## Appendix E: Word clouds

### E1 – Method for generating word clouds

Word clouds can be an effective way to visualize and contrast views articulated by many participants, as well as views less commonly put forward. Using the documents containing the written submission of essential requirements and the reason for each requirement from all participants as a starting point, two team members independently and manually coded the input to generate word clouds in English and French. This method was necessary to ensure that: i) ideas that had the same essence were assigned the same code (e.g., “transparency” and “transparent” were coded both as “transparency” because more participants used the word “transparency” than “transparent”); and ii) ideas that contained similar words were coded differently to preserve their meaning (e.g., in certain cases, “secure” and “security” were coded separately to reflect the participants’ original submission, based on context). Differences in coding between the two team members were minor and resolved through discussion.

Word clouds were created separately in English and French (i.e., one word cloud was not a translation of the other) using Poll Everywhere. The word cloud function of Poll Everywhere was selected because it presented words as entered and allowed multi-word codes to be kept together. Two word clouds (in each language) were generated: one that included all codes for participants' requirements, and one that included codes representing minority views, meaning codes that were present only one or two times.

### E2 – Words and phrases for all requirements seen as essential for health data social licence in participants’ own words **This image shows words and phrases for all requirements seen as essential for health data social licence in participants' own words. As an alternative to this visual presentation which uses larger fonts for the words and phrases expressed most frequently, section E4 presents the frequency of each coded words and phrases.**

### E3 – Words and phrases for all requirements seen as essential for health data social licence expressed one or two times in participants’ own words

This images shows words and phrases for all requirements seen as essential for health data social licence expressed one or two times in participants' own words. 

As an alternative to this visual presentation which uses larger fonts for the words and phrases expressed most frequently, section E4 presents the frequency of each coded words and phrases.

### E4 – Counts of words and phrases used to construct word clouds

| **Code word or phrase** | **Number of times expressed in participant’s requirements** |
| --- | --- |
| Health data | Participants referred to “health data” hundreds of times. To ensure that the words “health data” were not so large that other words and phrases were too small to be visible, a count of 37 was used for “health data” when constructing the word cloud. |
| Use | 26 |
| Patients | 21 |
| Benefits | 19 |
| Share | 18 |
| Access | 17 |
| Trust | 15 |
| Security | 14 |
| Research | 13 |
| Transparency | 12 |
| Public | 11 |
| Collect | 11 |
| Control | 10 |
| Confidentiality | 8 |
| Personal | 8 |
| Individuals | 8 |
| Reuse | 8 |
| Accountable | 7 |
| Commercial | 7 |
| Protect | 7 |
| Privacy | 7 |
| Consent | 7 |
| Discrimination | 6 |
| Misuse | 5 |
| Confidence | 5 |
| Regulation | 5 |
| Safeguard | 5 |
| Correct errors | 4 |
| Decision | 4 |
| Government | 4 |
| Patient access to their own information | 4 |
| Limits | 4 |
| Independent oversight | 4 |
| Canadian | 4 |
| Penalties | 3 |
| Compliance | 3 |
| Language | 3 |
| Communication | 3 |
| Stakeholders | 3 |
| Data | 3 |
| De-identify | 3 |
| Involve | 3 |
| Opt in | 3 |
| Best interest | 3 |
| Storing | 2 |
| Governance | 2 |
| Automated collection | 2 |
| Provincial system | 2 |
| Anonymized | 2 |
| Fraud | 2 |
| Inform | 2 |
| Healthcare | 2 |
| Third party | 2 |
| Communities | 2 |
| Improve | 2 |
| Identifiable | 2 |
| Disseminate | 2 |
| Applications | 2 |
| Insurance companies | 2 |
| Authorized | 2 |
| Safety | 2 |
| Subpopulations | 2 |
| Informed consent | 2 |
| Lived experience | 2 |
| Useful | 2 |
| Quality | 2 |
| Right | 2 |
| Publish | 2 |
| Breach | 2 |
| Opt out | 2 |
| Voluntary | 2 |
| Disability | 2 |
| Valuable | 2 |
| Integrity | 2 |
| Beneficiaries | 2 |
| Inaccurate | 2 |
| Intent | 2 |
| Healthcare system | 2 |
| Plain language | 2 |
| Understandable language | 2 |
| Health system | 2 |
| Genetic data | 2 |
| Sale | 2 |
| Private sector | 2 |
| Indigenous Peoples | 2 |
| Results | 2 |