book. If the purpose of the art image is to evoke emotions and create immersion, the description can be more interpretive. When describing art, a vivid language that evokes internal images and pictures in the target audience should be used.

- State the title of the artwork and the artist's name if that information is not in the caption.
- Describe the image's motif (what the image depicts: people, objects, environments, events).
- Describe the image's composition: what is in the foreground, background, to the left and right in the image, if there is a clear direction in the image.
- State whether the image is in color or black and white. If relevant, describe the color. If relevant, describe the light and dark in the image.
- Try to capture the mood the image conveys and describe what has contributed to it. Use vivid language.
- Depending on the context in which the image will be used, it may be relevant to state the image's format, material, and style.

- Consider the user's knowledge level and the context.

Example (when the purpose is to evoke emotions and create immersion):

Description of a woman's face: "the woman smiles thoughtfully, with a dreamy look"

Description of the light and mood: "In backlight, a shimmer of sunlight is created."

4.1.3 Children's Book Illustrations

Children's book illustrations are a representational type of images that resemble reality or depict a fantasy world and fantasy characters. Children's book illustrations create a parallel story to the text or complement and clarify what is told in the text. Therefore, they often need to be described. Children's book illustrations often add details about characters, objects, environments, events, relationships between characters, and paint the mood of the book.

- The image should first be described with a brief overview and then with details.
- Start with the image's composition, for example: "on the left side of the spread it is winter, on the right it is summer."
- Describe objects in the foreground first, then in the background, for example: "In the foreground, there are many flowers. In the background by the bike shed, a dog is sitting. To the left in the periphery, a person with a beret is walking."
- Size relationships: Many fairy tales involve characters changing size and becoming a miniature person or a giant; this must be indicated. If the relationship exists throughout the story, it can be stated initially, but if a change occurs in the middle of the story, this must be pointed out. Use objects in the surroundings for comparison.

Example: "A giant walking through the city is so big that the houses do not reach higher than the giant's soles" while Putte in *Putte's Adventure* in the Blueberry Forest "is so small that he only reaches halfway up the blueberry bush's crown."

- Distinguish between known and new information. Start with a general description and use the indefinite form of nouns when describing a character, object, or referent for the first time. For example: a middle-aged man, a chair. When the referent has been introduced, you can use the definite form the next time. Then the referent is known to the user, as it has already been introduced. This creates continuity in the image description. Use the same principle when, for example, an object has been introduced in the text ("a red ball") and you mention it later in the image description (the red ball").
- Do not repeat what is already in the text.

- Do not anthropomorphize animals and their actions (do not write: "the crow has its arms crossed; the cat looks over its shoulder").
- Avoid using evaluative expressions, such as "he behaves like a real scoundrel", similes ("The riding path looks like a yin and yang symbol"), and visual references "In the image you see ...".
- Do not comment "the image is difficult to describe." Avoid commenting on the layout of the original, for example: "The text is mostly in the middle or at the top of the pages."
- Adapt the language to the target audience and the style to the surrounding text.

MTM makes various types of books accessible for children of different ages with visual impairments/blindness, but also books where an adult braille reader can read the text while a sighted child looks at the pictures. By describing the images, the child and the adult can talk about the images. For the reader with visual impairment to get an idea of the character, the appearance of the images, etc., the book should initially be presented with these aspects in mind.

When there are multiple target audiences who will take part in the image description, for example, a sighted and a non-sighted person as in the case above, it may sometimes be relevant to add information about, for example, colors and similar details that are not relevant in a standard image description. When the adult is the target audience, the language should be adapted to the adult, not the child.

Image descriptions for interleaved books as well as for productions that have the original book in a pocket should follow the same rules as image descriptions in general.

Here is a brief description of the different variants that may occur.

4.1.3.1 Genre Picture Book

A picture book is a children's book whose text is usually accompanied by a large number of drawings or photographs that illustrate the story. It is recommended to begin with an overall description of the book's images. Describe recurring characters and environments. Also describe the technique in which the images are executed.

Then continue with the description of individual images. In picture books with spreads entirely without text, where the story is told solely through the images, the action in the image should primarily be described (rather than a pure image description). There are also children's books with advanced content aimed at a younger audience. When describing images, it is important to consider the potential prior knowledge and reference frames of the target audience.

For overall descriptions in graphic novels, role-playing books, and similar, the client decides if it is needed.

Interleaved Book

A picture book where the book's text and image descriptions are reproduced in braille on plastic sheets between the original book's pages. These books can be read by an adult with a visual impairment or blindness, who reads to a sighted child.

Side-by-Side Book in Pocket

The picture book's text and image descriptions are printed in braille. In the back cover, there is a pocket with the picture book. These books can be read by an adult with a visual impairment or blindness, who reads to a sighted child. These books can also be read by children with a visual impairment or blindness.

Example of Overall Image Description:

"The images in the book are drawings colored with watercolor. The main characters of the story are two children around five years old. The images show both the children's realistic environments and their fantasy worlds. In the fantasy worlds, the colors are stronger and the proportions different. Plants and animals are larger than in reality."

4.1.3.2 Genre Chapter Book

Chapter books, unlike picture books, tell the story primarily through the text. Chapter books can contain plenty of illustrations, but usually decorative images that do not need to be described in detail. It can be a gallery of characters with names and professions or maps. These books can be read by children with a visual impairment or blindness.

Other Image Production Techniques

Other image production techniques can be artistic techniques such as drawings in pencil or ink, paintings in watercolor or acrylic paint, copperplate prints, or computer-generated images. The style or manner of the images is more or less naturalistic but more stripped down compared to photographs. This group includes everything from line drawings to very lifelike depictions of reality or completely abstract images. This type of image can be significant or decorative in context, which determines whether to describe them or not. The images can also be made accessible by adding additional information to an existing caption.

- The image needs to be described if it provides information necessary to understand the subject.
- Study the image in context and assess whether it is significant or decorative. If the image is decorative, no description is needed.
- If the image adds value to the message and its content is not described in the surrounding text, an image description is needed. Describe the most important aspects of the image and then include details if necessary.

- Read the caption. If the content of the image is summarized in the caption, no image description is needed. A solution can be to supplement the caption with additional information. Do not introduce new terms or concepts that do not appear in the surrounding text.
- Check if the image contains any embedded text, that is, text that is part of the image. If there is embedded text, it also needs to be included as part of the image description.

4.2 Schematic and Factual Images

Schematic images (such as diagrams, graphs, maps, tables, timelines, and flowcharts) do not resemble the external world but have partially preserved some structural and conceptual relationships. For example, a map does not resemble a landscape but preserves spatial relationships.

Schematic images show objects and processes, show relationships between their parts and the whole, illustrate a relationship or phenomenon. They show trends, changes over time, and relationships between different variables. Reading schematic images is governed by rules and conventions.

When describing schematic and factual images, the main rule is to describe the content that the image conveys and not how the image is graphically designed. If the schematic or factual image is used in an exercise, it is important to ensure that the image description helps but does not reveal the answer.

Do not create a table that did not previously exist in the book. It is problematic in terms of production. The material you receive as an image describer is already marked, and the "new" table would not be correctly marked and accessible.

4.2.1 Diagrams

Diagrams are a type of schematic image with compressed data that is factual. There are different types of diagrams: for example, bar charts, pie charts, line charts, scatter plots, and Venn diagrams. The image describer must assess the purpose of the diagram and how it will be used: Does the reader need a summary, or should the facts presented in the image be used to solve a task?

First, a brief functional characterization of the different types of diagrams found in MTM's books is listed. Then guidelines that apply to all types of diagrams follow.

Bar Charts

Bar charts show the values of different factors through the height of the bars. This type of diagram is best suited for comparing these values.

Pie Charts

Pie charts show portions of a total amount, often using percentages.

Line Charts

Line charts show linear relationships between two values. The value of one variable is shown on the vertical axis (y-axis) and the value of the other variable on the horizontal axis (x-axis). The line chart shows how a variable increase,

decrease, remains constant, or fluctuates over time, in linear relationships. The line chart is best suited for presenting trends.

Scatter Plots

Scatter plots show relationships between two variables. Often it is about experimental values where it is not known if there is a mathematical relationship to describe their relationships.

Venn Diagrams

Venn diagrams are used in set theory to show overlap between values in different sets. They are often shown as circles that overlap each other to varying degrees.

The guidelines below address all types of diagrams:

- State what type of diagram it is (for example, bar chart, pie chart, line chart, scatter plot, or Venn diagram).
- Mention the labels on the x- and y-axes in graphs and the scales in the diagram.
- Focus on the content and data of the diagram, not on the appearance of the diagram.
- Start with an overview and then, if necessary, continue with a more detailed rendering.
- Present exact data in short sentences if the user is to perform a task and manipulate data.
- Provide a summary in the form of patterns or trends that are directly apparent.
- Base the level of detail in the description on the needs and prior knowledge of the target audience.
- Assess whether the task requires specific values to be read. If so, the description should include detailed information.
- Adapt the description to the target audience.
- It should be clear if the values shown in the diagram are approximate.
- Avoid evaluative descriptions. Consider the choice of words when characterizing data in relation to each other. The description "a steeply rising curve" can be helpful in context and is relatively neutral.
- The description "the number of cases has risen drastically," on the other hand, is subjectively evaluative and should be avoided.

4.2.2 Tables

Tables contain compressed data and are factual. For the description, it is crucial how the table will be used. This, in turn, determines what the user needs to know and what the description must contain.

- Provide a brief account of the table's structure. State the table headings, that is, column and row headings. Always start with the column headings. The number of columns and rows and their possible headings are often enough to give an overview. The results can be mentioned in the form of a list.

Example:

- "A table with two columns, the first refers to the country and the second to the number of traffic fatalities per 100,000 inhabitants. Bulgaria 9.6, Latvia and Romania 9.5, Lithuania 8.3, Croatia 8.2, Poland 7.7 ...".
- Reproduce more complicated tables as long as the description does not become too long so that it no longer serves its purpose of providing an overview.
- A short, very simple list-type table does not need a description.
- Is there a summary of the information that is sufficiently good in the text? If so, only a short note is needed.
- Does the table contain information that is not included in the accompanying text? If so, the data in the table must be provided along with the information in the text.
- State if there is any clear pattern.
- If there are any important rows, such as a total row, they must be included in the image description.

4.2.3 Maps

A map is a conventional representation that partially resembles reality. There are different types of maps: geographical maps, drawings, bus routes, and more.

They serve different purposes: they help the user navigate, orient themselves in an area, find a specific place; they illustrate distances, routes, spatial relationships, and distances between places and show areas and their boundaries (county, communal, municipal, and state boundaries).

Thematic maps also contain additional information, for example, how the land is used by agriculture, how densely populated a geographical area is, what the weather is like in different areas, what historical events occurred in certain places, how certain areas are valued in terms of, for example, political stability, and more.

Maps are made accessible through different levels of descriptive text, depending on the purpose of the map. A short description is suitable for a summary of the size of an empire, for example: "Spain's North American empire stretched in 1770 over the southern part of the continent, from sea to sea." A longer description is needed, however, if one needs to identify all areas that were part of the Spanish empire.

It may be sufficient to reproduce a very simplified map image with an image description that focuses on spatial relationships and the placement of cities, for example, "Smygehuk is located between Ystad and Trelleborg."

- Consider whether the map is significant for the text and message and therefore needs to be described.
- If the map provides essential information, it should be described.
- Identify what type of map it is, and the area covered.
- State the title, describe the areas, regions, and relevant details on the map.
- Do not state the colors in which the data on the map is coded. Instead, describe the categories into which the map is divided.
- Describe the most important content and any patterns based on the map's significance for the text and message.
- Use a summarizing and narrative style if the map is to convey a general understanding rather than specific information ("Spain's North American empire stretched in 1770 over the southern part of the continent, from sea to sea.").
- In cases where all necessary information is provided in the text, no image description is needed. This can be in, for example, chapter books.

Example of Route Description

A map of Dias and de Gama's ocean routes around 1500 ...

"Dias's route starts in Portugal, goes along the west coast of Africa, and ends at the Cape of Good Hope. De Gama's route also starts in Portugal, goes around the Cape of Good Hope, goes along the east coast of Africa, crosses the Indian Ocean, and ends in India."

Example of Description of a Thematic Map

A thematic map of land and resource use. "The map shows the distribution of agriculture, forest, and pasture. The western part of the country shows more pasture, and the eastern part shows more forest..."

4.2.4 Flowchart

Flowcharts convey complex relationships, illustrate conditions or phenomena, and depict abstract reasoning. Flowcharts often contain multiple entry points and reading paths.

- When describing flowcharts, focus on describing the flow. Information about the graphic appearance of the image is not relevant.
- A simple flowchart can be reproduced in the form of a list, either as numbered options or in text form.

4.2.5 Tree Diagrams and Family Trees

Hierarchical tree diagrams are often used to show relationships between different types of activities, ideas, or areas, such as roles in an organization or family relationships. Since the structure of these diagrams can become complex,

it is important to know the purpose of the diagram and the intended target audience.

- Start by stating the title, a general overview of the tree diagram. Mention its different levels and parts.
- Approach the diagram from top to bottom and from left to right.
- Base the description on the target audience's age and knowledge level and the purpose of the diagram.
- Then choose either a factual or a narrative style when describing the tree diagram.

Example of Factual Style in Describing a Family Tree:

"Zalman Tukatsier 1854–1929 married to Hanna 1854–1937. Four children: Berl Tukatsier 1871–1941, Meier Tokazier 1880–1966 married to Sara Tokazier (Leffkovitsch) 1992–1970, Abram Tukatsier 1885–1927 married to Vera Tukatsier."

Example of Narrative Style in Describing a Family Tree:

"The large tree begins with the chocolate maker who was married twice, first to Ketevan and then to Lara. With Ketevan, he had children Lida, Meri, and Stasia (born 1900). With Lara, he had a daughter Christine (born 1907). Christine married Ramas. Stasia married Simon Jasji, and they had a son Kostja (born 1921) and a daughter Kitty Jasji (born 1924)."

4.2.6 Timeline

Timelines allow one to follow events, processes, epochs, and their relationships to each other from a chronological perspective. Timelines can illustrate both long and short periods. They help create a concrete picture of chronology, show how events overlap, and place historical events in a larger context.

- State the title and describe the most important parts of the timeline (dates and events) in a specific order.
- Describe the timeline chronologically, from the earliest to the latest events.
- If there is a general trend, such as several events clustered around a certain timeframe or date, or a significant gap in the sequence of events, a summarizing sentence can be useful.

Example:

"A timeline shows two major events in Kentucky. About 12,000 years ago, people arrived in Kentucky. About 3,000 years ago, the native population began cultivating corn and beans."

4.2.7 Mathematical Expressions

If equations or other mathematical expressions appear as images, they should generally not be described. Exceptions are made if:

- Instructions are given to describe equations in connection with the purchase order, or
- The expression appears in a figure (e.g., a diagram) and is not evident from the caption.

When mathematical expressions are described, they should generally be transcribed into words. Parentheses and mathematical symbols should be spelled out.

Example of an Equation Transcribed into Words:

$$f(x) = \sin\left(\frac{\pi x^2}{2}\right)$$

"f of x equals sine start parenthesis, division with numerator pi x squared, and denominator 2, end parenthesis."

This is to help the synthetic voice when creating audiobooks with synthetic speech. It cannot decode mathematical symbols. Note that it is important for the description to be unambiguous. It may need to be clarified where, for example, exponents, root signs, and divisions begin and end for the expression to be understood even without seeing it. Commas make the speech synthesis take a short pause and can be inserted where it would feel natural to breathe in a reading.

There is an exception when expressions do not need to be transcribed into words. This applies to books in TEXT format, which are to be produced without speech. In that case, mathematical expressions in graphs or diagrams should be marked up with MathML in the image description, as far as possible. Any expressions captured as both image and MathML only need to be described briefly.

Example of a Brief Description of Matrix Multiplication:

$$A = \begin{bmatrix} 6 & -2 & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \\ \textcircled{1} & -\frac{1}{3} & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{bmatrix} \leftarrow \text{multiplier} = \frac{1}{6}$$

"Row 1 is multiplied by 1/6, yielding 1 in column 1, row 1."

This assumes that all values in the matrix are also represented as MathML outside the image. It is only the "extra" information in the form of, for example, comments or markings that needs to be described.

Character Selection

Use standard Unicode characters for letters that appear in a mathematical context, regardless of whether MathML is used or not. That is, not [Mathematical Alphanumeric Symbols](#), but [Basic Latin](#) or [Greek](#). This is because the speech synthesis cannot decode them.

4.2.8 Musical Notation

Musical notation in running text is generally considered images and should therefore be described. General information about how the notation system is constructed usually does not need to be provided. Simpler notation can be fully reproduced. For more complicated notation examples, scores, and so on, a selection should be made, and only typical motifs, themes, or other typical features should be presented briefly.

5 AI-Generated Image Descriptions

Artificial intelligence (AI) has made significant progress in various fields in recent years. In the field of image description, deep learning algorithms have been used to classify images, which is necessary to create systems and applications that help visually impaired people perceive their surroundings (e.g., Seeing AI, Be My Eyes). AI tools can currently be used to generate descriptive sentences from images, helping blind people understand visual content in everyday life. Experiments have also begun with generating diagram descriptions using labels and surrounding text, but the process is not fully automatic and still requires human expertise.

It is **not allowed** to use AI tools to create image descriptions for books produced for MTM. MTM's mission to produce accessible literature is based on an exception to copyright law, and it is of utmost importance that all suppliers working on behalf of MTM respect this and are aware that we work with copyrighted material.

AI tools *can* be a good aid, but they should always be linked to an editorial resource that reviews and adjusts the automatic image descriptions. The editor would greatly benefit from following the guidelines and recommendations in the example collection.

Below, we provide examples of some areas where AI algorithms are currently inferior to human ability to describe images.

- The AI tool has difficulties assessing the recipient's age, prior knowledge, and needs. It can fail to choose appropriate language, appropriate level, or adapt the content to the user.
- The AI tool has difficulties assessing what is most important in the image. AI tools do not evaluate what is most important/most significant in the image. They describe everything in the image at the same level of detail.
- The AI tool does not consider the image's context if not prompted to so and provided with all the text. It does not read whether the image is already described in the main text or in the image descriptions. It cannot assess the purpose of the image and how it complements the text.
- The AI tool does not consider how the image will be used. For example, if the image is a diagram to be used to solve a task later, the answer to the task cannot be revealed in the image description.
- The AI tool has difficulties assessing what has been described previously. The AI tool treats images in isolation. It will describe each new image from the beginning, even if things and people have been mentioned before, which creates redundant information and cognitively burdens the user.
- The AI tool has difficulties to notice the unusual. It cannot contrast things against how they usually/should look or if an image is created to show such a contrast.

- The AI tool has difficulties interpreting sequences. It cannot interpret sequences and parallel actions, for example, if the same person/character appears in several places in an image. This type of image is common in picture books, such as the books about Pettson and Findus.
- The AI tool has difficulties interpreting interaction between different image elements. Images of people often convey interaction and relationships that can be important for understanding the image. The AI tool often describes the image statically. There is a lack of dynamic description of events, for example, how something happens and how characters interact with each other and the environment.