Writing for every reader

Last modified: 2026-11-13 12:24:54 (PST)

> **AI-Assisted Authorship**
>
> This chapter was written with the assistance of GitHub Copilot, which expanded on outlined ideas and draft notes provided by the author. The content represents the author's perspective and has been reviewed for accuracy, but the detailed prose was generated through AI assistance.

Write simple, direct prose that every reader can follow. Some of your readers learned English as a second language. Some are neurodiverse. For example, some autistic readers take jokes and similes literally (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021, Use Literal Language). Some readers have a memory impairment, so a long sentence with nested clauses is hard for them to follow (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021, Avoid Double Negatives or Nested Clauses). Plain wording helps these readers most, and it helps every other reader too.

This is an equity issue. Idioms, slang, clichés, unnecessary jargon and other specialized language exclude readers who do not share the writer's background. Those readers must first work out what the words mean, and only then can they think about the ideas. The plain language guidelines put it this way: "The first rule of plain language is: **write for your audience**" (Plain Language Action and Information Network 2026, Audience). Your audience includes all of these readers.

The rules in this chapter come mostly from three sources:

- the W3C guidance on making content usable for people with cognitive and learning disabilities (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021);
- the US government's plain language guidelines (Plain Language Action and Information Network 2026);
- the GitHub Docs style guide, which is written for a global audience and for translation (GitHub 2026a, 2026b).

Each rule below comes with an example. Several rules extend sections in other chapters, and they link to those sections instead of repeating them.

Use literal language

Say exactly what you mean. Do not use idioms, clichés, slang, jokes or sarcasm. An idiom is a phrase whose meaning is different from the meanings of its words, such as "blow up" for "become very large". A reader who has not encountered the phrase before has to guess what it means. A reader who takes words literally may guess wrong. The W3C guidance says: "Use literal and concrete language" and "Do not use metaphors and similes unless you include an explanation" (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021, Use Literal Language). GitHub Docs says: "Avoid turns of phrase, idioms, and slang that are specific to a particular region or country" (GitHub 2026a).

✗ Which way does this course lean?

✓ Does this course put prediction or inference first?

The word "lean" in the first version is an idiom. The second version asks the literal question.

Table 1 lists more examples from statistics and machine learning.

Table 1: Idioms, clichés and slang, with literal alternatives

Idiom, cliché or slang	Literal wording
the estimates blow up	the estimates become very large
the model is a black box	we cannot read the model's decision rule from its parameters
flexibility buys lower bias	more flexible models tend to have lower bias
this assumption costs nothing	this assumption loses no generality
a mean of zero buys one identity	a mean of zero implies one identity
that trade is the whole bargain of gradient descent	gradient descent decreases the objective function only when each step is small enough
the recommendation is trustworthy only as far as the step is small	the objective function decreases only for steps that are small enough
Fix $\varepsilon > 0$.	Let $\varepsilon > 0$.
Fix a value t.	Choose a value t.
a ton of data	a large data set
the method falls flat on small samples	the method has high error on small samples
at the end of the day, we want low test error	our goal is low test error

Idiom, cliché or slang	Literal wording
Write the bias of an estimator.	Derive an expression for the bias of an estimator.
Write X in terms of Y.	Express X in terms of Y.
You will meet this notation again.	You will encounter this notation again.
We first meet the odds in logistic regression.	We first encounter the odds in logistic regression.
Take a sample of size n and call the sample mean $\bar{x}$.	Consider a sample of size n and denote the sample mean by $\bar{x}$.
the residual lies in $(-1, 1)$	the residual is in $(-1, 1)$
the certain distribution sits at 0, the uniform distribution at $\log 4$, and the skewed distribution lies between them	the certain distribution has entropy 0, the uniform distribution has entropy $\log 4$, and the skewed distribution has entropy between those two values
the car sits at height 1	the car is at height 1
the car is drawn at height 1	the car is plotted at height 1
the origin sits on the side of the dome	the origin is on the side of the dome
the two discriminant scores agree at x	the two discriminant scores are equal at x
the derivative vanishes at $\hat{\theta}$	the derivative equals 0 at $\hat{\theta}$
setting $K = 2$ recovers the binary log-odds	setting $K = 2$ gives the binary log-odds
the probability p lives in $(0, 1)$	the probability p is in $(0, 1)$
the linear model knows nothing about the range $[0, 1]$	the linear model does not constrain predictions to $[0, 1]$
Say which class the classifier predicts at $x = 2$.	State which class the classifier predicts at $x = 2$.
the noise model says that the errors have mean zero	in the noise model, the errors have mean zero
as Theorem 2 says, the estimator is unbiased	by Theorem 2, the estimator is unbiased
the table reads as a list of errors	the table seems to be a list of errors
the formula comes across as a definition	the formula seems to be a definition

For a reader's first contact with a term or notation, write "encounter". Do not write "meet", which treats the idea as a person. Do not write "see" either, because it names one sensory modality, and readers also listen to text or read it in braille.

Mathematical writing has idioms too. In an exercise, "Write the bias of an estimator" means "derive an expression for the bias of an estimator", and "Say which estimator has smaller

variance" means "state which estimator has smaller variance". A reader who takes "write" or "say" literally may copy the definition or give an informal remark instead of deriving or stating the result. Name the action instead:

- "derive" for a result the reader must work out;
- "express" for a rewrite in terms of other quantities;
- "state" for a result the reader recalls;
- "give" for a value the reader reads off.

Idioms are a kind of metaphor, so Avoid vague and metaphorical language applies to them too, and gives more examples.

Use common words, and define technical ones

Use the most common word that is still correct. The W3C guidance says to "Use common and clear words in all content" and to "Remove or explain uncommon acronyms, abbreviations, and jargon" (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021, Use Clear Words).

> ✗ We utilize a heuristic to ascertain the optimal number of clusters.
>
> ✓ We use the "elbow" method to choose the number of clusters: we plot the within-cluster sum of squares against the number of clusters, and pick the point where the curve stops falling steeply.

The second version uses common words, and it says what the method does instead of calling it a "heuristic".

Some technical terms are necessary. Minimize unnecessary jargon explains when to keep a technical term and how to define it.

Say when a common word has a technical meaning

Many statistical terms are everyday words with a different technical meaning. Examples include *bias*, *significant*, *confidence*, *normal*, *regression*, *error* and *likelihood*. A reader who knows only the everyday meaning will read the sentence as a different claim. The W3C guidance says: "Do not invent new words or give words new meanings" and "Do not expect people to learn new meanings for words just to use your content" (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021, Use Clear Words). A technical course needs these terms. So say that the technical meaning differs, and give the technical meaning in plain words.

✗ The model is biased.

✓ The model is biased: at a given input, its average prediction over many training sets differs from the true mean outcome at that input. (Here *bias* has its statistical meaning; it does not mean the model is unfair.)

Once you define the term, use the same word for the same idea every time. The plain language guidelines say that using a different term "may cause the reader to wonder if you're referring to the same group" (Plain Language Action and Information Network 2026, Use the same terms consistently).

✗ We split the data into a training set and a test set. We then compute the error on the held-out data.

✓ We split the data into a training set and a test set. We then compute the error on the test set.

Spell out abbreviations

An acronym or initialism, such as "MSE", is allowed; do not clip a term covers shortened phrases. The plain language guidelines say abbreviations "constantly require the reader to look back to earlier pages" (Plain Language Action and Information Network 2026, Minimize abbreviations). Do not use the Latin abbreviations *e.g.* and *i.e.*: "Few people know what they mean, and they often confuse the two" (Plain Language Action and Information Network 2026, Use examples). Write "for example" or "that is".

✗ Tree ensembles, e.g. RF and GBM, usually reduce the MSE of a single CART.

✓ Ensembles of trees, for example random forests and boosted trees, usually have lower mean squared error than a single regression tree.

Do not clip a term

Do not shorten a term by dropping words, even after introducing the short form. This covers clipped phrases, such as "objective" for "objective function". Write the full term at every use. Do not leave out words that the reader must supply either, as in "the cost is convex", which leaves out "function" and what it is a function of. A leftover gap of this kind is a semantic ellipsis. An acronym or initialism, such as "MSE" for "mean squared error", is neither clipped nor elliptical, and this rule allows it. Write the full phrase says when a clipped name is correct: only for the value of a function at a specific input.

✗ An *objective function*, or *objective* for short, is the quantity we minimize. The objective as a function of the parameters is a paraboloid.

✓ An *objective function* is the quantity we minimize. The objective function of the parameters is a paraboloid.

Write positive statements

State what is true, not what is not false. A double negative makes the reader turn "no" into "yes" before they can use the sentence. The W3C guidance says not to "use a double negative to express a positive" (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021, Avoid Double Negatives or Nested Clauses), and the plain language guidelines give the same advice (Plain Language Action and Information Network 2026, Use positive language).

✗ We do not reject the null hypothesis unless the p-value is not above 0.05.

✓ We reject the null hypothesis when the p-value is 0.05 or less.

Rewrite noun strings and hidden verbs

A noun string is a row of nouns that act as adjectives, such as "test set error rate estimate". The reader has to work out which noun modifies which. A hidden verb (also called a *nominalization*) is a verb turned into a noun, such as "perform an estimation of" in place of "estimate". GitHub Docs advises against both. It says noun strings "can lead to incorrect translations", and nominalizations make sentences "longer, harder to understand, and harder to translate" (GitHub 2026a). The plain language guidelines say to open up a noun string by "using more prepositions and articles" (Plain Language Action and Information Network 2026, Avoid noun strings).

✗ We report the test set error rate estimate variance.

✓ We report the variance of the estimated error rate on the test set.

✗ We performed an evaluation of the model's prediction accuracy.

✓ We evaluated how accurately the model predicts.

Put one idea in each sentence

Long sentences with nested clauses make the reader hold several unfinished ideas at once. That is hard for readers with a memory impairment (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021, Keep Text Succinct), and for readers who are translating as they go. The W3C guidance says: "Use short sentences. Have only one point per sentence" (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021, Keep Text Succinct). It also says that "Sentences that have more than one point usually have more than one linking word such as 'and' or 'but' ".

> ✗ Because the lasso penalty, which is the sum of the absolute values of the coefficients, is not smooth at zero, some coefficients, especially when λ is large, are set exactly to zero.
>
> ✓ The lasso penalty is the sum of the absolute values of the coefficients. The absolute value function is not smooth at zero. Because of this, the lasso sets some coefficients exactly to zero. Usually, more coefficients are zero when λ is larger.

A sentence that states a result and then adds "which" and "and" clauses about it also holds several points. Give each fact its own sentence, in the order a reader needs them.

> ✗ The constant prediction has squared-error risk 5, which is larger than the risk 1 of the conditional mean, and strictly so because the two predictions differ.
>
> ✓ Consider the constant prediction. Its squared-error risk is 5. The risk of the conditional mean is 1. So the constant prediction has the larger risk. The inequality is strict because the two predictions differ.

Limit complex sentence structures and Lead with the subject give more examples.

State every step and every point

Write every step of an argument or a procedure. Do not leave a step for the reader to fill in, and do not hint at a point you could state. The W3C guidance says readers may need "definitions or explanations for implied or ambiguous information" such as jokes, sarcasm and metaphors (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021, Explain Implied Content). In an experiment, rewriting an instructional text to state each link that readers had to infer greatly improved what students recalled of it (Britton and Gülgöz 1991).

> ✗ Of course, nobody would use accuracy here.
>
> ✓ Do not judge this classifier by accuracy alone. Only 1% of the emails are spam, so a classifier that labels every email "not spam" has 99% accuracy but finds no spam.

The first version implies a reason and never states it. A reader who does not already know the reason learns nothing. Words like "obviously" and "clearly" hide a step in the same way; Do not substitute flippancy for support explains what to write instead.

Say what numbers and symbols mean

Do not leave a number or a formula for the reader to interpret. Some readers find numbers hard to process (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021, Provide Alternatives for Numerical Concepts). Text-to-speech tools can also read numbers and symbols wrongly (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021, Use Clear, Unambiguous Formatting and Punctuation). The CDC Clear Communication Index scores a document on whether it explains what its numbers mean (Centers for Disease Control and Prevention 2026).

For a number, say what it means in words:

> ✗ The AUC is 0.83.
>
> ✓ The AUC is 0.83. That is, if we pick one spam email and one non-spam email at random, the model gives the spam email the higher score 83% of the time.

For a formula, say each symbol's meaning in words at first use, and then say what the formula does in a sentence:

> ✗ $\hat{\beta} = (X^\top X)^{-1} X^\top y$.
>
> ✓ Let X be the $n \times p$ matrix of inputs, with linearly independent columns, and let y be the vector of n outcomes. Let $X^\top$ be the transpose of X. The least squares estimate of the coefficients is $\hat{\beta} = (X^\top X)^{-1} X^\top y$. In words, $\hat{\beta}$ is the coefficient vector that makes the sum of squared prediction errors on the training data as small as possible.

The Mathematical notation chapter has more rules for notation.

Use examples that every reader knows

Choose examples that do not need a particular culture, country, sport or hobby to understand. GitHub Docs says to "Use examples that are generic and can be understood by most people" and to "Avoid examples that are controversial or culturally specific to a group" (GitHub 2026b).

✗ Predict a quarterback's passer rating from his completion percentage.

✓ Predict whether an email is spam from the words it contains.

The first example needs knowledge of American football before the reader can think about the statistics. It also assumes the player is a man. GitHub Docs also says to "Avoid gender-specific words" (GitHub 2026b). Use "they" for a person whose gender you do not know:

✗ Each student submits his code by Friday.

✓ Each student submits their code by Friday.

Keep a predictable structure

Readers find text easier to follow when each part has one purpose, and when parts of the same kind always look the same. The W3C guidance says to "Keep paragraphs short. Have only one topic in each paragraph", to "Try to have the aim of the paragraph or chunk at the beginning", and to "Use bulleted or numbered lists" and "short descriptive headings" (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021, Keep Text Succinct). It also recommends a short summary at the start of a long document (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021, Provide Summary of Long Documents and Media). In course materials, the same advice applies across pages: the Cornell Center for Teaching Innovation suggests "a predictable weekly structure and consistent layout" (Cornell University Center for Teaching Innovation 2026).

✗ A paragraph that starts with the history of the method, moves on to its assumptions, and states its main use in the last sentence.

✓ A paragraph that starts with "Use cross-validation to choose a tuning parameter", and then explains how.

Put lists in bullet points and the Paper organization chapter cover lists and headings in more detail.

Present ideas in more than one form

Readers differ in what helps them understand. Universal Design for Learning (UDL) is a framework for teaching diverse learners. The Cornell Center for Teaching Innovation describes UDL as giving students several ways to "Deliver content (text, audio, visuals, notes)" (Cornell University Center for Teaching Innovation 2026). Learners understand words and pictures together better than words alone (Mayer 2020), and a practice guide from the Institute of Education Sciences recommends combining graphics with verbal descriptions (Pashler et al.

2007, recommendation 3). For a technical idea, give at least two of these forms: a sentence, a formula, a figure, a worked example with numbers, or code.

> ✗ A slide that gives only the formula for k-fold cross-validation.
>
> ✓ A slide that gives the formula, a figure of the data split into k folds, and one sentence saying what the formula averages.

Each figure also needs a caption and alt text, so that a reader who cannot see it gets the same information.

Test your writing with a reader

You cannot judge your own writing from the reader's side, because you already know what you meant. The W3C guidance ends with "Test with real users!" (W3C Cognitive and Learning Disabilities Accessibility Task Force 2021). The plain language guidelines describe *paraphrase testing*: ask a reader to explain the text in their own words, and compare their version with what you meant (Plain Language Action and Information Network 2026, Paraphrase testing).

> ✗ Reread your lecture notes yourself, and decide they are clear.
>
> ✓ Ask a student to explain a paragraph back to you. Where their version differs from what you meant, rewrite that paragraph.

Britton, Bruce K., and Sami Gülgöz. 1991. "Using Kintsch's Computational Model to Improve Instructional Text: Effects of Repairing Inference Calls on Recall and Cognitive Structures." *Journal of Educational Psychology* 83 (3): 329–45. https://doi.org/10.1037/0022-0663.83.3.329.

Centers for Disease Control and Prevention. 2026. *CDC Clear Communication Index*. https://www.cdc.gov/ccindex/index.html.

Cornell University Center for Teaching Innovation. 2026. *Universal Design for Learning*. https://teaching.cornell.edu/teaching-resources/building-inclusive-classrooms/universal-design-learning.

GitHub. 2026a. *Style Guide*. GitHub Docs. https://docs.github.com/en/contributing/style-guide-and-content-model/style-guide.

GitHub. 2026b. *Writing Content to Be Translated*. GitHub Docs. https://docs.github.com/en/contributing/writing-for-github-docs/writing-content-to-be-translated.

Mayer, Richard E. 2020. *Multimedia Learning*. 3rd ed. Cambridge University Press. https://doi.org/10.1017/9781316941355.

Pashler, Harold, Patrice M. Bain, Brian A. Bottge, et al. 2007. *Organizing Instruction and Study to Improve Student Learning*. {IES} Practice Guide NCER 2007-2004. National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/WWC/Docs/PracticeGuide/20072004.pdf.

Plain Language Action and Information Network. 2026. *Federal Plain Language Guidelines*. https://github.com/GSA/plainlanguage.gov/tree/main/_pages/guidelines.

W3C Cognitive and Learning Disabilities Accessibility Task Force. 2021. *Making Content Usable for People with Cognitive and Learning Disabilities*. W3C Working Group Note, 29 April 2021. https://www.w3.org/TR/coga-usable/.