

S1: Results of frequent bigrams

Table S1 Results of top 20 frequent bigrams

Bigram	Frequency	Number of agreements
third, party	1073	118
service, provider	560	98
contact, us	526	114
may, use	383	98
email, address	348	107
use, information	345	93
collect, use	329	102
information, may	316	91
use, cookie	265	82
IP, address	222	84
personal, identifiable	214	67
information, use	211	98
may, collect	203	71
identifiable, information	201	60
disclose, personal	192	65
terms, conditions	191	62
share, personal	185	59
google, analytics	178	55
help, us	162	64
may, include	157	76

S2: Results of frequent trigrams

Table S2 Results of frequent trigrams

Trigrams	Frequency	Number of agreements
personal, information, collect	147	51
disclose, personal, information	145	52
please, contact, us	128	68
inform, third, party	115	63
share, personal, information	111	37
use, personal, data	96	19
disclosure, personal, information	84	41
third, party, service	76	34
access, personal, information	75	37
protect, personal, information	72	40
personal, information, provide	72	42

name, email, address	69	47
provide, personal, information	67	49
process, personal, information	64	25
collect, personal, data	63	26
collect, use, disclosure	61	32
information, collect, use	60	46
use, disclosure, personal	59	32
change, privacy, policy	58	38
collect, use, disclose	57	45
share, personal, data	55	18
use, disclose, person	54	29
delete, person, inform	46	23

Table S3 Results of trigrams with 'may'

personal, information, may	102	50
may, use, information	67	37
may, also, use	60	46
may, also, collect	47	34
may, share, personal	47	28
third, party, may	46	35
may, disclose, personal	44	29
may, share, information	39	18
information, may, use	38	29
may, use, personal	36	28
may, use, cookie	26	21
may, collect, use	25	19
may, contact, us	24	20
may, provide, us	23	16

S3: Keyword Analysis for Predefined Questions

Question: After I upload data to the service providers servers, will it be possible to retrieve my original complete dataset in an original or equivalent format?

Table S4 Results of keyword analysis

Keyword	Frequency	Number of farm data agreements
access	1263	122
request	1128	121
control	591	103
available	359	91
view	207	117
download	153	68

electronic	145	70
digital	129	34
copy	109	63
written	99	38
original	80	44
portable	39	26

Observation: It is noted that access is the most frequent word (1263 times) and is mentioned in 86.52% of the policies followed by request (1128 times) in 85.81% of the policies. Control is used in around 73% of the policies and View is used in 82.87% of the policies. Download, electronic copy are used 153 times and 109 times respectively and was covered in around 46% of the policy.

Sample (DJI Camera):

You may contact us to have your information blocked, deleted or erased, or to opt out from profiling where your prior usage of DJI Products and Services is processed to provide you with customized or personalized products, services or advertisements, or to obtain copies of certain of your information in a **portable** format.

Sample (AGvisorPRO):

A printed version of these Terms of Service and of any notice given in **electronic** form will be admissible in judicial or administrative proceedings based upon or relating to these Terms of Service to the same extent and subject to the same conditions as other business documents and records originally generated and maintained in printed form.

Question: How are farm data protected in the case of cross-border data transfers? What policies apply to farm data?

Table S5 Results of keyword analysis

Keyword	Frequency	Number of farm data agreements
information/inform	7671	127
protect	921	119
transfer	463	97
country	259	68
european	138	43
jurisdiction	125	53
EEA	63	25
economic	46	30
crossborder	5	5
overseas	3	3

Observation: Transfer is the most frequent keyword used in 68 %of the policies. Country is used in 48% of the policies. European economic (EEA), jurisdiction are other relevant keywords which are used. crossborder, overseas are the least used keywords related to this query and is used in less than 3% of the policies.

Sample (Telus/AgIntegrated):

Cross-border Transfer of Information. As stated above, we generally maintain servers and systems in the United States and Canada hosted by service providers. We also may subcontract the processing of your information to, or otherwise share your personal information with, other parties in the United States, Canada, or countries other than your country of residence. As a result, where the personal information that we collect through or in connection with the Sites, Apps, or Services, or that we otherwise collect, is **transferred** to and processed in Canada, the United States or anywhere else outside the **European Economic Area (EEA)** or United Kingdom for the purposes described above, we will take steps to ensure that the information receives the same level of protection as if it remained within the **EEA** or United Kingdom, including entering into data transfer agreements using the EU Commission approved Standard Contractual Clauses or other mechanisms that provide an adequate level of protection for your personal information. You may have a right to details of the mechanisms under which your data is transferred outside the **EEA** or United Kingdom.

Sample (John Deere):**Transfer of Information & Data to Other Countries**

Deere & Company is headquartered in the United States of America and its controlled affiliates are located in various countries throughout the world. Personal Information and Machine Data may be accessed by or **transferred** to the United States or to our controlled affiliates and data processors elsewhere in the world for the Purposes described in this Privacy Statement. When we access or **transfer** Personal Information and Machine Data, we do so in compliance with applicable law. Additionally, we protect the privacy and security of Personal Information and Machine Data, regardless of where it is processed or stored.

Question: Who can we contact in the company if we have questions about farm data?

Table S6 Results of keyword analysis

Keyword	Frequency	Number of farm data agreements
contact	1313	127
request	1128	121
customer	931	112
help	492	99
question	266	102
directed	224	111
concern	153	69
inquiry	86	52

Contact is used 1313 times in 90% of the policies. *Request* is the second most frequent word for this query and is repeated 1128 times in 85% of the policies. *Help* is used in 70% of the policies. *Question* is also used frequently 266 times in 72% of the policies whereas *inquiry* is used 86 times only in 36% of the policies.

Observation:

Under GDPR and some codes of conduct, it is important to mention the contact information of the companies or service providers. Most of the agreements have a “contact us” section with all the

company's details where the user can contact them. Below are some sample texts that discuss contact information of service providers:

Sample text 1 related to Q7:

Please **contact us** if you have any **questions** or comments about our privacy practices or this Privacy Statement. You can reach us online at ...

Sample text 2 related to Q7:

Contact Us If you have any additional **questions** or concerns about these Terms, please **contact us** at:

Question: Do the service provider agreements establish how long my original datasets will be retained?

Table S7 Results of keyword analysis

Keyword	Frequency	Number of farm data agreements
time	948	122
store	585	109
delete	489	94
record	255	79
period	218	85
retain	199	73
keep	198	84
terminate	148	72
profile	133	44
retention	111	59
close	87	47
backup	41	26
destroy	24	20

From the results, it is noted that *time* is the most frequent keyword related to this query and is used 948 times in 86.52% of policies. The other frequent words related to this question are *store*, *record*, *retain*, *keep*, *retention*, and *backup*. *Store* is used 585 times in 77% of the policies. *Record*, *retain* and *keep* are used in approximately 56% of the policies. *Backup* is used less frequently (41 times in 18% of the policies only) as compared to other keywords and *retention* is used 111 times in 41% of the policies. *Terminate* was used in 51% of the policies.

Observation:

The analysis shows that although retention-related practices are mentioned in the agreements, the statements are written in an inconsistent and unclear manner. For instance, the agreements use words such as “as long as” or “reasonable” to refer to the period ATPs keep the data. Only limited agreements provide information regarding the criteria or conditions determining the retention period. Most of the agreements give a generalized statement without giving enough details of the retention practices.

Sample text 1 related to Q6:

“Our company” **retains** personal information for only **as long as** is reasonable to fulfill the purposes for which the information was collected, or for legal or business purposes.

Sample text 2 related to Q6:

“The Company” acquires and **retains** only the personal data that is required for effective operation of Raven or otherwise required by law. Access to such information is restricted internally to those with a recognized need to know.

Sample text 3 related to Q6:

We **retain** Personal Information for as long as needed or permitted in light of the purpose(s) for which it was obtained and consistent with applicable law.

The criteria used to determine our **retention** periods include: The length of time we have an ongoing relationship with you and provide the Services to you (for example, for as long as you have an account with us or keep using the Services);

Whether there is a legal obligation to which we are subject (for example, certain laws require us to keep records of your transactions for a certain period of time before we can delete them); or

Whether **retention** is advisable in light of our legal position (such as in regard to applicable statutes of limitations, litigation or regulatory investigations).