

# Annotation guidelines

Annotators are required to extract obligations, prohibitions, and permissions from acts of UK legislation and rephrase them into the following *structured templates*:

- (1) IT IS (OBLIGATORY or PROHIBITED or PERMITTED)  
FOR someone  
TO do something  
(WHEN/IF/WHERE or ONLY IF or BEFORE or AFTER or UNLESS)  
some conditions apply

The slots IT IS, FOR, and TO are *mandatory* (i.e., they occur in every annotations), while conditions are *optional*: annotations may include no conditions, a single condition, or multiple conditions. It is however rare, although not impossible, that annotations will include more than two conditions. Each condition is identified by either keyword WHEN/IF/WHERE, ONLY IF, BEFORE, AFTER, or UNLESS.

Annotators will have to *copy* the text from the UK acts and *paste* it into the template slots, while possibly rephrase it to make it grammatical and fluent. Furthermore, in cases where multiple options are available on a template's slot, annotators will have to *select* the option that better fits the intended meaning (e.g., first of all they will have to choose between OBLIGATORY, PROHIBITED, and PERMITTED).

A Web tool is provided for the task. The tool will save the annotations in a text file that the annotators can manually update, in cases where they change their decisions.

The annotations so created might be very useful for AI/LegalTech applications; for instance, they can be used to train Large Language Models such as OpenAI's GPT, with the aim of *automatically* analysing legislation.

The rest of this document shows examples of annotations as well as general directions that the annotators should follow to carry out the annotation task. In addition, two UK acts have been already fully annotated: the Equality Act 2010<sup>1</sup> and the Police, Crime, Sentencing and Courts Act 2022<sup>2</sup>. The annotations from these two acts are further examples made available to the annotators.

---

<sup>1</sup><https://www.legislation.gov.uk/ukpga/2010/15>

<sup>2</sup><https://www.legislation.gov.uk/ukpga/2022/32>

# 1 Some simple examples

Let's start with some simple examples of the task: the obligations for UK drivers and riders available at <https://www.gov.uk/legal-obligations-drivers-riders>. The text of the first norm is:

- (2) Before you drive or ride you must have the correct driving licence

The straightforward annotation of this example appears to be the following:

- (3) IT IS OBLIGATORY  
FOR you  
TO have the correct driving licence  
BEFORE you drive or ride

However, (3) is not the single way in which the template might be filled with respect to the example in (2). For instance, the pronoun “you” does not only refer to the reader(s) of the norm; in the intended meaning of (2) is that the *everyone* is subject to the obligation. In light of this, an annotator could decide to *rephrase* the copied and pasted text as follows:

- (4) IT IS OBLIGATORY  
FOR everyone  
TO have the correct driving licence  
BEFORE driving or riding

It is even possible to rephrase the obligation in (4) into a prohibition, while changing the keyword on the condition from BEFORE to UNLESS and by further rephrasing the texts in the slots, as follows:

- (5) IT IS PROHIBITED  
FOR a person  
TO drive or ride  
UNLESS the person has the correct driving licence

The annotators must be aware that:

**IMPORTANT:** all variants listed in (3)-(5) are correct annotations of (2): they are all paraphrases of the same meaning conveyed by the sentence. Still, since the task should be “as natural as possible”, annotators are required to annotate the obligations, permissions, and prohibitions *while minimally changing the original wording*. In other words, they should not “think too much” about the proper way to fill in the form’s fields. In light of this, they should prefer the annotation in (3).

As a second example, consider (6), which is also taken from <https://www.gov.uk/legal-obligations-drivers-riders>:

(6) **Learner drivers**

You must:

- be supervised by a qualified driver (except if riding a motorcycle)

Contrary to the previous example, the template cannot be filled by simply copying and pasting text from the obligation in (6): the text among round brackets, which denotes an **UNLESS** condition, must be rephrased depending on the text chosen for the **FOR** slot. A natural way to fill the template from (6) could be then:

(7) **IT IS OBLIGATORY**

**FOR** learner drivers

**TO** be supervised by a qualified driver

**UNLESS** they ride a motorcycle.

but also:

(8) **IT IS OBLIGATORY**

**FOR** a learner driver

**TO** be supervised by a qualified driver

**UNLESS** the learner driver rides a motorcycle.

In which the text in the **UNLESS** condition depends on the plural/singular choice for the text in the **FOR** slot.

Of course many more way to annotate (6) into the template are possible, for instance:

(9) **IT IS OBLIGATORY**

**FOR** you

**TO** be supervised by a qualified driver

**WHEN/IF/WHERE** you are a learner driver

**UNLESS** you ride a motorcycle.

But these looks less “natural” than (7) or (8), so that one of the latter should be preferred.

In any case:

**IMPORTANT:** the text of the final annotations must always read grammatically and fluently. To double-check so, annotators are invited to read the final template as a whole, e.g.: “it is obligatory for a learner driver to be supervised etc.”. In cases where the text in the template contains grammatical errors (e.g., misaligned plurals and singulars) or it does not read smoothly, annotators are required to rephrase it in order to make it grammatical and fluent.

Sometimes, annotators could be “tempted” to paraphrase the text because the obligation is in *passive form*, and the subject is an *inanimate object*, which of course cannot be in charge of complying with the obligation. Consider:

(10) The vehicle must:

- be registered with DVLA

The obligation of course applies to *the owner* of the vehicle, not to the vehicle itself, so that a meticulous annotator could decide to spell this out in the **FOR** slot.

Although such a choice would make perfectly sense, annotators are again invited to avoid spelling out the implicit text (unless this is strictly necessary) in order to minimize manual amending. In light of this, the suggested annotation of the obligation in (10) is:

(11) **IT IS OBLIGATORY**  
**FOR** the vehicle  
**TO** be registered with DVLA.

i.e., with the **FOR** slot applying to the (inanimate) subject of the passive verb, even if it appears a bit “weird” for an inanimate object to be obliged to do something.

Note also that the annotation in (11) does not have any condition; as explained earlier, only the first three slots of the template are mandatory while conditions are optional.

The rest of the document will show annotations of more complex examples, taken from the Equality Act 2010<sup>3</sup> and the Police, Crime, Sentencing and Courts Act 2022<sup>4</sup>.

## 2 Regulative and constitutive norms

Legal theory usually distinguishes between *regulative norms* and *constitutive norms*. The former are obligations, permissions, and prohibitions; the latter are definitions of the legal concepts used in the regulative norms.

**IMPORTANT:** only regulative norms must be annotated. Moreover, again to minimize copying, pasting, and rephrasing, the definitions of the legal concepts used in the regulative norms should not be copied and pasted in the template’s slots.

Let’s see an example: the annotation of the first three paragraphs of Article 1 of the Police, Crime, Sentencing and Courts Act 2022:

---

<sup>3</sup><https://www.legislation.gov.uk/ukpga/2010/15>

<sup>4</sup><https://www.legislation.gov.uk/ukpga/2022/32>

- (12) (1) The Secretary of State must in each financial year—
- (a) prepare a police covenant report, and
  - (b) lay a copy of the report before Parliament.
- (2) A police covenant report is a report about—
- (a) the health and well-being of members and former members of the police workforce,
  - (b) the physical protection of such persons,
  - (c) the support required by members of their families, and
  - (d) any other matter in relation to members or former members of the police workforce, or a particular description of such persons, that the Secretary of State considers appropriate, so far as these matters relate to the fact that the persons concerned are members or former members of the police workforce.
- (3) In preparing a police covenant report the Secretary of State must have regard in particular to—
- (a) the obligations of and sacrifices made by members of the police workforce, and
  - (b) the principle that it is desirable to remove any disadvantage for members or former members of the police workforce arising from their membership or former membership.

The first and the third paragraphs in (12) convey two obligations for the Secretary of State. These might be respectively annotated as in (13) and (14):

- (13) **IT IS OBLIGATORY**  
**FOR** the Secretary of State  
**TO** (a) prepare a police covenant report, and (b) lay a copy of the report before Parliament.  
**BEFORE** the financial year ends.
- (14) **IT IS OBLIGATORY**  
**FOR** the Secretary of State  
**TO** have regard in particular to (a) the obligations of and sacrifices made by members of the police workforce, and (b) the principle that it is desirable to remove any disadvantage for members or former members of the police workforce arising from their membership or former membership.  
**WHEN/IF/WHERE** preparing a police covenant report.

On the other hand, the second paragraph in (12) does not contain any obligation, prohibition, or permissions but it only *defines* what a police covenant report is. In other words, the second paragraph only adds further details to describe a concept used in the other

two paragraphs. In the literature it is usually said that the second paragraph contains a *constitutive norm*, i.e., a definition of a legal concept used in the regulative norms.

As stated earlier, the annotation task only focuses on regulative norms while definitions, i.e., constitutive norms, must be ignored. Therefore, in the example above a “police covenant report” can be taken for granted and so there is no need to copy and paste its definition in the template’s slots. In other words, annotators should **NOT** expand the condition in (14) as follows:

**WHEN/IF/WHERE** preparing a police covenant report, *which is a report about— (a) the health and well-being of members and former members etc.*

but they can just assume that it is well-known what a police covenant report is.

### 3 More complex examples - different conditions

It has been explained above that the annotation task consists of identifying regulative norms from UK acts and filling the provided template by copying and pasting text from these acts. The final text of the template must read grammatically and fluently, which may require annotators to rephrase parts of it. Nevertheless, annotators should try to *minimally* do so, i.e., to keep the text as close as possible to the original one.

Sometimes, it could be hard to fill the template because the original norms contain particular linguistic constructions that mix the text of the conditions with the text of the other slots. For this reason, the template specifies different keywords for the conditions, intended to help the annotators identify them.

The most frequent type of condition occurring in the norms is the one associated with the keyword **WHEN/IF/WHERE**. This keyword actually refers to the *three* distinct keywords, which have been merged together in that they convey the same meaning.

Sometimes, the first keyword “when” occurs in the original text, as in the following regulative norm from the Equality Act 2010:

- (15) An authority to which this section applies must, *when* making decisions of a strategic nature about how to exercise its functions, have due regard to the desirability of exercising them in a way that is designed to reduce the inequalities of outcome which result from socio-economic disadvantage.

Annotated as follows:

- (16) **IT IS OBLIGATORY**  
**FOR** an authority to which this section applies  
**TO** have due regard to the desirability of exercising its functions in a way that is designed to reduce the inequalities of outcome which result from socio-economic disadvantage  
**WHEN/IF/WHERE** making decisions of a strategic nature about how to exercise its functions.

Sometimes it is the keyword “if” that occurs in the norm’s text, as in:

- (17) *If it appears to a court in which proceedings are pending that a claim or counter-claim relating to an equality clause or rule could more conveniently be determined by an employment tribunal, the court may strike out the claim or counter-claim.*

Annotated as follows:

- (18) **IT IS PERMITTED**  
**FOR** a court in which proceedings are pending  
**TO** strike out a claim or counter-claim relating to an equality clause or rule  
**WHEN/IF/WHERE** it appears to the court that the claim or counter-claim could more conveniently be determined by an employment tribunal.

Sometimes it is the keyword “where” that occurs in the norm’s text, as in:

- (19) The regulations may provide for an employment tribunal to have power, *where* a person fails to comply with an order to carry out an equal pay audit, to order that person to pay a penalty to the Secretary of State of not more than an amount specified in the regulations.

Annotated as follows:

- (20) **IT IS PERMITTED**  
**FOR** regulations  
**TO** provide for an employment tribunal to have power to order that person to pay a penalty to the Secretary of State of not more than an amount specified in the regulations  
**WHEN/IF/WHERE** a person fails to comply with an order to carry out an equal pay audit.

Sometimes, although neither of the three keywords occurs in the text of the condition, the latter might be easily framed into one of them; this is often the case when the preposition “in” is followed by a gerund verb, as it has been already exemplified above in (14). Another example is:

- (21) An authorised person must have regard to the code of practice for the time being in force under section 42 *in exercising, or deciding whether to exercise*, the power in subsection (1).

Annotated as follows:

- (22) **IT IS OBLIGATORY**  
**FOR** An authorised person  
**TO** have regard to the code of practice for the time being in force under section 42  
**WHEN/IF/WHERE** exercising, or deciding whether to exercise, the power in subsection (1).

WHEN/IF/WHERE is not the only keyword marking conditions. The others are:

- BEFORE or AFTER
- UNLESS
- ONLY IF

An example with BEFORE has been already shown above in (2); often, and this was indeed the case in (2), the keywords “before” and “after” occur in the text of the norm. However, again this is not always the case, as already exemplified above in (13), where “in each financial year” has been rephrased into “BEFORE the financial year ends”. Another example is:

- (23) Once a strategy has been prepared under this section for a local government area, the specified authorities for the area must—
- (a) publish the strategy,
  - (b) keep the strategy under review, and
  - (c) from time to time prepare and implement a revised strategy.

Annotated as follows:

- (24) IT IS OBLIGATORY  
FOR The specified authorities for a local government area  
TO
- (a) publish a strategy prepared under this section for the local government area
  - (b) keep the strategy under review, and
  - (c) from time to time prepare and implement a revised strategy
- AFTER the strategy has been prepared.

In (24), note that not only “once” has been rephrased into “after” but the whole text has been rephrased in order to make it grammatical and fluent. Specifically, since the keyword AFTER occurs after the keyword TO, the enumerated item “(a)” becomes the first mention of the strategy referred to by the norm; for that reason, “the strategy” has been replaced as “a strategy prepared under this section for the local government area” in the enumerated item “(a)”, while “the strategy” has been used in all subsequent mentions.

Concerning UNLESS, an example has been provided above in (7). In this example, the norm did not contain the word “unless”, but rather the equivalent locution “except of”. More complex examples associated with this keyword might be found in legislation, because it is often the case that a norm overrides another norm, but the text of the two norms is found in different subsections. For instance, consider Article 63, subsections (1) and (2), of the Police, Crime, Sentencing and Courts Act 2022:



- (25) (1) A person commits an offence if they trespass on land with the intention of—
- (a) using a dog to search for or to pursue a hare,
  - (b) facilitating or encouraging the use of a dog to search for or to pursue a hare, or
  - (c) enabling another person to observe the use of a dog to search for or to pursue a hare.
- (2) It is a defence for a person charged with an offence under subsection (1) to prove that they had a reasonable excuse for the trespass mentioned in that subsection.

Subsection (2) in (25) conveys an exception of the prohibition conveyed by subsection (1). Therefore, the text of subsection (2) has been integrated in the annotation of subsection (1) and rephrased as an **UNLESS** clause, as follows:

- (26) **IT IS PROHIBITED**  
**FOR** A person  
**TO** commit an offence by trespassing on land with the intention of—
- (a) using a dog to search for or to pursue a hare,
  - (b) facilitating or encouraging the use of a dog to search for or to pursue a hare,  
or
  - (c) enabling another person to observe the use of a dog to search for or to pursue a hare
- UNLESS** they had a reasonable excuse for that trespass.

Note also that in (26) the linguistic construction “commits an offence if” has been rephrased as “it is prohibited to”; of course, other choices are possible, e.g., rephrasing the linguistic construction as “it is obligatory to not”. Rephrasing the norms in (25) as a prohibition rather than an obligation appeared to be more “natural”.

Let’s now show an example involving the keyword **ONLY IF**. This keyword normally occurs as it is in the norms’ text, e.g., in Article 146, subsection (1) and (2), of the Equality Act 2010:

- (27) (1) A qualifying person (P) may make a complaint to an employment tribunal that a term is void, or that a rule is unenforceable, as a result of section 145.
- (2) But subsection (1) applies only if—
- (a) the term or rule may in the future have effect in relation to P, and
  - (b) where the complaint alleges that the term or rule provides for treatment of a description prohibited by this Act, P may in the future be subjected to treatment that would (if P were subjected to it in present circumstances) be of that description.

Annotated as follows:

(28) **IT IS PERMITTED**

**FOR** A qualifying person (P)

**TO** make a complaint to an employment tribunal that a term is void, or that a rule is unenforceable, as a result of section 145

**ONLY IF**

- (a) the term or rule may in the future have effect in relation to P, and
- (b) where the complaint alleges that the term or rule provides for treatment of a description prohibited by this Act, P may in the future be subjected to treatment that would (if P were subjected to it in present circumstances) be of that description.

## 4 Even more complex examples - multiple conditions

It has been explained above that conditions are *optional*, i.e., that there could be instantiations of the template that do not include any condition, e.g., example (11) above. The previous section, on the other hand, only showed examples including a single condition.

This section will present and discuss examples including more than one condition. The annotations of the Equality Act 2010's and the the Police, Crime, Sentencing and Courts Act 2022's norms, which are provided together with these guidelines, include several examples specifying two conditions, few annotations including three conditions, and a couple including four conditions. Norms that require the template to specify more than four conditions appear to be extremely rare and they have not been found yet.

Multiple conditions usually took place when certain subsection specified when other subsections apply or do not apply. An example is provided by the five subsections of Article 18 of the Police, Crime, Sentencing and Courts Act 2022:

- (29) (1) Subsection (2) applies if the Secretary of State is satisfied that—
- (a) a specified authority has failed to discharge a duty imposed on it by section 8, 14(6), 15(3) or 17(4), or
  - (b) an educational authority, prison authority or youth custody authority has failed to discharge a duty imposed on it by section 15(3), (4) or (5)(b) or 17(4).
- (2) The Secretary of State may give directions to the authority for the purpose of securing compliance with the duty.
- (3) ...
- (4) ...
- (5) This section does not apply in relation to—
- (a) a provider of probation services if that provider is the Secretary of State,
  - (b) the governor of a prison, young offender institution or secure training centre, or
  - (c) the principal of a directly managed secure college as defined in paragraph 27 of Schedule 10 to the Criminal Justice and Courts Act 2015.

Subsection (1) conveys a condition for apply subsection (2) while subsection (5) conveys an exception to subsection (2), i.e., a condition for *not* applying it.

The whole Article 18 of the Police, Crime, Sentencing and Courts Act 2022 is therefore annotated as follows:

(30) **IT IS PERMITTED**

**FOR** The Secretary of State

**TO** give directions to the authority for the purpose of securing compliance with the duty

**WHEN/IF/WHERE** the Secretary of State is satisfied that—

- (a) a specified authority has failed to discharge a duty imposed on it by section 8, 14(6), 15(3) or 17(4), or
- (b) an educational authority, prison authority or youth custody authority has failed to discharge a duty imposed on it by section 15(3), (4) or (5)(b) or 17(4).

**UNLESS** this is done in relation to—

- (a) a provider of probation services if that provider is the Secretary of State,
- (b) the governor of a prison, young offender institution or secure training centre, or
- (c) the principal of a directly managed secure college as defined in paragraph 27 of Schedule 10 to the Criminal Justice and Courts Act 2015.

Another example is provided by first two subsections of Article 15 of the Equality Act 2010:

- (31) (1) A person (A) discriminates against a disabled person (B) if—
- (a) A treats B unfavourably because of something arising in consequence of B's disability, and
  - (b) A cannot show that the treatment is a proportionate means of achieving a legitimate aim.
- (2) Subsection (1) does not apply if A shows that A did not know, and could not reasonably have been expected to know, that B had the disability.

Contrary to the previous example, in (31) also paragraph (b) of subsection (1) conveys an exception to the main prohibition (conveyed by paragraph (a) of subsection (1)). Therefore, these two subsections have been annotated as follows:

(32) **IT IS PROHIBITED**

**FOR** A person (A)

**TO** discriminate a disabled person (B) by treating B unfavourably because of something arising in consequence of B's disability

**UNLESS** A can show that the treatment is a proportionate means of achieving a legitimate aim

**UNLESS** A shows that A did not know, and could not reasonably have been expected to know, that B had the disability.

An example of template with three conditions is the one associated with subsections (1), (2), (4), and (6) of Article 106 of the Equality Act 2010:

- (33) (1) This section applies to an association which is a registered political party.
- (2) If the party had candidates at a relevant election, the party must, in accordance with regulations, publish information relating to protected characteristics of persons who come within a description prescribed in the regulations in accordance with subsection (3).
- (3) ...
- (4) The duty imposed by subsection (2) applies only in so far as it is possible to publish information in a manner that ensures that no person to whom the information relates can be identified from that information.
- (5) ...
- (6) This section does not apply to the following protected characteristics—
- (a) marriage and civil partnership;
  - (b) pregnancy and maternity.

Annotated as follows:

(34) **IT IS OBLIGATORY**

**FOR** a registered political party

**TO**, in accordance with regulations, publish information relating to protected characteristics of persons who come within a description prescribed in the regulations in accordance with subsection (3)

**WHEN/IF/WHERE** the party had candidates at a relevant election

**ONLY IF** it is possible to publish information in a manner that ensures that no person to whom the information relates can be identified from that information

**UNLESS** the published information concern the following protected characteristics—

- (a) marriage and civil partnership;
- (b) pregnancy and maternity.

Very rarely, the conditions for the obligations, permissions, and prohibitions might be even found in subsequent articles. For instance, consider the first four subsections of Article 41 of the Police, Crime, Sentencing and Courts Act 2022:

- (35) (1) An authorised person may extract information stored on an electronic device from that device if—
  - (a) a person who was a user of the electronic device has died, and
  - (b) the person was a user of the device immediately before their death
- (2) The power in subsection (1) may be exercised only for the purposes of—
  - (a) an investigation into the person’s death under Chapter 1 of Part 1 of the Coroners and Justice Act 2009,
  - (b) an inquest into the person’s death under the Coroners Act (Northern Ireland) 1959, or
  - (c) an investigation into the person’s death by the Lord Advocate
- (3) ...
- (4) An authorised person may exercise the power in subsection (1) only if—
  - (a) the authorised person reasonably believes that information stored on the electronic device is relevant to a purpose within subsection (2), and
  - (b) the authorised person is satisfied that exercise of the power is necessary and proportionate to achieve that purpose.

Later in the act, subsection (3) of Article 44 specifies that:

- (36) (4) The power in section 41(1) may be exercised only by a person listed in Part 1 of Schedule 3.

Altogether, the text in (35) and (36) has been tagged as in (37). The annotation in (37) is one of the very few examples that includes four conditions.

(37) **IT IS PERMITTED**

**FOR** An authorised person

**TO** extract information stored on an electronic device from that device

**WHEN/IF/WHERE**

- (a) a person who was a user of the electronic device has died, and
- (b) the person was a user of the device immediately before their death

**ONLY IF** the purpose is

- (a) an investigation into the person's death under Chapter 1 of Part 1 of the Coroners and Justice Act 2009,
- (b) an inquest into the person's death under the Coroners Act (Northern Ireland) 1959, or
- (c) an investigation into the person's death by the Lord Advocate

**ONLY IF**

- (a) the authorised person reasonably believes that information stored on the electronic device is relevant to a purpose within subsection (2), and
- (b) the authorised person is satisfied that exercise of the power is necessary and proportionate to achieve that purpose

**ONLY IF** the person is a person listed in Part 1 of Schedule 3.