FDSN Framework Proposal for a
Metadata Standard for Legacy Seismic Data

Proposal Phase – Community Consensus Phase Type B: new FDSN standard
Submitted by Tim Ahern, IRIS Emeritus and Lorraine Hwang, UC Davis

Background: In September 2019, a workshop was held in Albuquerque, New Mexico where 29 interested parties discussed many topics related to legacy seismic data. A significant outcome of this workshop was the identification of important metadata needed to assist in the discovery and use of legacy seismic data. The workshop report can be found at https://eartharxiv.org/repository/view/418/. Workshop participants began with a set of metadata elements that have been used in a variety of data rescue projects around the world. Participants reviewed these elements and gave their opinion as to whether each element should be included in any metadata standard developed for legacy data. Workshop participants also suggested modifications to proposed elements and many made suggestions for additional elements that should be considered.

In March 2021, an expanded survey, including all input from the Albuquerque workshop, was sent to several email lists and/or newsletters in the international seismological community including FDSN, IASPEI, IRIS, SCEC, and AGU. The metadata effort was also announced in presentations at the 2021 Seismological Society of America meeting by Lorraine Hwang.

What this project is not about. This portion of the legacy data effort only focusses on identifying the key metadata to aid in the discovery and use of digitized objects such as scanned images of paper records. It does not try to address the actual process and funding of the digitization of analog data which can entail scanning paper records, transcribing data from analog tapes, or even vectorizing scanned images of paper records. Having comprehensive metadata can help others in data transcription and vectorizing data, but those activities are not part of this effort. Existing and new efforts will address the digitization effort needed separately. This is the first attempt we are aware of to try and standardize the metadata that will be needed to support a distributed, federated system of digital legacy seismic data objects generated from analog recordings. This project, and perhaps future efforts, will enable enhanced search and discovery, improved usability and enable access to all collections that adopt it. We do not believe this level of international coordination has been attempted before.

Metadata in Context: Survey takers were asked to review the 56 metadata elements shown in Appendix A, and indicate whether specific metadata elements should be **Required, Recommended, Optional, or Not Included** in a legacy metadata standard. The recommendations matter since:

- Required metadata MUST be provided by all centers wishing to meet FDSN standards and this will require effort. FDSN Data Center tools must allow search capability for all required metadata and return all required metadata resulting from user queries.

- Recommended metadata SHOULD be included if available and will improve discovery and document key metadata for use of the data by third parties. Future tools should support these elements in search and discovery and for usability of the data even though they are not required and may not be available for some collections.
- Optional metadata MAY be provided but collections making their data available and centers that adopt the FDSN standard and manage legacy data within the standard will at their discretion include or not include optional metadata. All centers managing optional metadata should have systems that capture and make available optional data.
- Metadata elements that are determined NOT USEFUL will be omitted in the resulting standard.

Level of Effort

The final product of the metadata for legacy seismic data effort is the development of an FDSN standard set of metadata and associated documentation. Once that is in place, it would be possible to further enhance the management of legacy data when and if there are funds available to support this effort.

Reducing the amount of metadata eases both the work of those at collection centers creating the metadata and data centers that manage metadata and corresponding legacy data. FDSN working groups should balance the amount of effort needed both at the collection centers and FDSN data centers and still be able to support enhanced discovery systems and support full usability of the data.

We envision a distributed effort where owners of digital collections provide metadata in the new FDSN standard format as well as perform the actual digitization of the analog data (we focus on scanning and not digitization of traces in this effort). We assume that traces vectorized from the original scanned images would be put into FDSN SEED formats. We believe that most owners of collections would opt to maintain the metadata and scanned images locally, but it would be their discretion if they prefer to have another center manage the digital data remotely.

Remember that curators of legacy data collections will be responsible for providing all Required metadata elements and the available Recommended metadata elements. The amount of effort will be related to the number of elements so that should be considered as FDSN working groups categorize the various elements. Requiring too many metadata elements will discourage those with collections from participating as well as increasing the burden for data centers managing the metadata. The development of software tools and services could also be impacted. Not having complete REQUIRED metadata will result in such datasets not being included an FDSN system.

The Working Group might consider excluding metadata that is available through other means for instance the size of the object being managed can be determined directly and therefore not included as Required metadata. Similarly, the file type may be

identified in the file name rather than something provided by the collection manager as metadata.

The objective of the effort should be to manage the level of effort required for data collections to participate but also ensure that enough metadata is required to enable future developments related to making the data FAIR.

Findable, Accessible, Interoperable and Reusable (FAIR) Data
The goal is to make initial progress toward making legacy seismic data FAIR. However, it is important to keep the required effort for those with existing collections reasonable and affordable.

- Findable. One of the most important outcomes of this effort is having metadata that would enable discovery of related digital objects across a federated system of legacy seismic data centers. By standardizing and making metadata such as the time and coordinates of the original analog data available, intelligent searches would be enabled. By federating the system, all data managed at multiple data centers would also be findable even to researchers unaware of the existence of some data collections.

- Accessible. We foresee the development of an FDSN system to manage objects containing information in a digital format such as scanned images of paper records, or transcriptions of data from analog tapes. Including a DOI as an element of the metadata should make the data accessible in that the digital objects of interest could be found and transferred to a user's system using very simple file transfer systems that exist. Later systems, perhaps based on web-services that are more powerful could support added features such as format conversion, file compression, image sub setting, etc. The envisioned system will make legacy seismic data more accessible but future efforts could significantly improve accessibility.

- Interoperable. Traditional digital seismic data interoperability was greatly enhanced through service-based approaches within the FDSN in the past. Some of these services were able to improve interoperability for instance by supporting easy to understand formats that are not esoteric domain-based formats. The metadata in this project would use international standards when they exist to identify metadata such as the time and coordinates of the legacy data, again improving interoperability. Capturing required, recommended, and other identified metadata inherently make these data more interoperable within and outside the seismological domain.

- Reusable. A system that manages the metadata as well as the digital representations of legacy seismic data has the potential for allowing others to reuse the data and reproduce the work using the same metadata as well as the

same digital object. The data in such a system would significantly improve the reusability of the data.

The Survey

There are 7 broad categories in which the 56 survey metadata elements lie in addition to 19 new elements suggested by those completing the latest survey.

Metadata Elements in Survey by Metadata Category

	Metadata Category	Elements in Survey	Suggested by Survey Respondents
1	Time of Data	3	1
2	Station or Channel	12	4
3	Sensor	8	4
4	Recording System	4	2
5	Helical Drum Recorder (TESEO)	6	6
6	Image File	20	1
7	Miscellaneous	3	1
8	New Category Suggestions*		0
	TOTAL NUMBER OF ELEMENTS	56	19

*While some users made suggestions for new metadata categories, all suggested new categories could be easily mapped into the original 7 categories.

Appendix A is included and identifies the Category, metadata element name, short description and possible units. **Appendix B** shows the Final Survey Results for the 56 metadata elements by Metadata Category Type. Results are presented both in graphical and numerical format. **Appendix C** contains a summary of New Metadata Elements Recommended in Latest Survey including the category, element name, a short description and units when applicable.

Proposed Procedure – follow the FDSN Framework

Community Consensus Phase

1. Form a Proposal Review Team
2. Proposal Review Team should be cognizant of the survey results but determine a final list of Required, Recommended, Optional, and Not Included parameters keeping in mind the ramifications of having too many or too few parameters.
3. WG II should vote on the Review team recommendations
4. If approved, the WGII recommendation should be sent to the Steering Committee for formal adoption by the FDSN

Evaluation and Adoption Phase

5. If the decision is made to accept the standard, then
    a. WGII should begin the formal definition of an XML Schema to capture all accepted elements categorized (e.g., Required, Recommended, Optional)

b. FDSN WGII could identify tools and systems to create, manage, and disseminate the information in the standard.
c. Service-based systems to access data across a Federated system of Legacy Data Collections/Centers should be discussed and perhaps developed.

Appendix A – Metadata elements in the Final Survey

Appendix A shows the Metadata Category, the Element Name, a short description and units provided to those that took the final survey. They are provided here to help the Working Group in their efforts. The green elements identify metadata elements that were either new or modified by participants in the Albuquerque workshop.

	56		All Entries are ASCII
	Metadata Element	Metadata Element Description	Units or Format
Time of Data	3		
	Start Time	the time of the first sample in the image	YYYY-MM-DD-THH:MM:SS.FFFFZ format
	End Time	the time of the last sample in the image	YYYY-MM-DD-THH:MM:SS.FFFFZ format
	Time Correction	any time correction applied to the data	Seconds (nn.nnn)
Station Channel Details	12		
	Latitude	latitude using WGS84 datum	SEED format convention
	Longitude	longitude using WGS84 datum	SEED format convention
	Elevation	Elevation above (+) or below (-) sea level	real in meters
	Depth of sensor below ground surface	depth below ground surface at specified longitude and latitude	real in meters
	Network Name	network to which the station belongs (e.g. WWSSN,GSN, EREBUS)	text
	FDSN Network Code	FDSN network code- Earliest FDSN Code in use for the station (use SS if not associated with a network)	text
	Site Name	site name (e.g. Albuquerque, New Mexico, USA)	text
	IR Station Code	station's code in the International Registry (ISC)	text
	Channel/component	channel code as in SEED format	text as in SEED Manual Appendix A
	Open Date	date when station was opened	YYYY-MM-DD-THH:MM:SS.FFFFZ format
	Close Date	if closed, Date when station was closed. Leave empty if still	YYYY-MM-DD-THH:MM:SS.FFFFZ format
	FDSN TimeSeries Identifier	Proposed new FDSN Time series identifier	
Sensor	8		
	Type of sensor	type of sensing instrument (e.g. Streckheisn STS-2, Ewing,	text
	Sensor serial number	manufacturer's serial number of seismometer if known	text
	Galvo Free period	the free period of the instrument	real in seconds (nn.nnn)
	Galvo Damping constant	the instrument's damping constant	real dimensionless (0.nnn)
	Horizontal 1 dip/azimuth	the dip/azimuth of the first horizontal	SEED convention
	Horizontal 2 dip/azimuth	the dip/azimuth of the second horizontal	SEED convention
	Vertical dip/azimuth	the dip/azimuth of the vertical channel	SEED convention
	Nature of instrument	Mechanical (e.g., Wiechert) or electromagnetic (e.g., Golitsyn).	Mechanical (enter T, V, epsilon) or electromechanical (enter Tp, hp, Tg, hg, mu, Vmax)
Recording	4	Many of these are parameters used in Teseo. Some of these	
	Type of recording system	type of recording system (e.g. Teledyne helicorder)	text
	Recording system serial number	manufacturers serial number if known	text
	Scale/gain/amplification	scale or gain factor (scaler)	real dimensionless
	Period of scale/gain	period at which the gain is valid	real in seconds
TESEO	6		
	Paper speed	paper speed (linear velocity of paper)	real in mm/min
	R	length of the writing arm, from its rotating axis to the tip of	real in centimeters
	r	radius of the drive cylinder bearing the smoked paper	real in centimeters
	a	distance from the rotating arm axis to the driving cylinder	real in centimeters
	b	shift of the arm axis to the base line on the smoked paper	real in millimeters
	d	length of 1 minute on paper	real in millimeters
Image file details	20		
	DOI of scanned Image	Ienter the DOI if one has been assigned	DOI
	Date of Scanning	the data the image was scanned	YYYY-MM-DD-THH:MM:SS.FFFFZ
	Resolution	the resolution of the scanned image	pixels per inch
	Vertical pixels	the number of pixels in the vertical dimension	number of pixels
	horizontal pixels	the number of pixels in the horizontal dimension	number of pixels
	Image format	image file type	e.g. heic,jpeg, jpeg-2000, openEXR, pdf, png, tiff
	image size	the total size of the image in bytes	integer
	Analog image length	length of the original document	real in meters
	Analog image width	width of the original document	real in meters
	Color depth	the color depth of the scanner if applicable	integer
	Phase Markings present	Indicate if phase notations were placed in the image	True or False
	Associated Bulletin	Earthquake phases are present on image. Phases were reported to a bulletin or otherwise publlished.	if True provide Bulletin Name, URL, or DOI of the bulletin
	Occlusions	Indicate true if tears or other flaws obscure trace data	True or False
	Earthquake signal	Indicate true if an earthquake signal is present	True/False
	Timemark Format	positive real to indicate length of vertically offset	real numbers (pixels?)
	Polarity of recording	Ground motion up = up on paper or down on paper	up or down
	Original recording type	Photographic paper, drum recordings (smoke, hot stylus, ink)	text
	Location of original record	country, state or province, city, institution, room of original analog document when scanned	text

	Vectorized_trace	The trace in this image has been vecotrized.	URL or DOI of the vectorized trace
	Contact information of owner	contact information for the original owner of the data	text

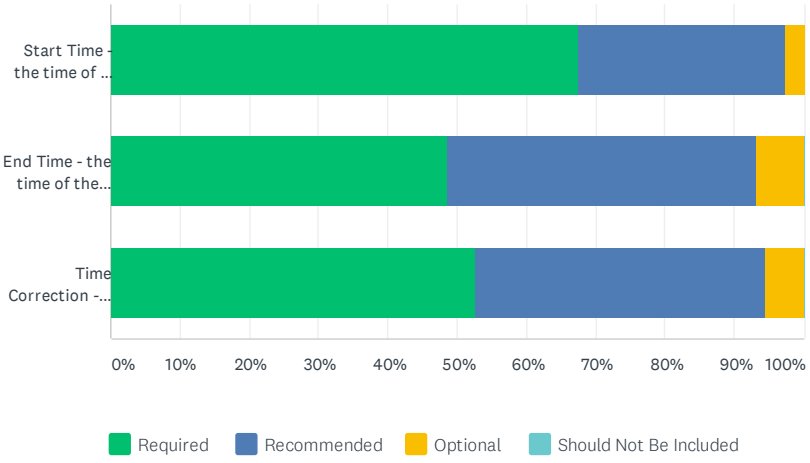
Additional Metadata	3		
	Notes and Comments	optional Notes and/or comments	text
	Source of information	Information about source of metadata entered - e.g., lat/lon adopted from a published source, or response assumed based on X information or publication.	freeform text
	Date of metadata creation	date and time when the metadata was created or last updated	YYYY-MM-DD-THH:MM:SS.FFFFZ

Appendix B – Final Survey Results for 56 metadata elements by Category Type

Appendix B shows preliminary results from the survey as of July 18, 2021.

Q7 Metadata related to the time of the data captured

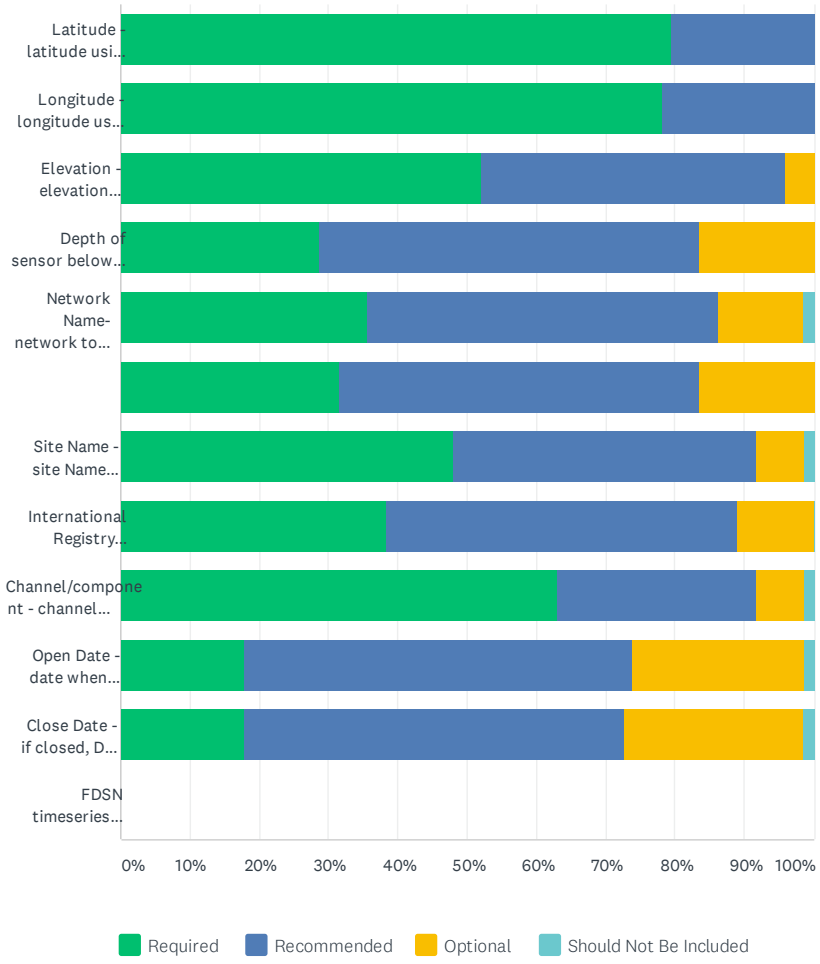
Answered: 74 Skipped: 1



	REQUIRED	RECOMMENDED	OPTIONAL	SHOULD NOT BE INCLUDED	TOTAL
Start Time - the time of the first sample in the image in YYYY-MM-DD-T00:00:00.0000Z format	67.57% 50	29.73% 22	2.70% 2	0.00% 0	74
End Time - the time of the last sample in the image in YYYY-MM-DD-T00:00:00.0000Z format	48.65% 36	44.59% 33	6.76% 5	0.00% 0	74
Time Correction - any time correction applied to the data in YYYY-MM-DD-T00:00:00.0000Z format	52.70% 39	41.89% 31	5.41% 4	0.00% 0	74

Q8 Station and Channel metadata

Answered: 73 Skipped: 2

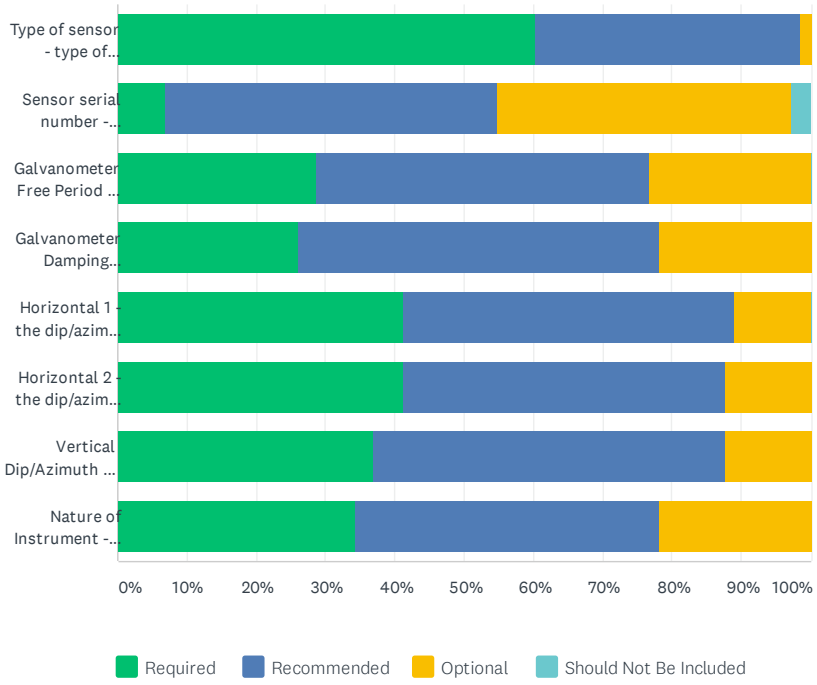


Securing Legacy Seismic Data to Enable Future Discoveries - Version 2 (Feb 2021)

	REQUIRED	RECOMMENDED	OPTIONAL	SHOULD NOT BE INCLUDED	TOTAL
Latitude - latitude using WGS84 datum - SEED format convention	79.45% 58	20.55% 15	0.00% 0	0.00% 0	73
Longitude - longitude using WGS84 datum - - SEED format convention	78.08% 57	21.92% 16	0.00% 0	0.00% 0	73
Elevation - elevation meters above (+) or below (-) sea level - real in meters	52.05% 38	43.84% 32	4.11% 3	0.00% 0	73
Depth of sensor below ground surface - depth below ground surface at specified longitude and latitude - real in meters	28.77% 21	54.79% 40	16.44% 12	0.00% 0	73
Network Name- network to which the station belongs (e.g., WWSSN,GSN, EREBUS) - text	35.62% 26	50.68% 37	12.33% 9	1.37% 1	73
FDSN Network Code - FDSN network code. Earliest FDSN Code in use for the station. (use SS if not associated with a network) - text	31.51% 23	52.05% 38	16.44% 12	0.00% 0	73
Site Name - site Name (e.g., Albuquerque, New Mexico, USA) - text	47.95% 35	43.84% 32	6.85% 5	1.37% 1	73
International Registry Station Code - station's Code in the International Registry (ISC) -text	38.36% 28	50.68% 37	10.96% 8	0.00% 0	73
Channel/component - channel code as in the SEED format - text as in SEED Manual Appendix A	63.01% 46	28.77% 21	6.85% 5	1.37% 1	73
Open Date - date when station was opened - YYYY-MM-DD-T00:00Z format	17.81% 13	56.16% 41	24.66% 18	1.37% 1	73
Close Date - if closed, Date when station was closed. Leave empty if still operating or not known - YYYY-MM-DD-T00:00Z format	17.81% 13	54.79% 40	26.03% 19	1.37% 1	73
FDSN timeseries identifier - https://iris-edu.github.io/xseed-specification/docs/xFDSNSourceIdentifiers-DRAFT20190520.pdf	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0

Q9 Sensor Metadata

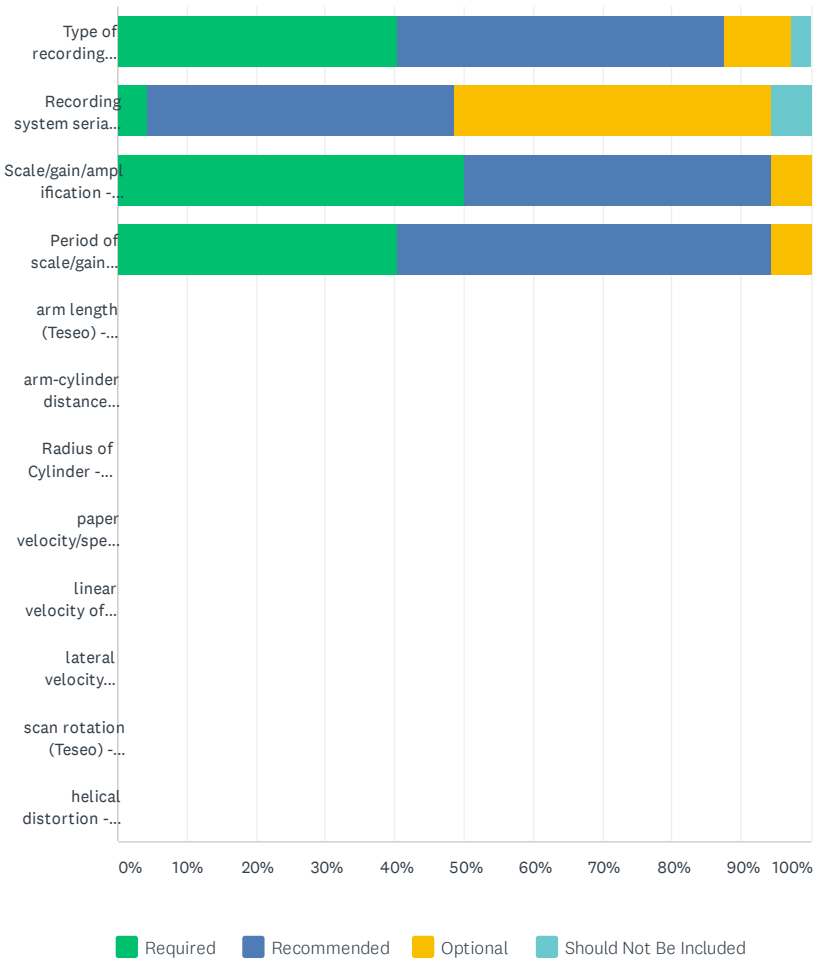
Answered: 73 Skipped: 2



	REQUIRED	RECOMMENDED	OPTIONAL	SHOULD NOT BE INCLUDED	TOTAL
Type of sensor - type of sensing instrument (e.g., Press Ewing, Benioff) - text	60.27% 44	38.36% 28	1.37% 1	0.00% 0	73
Sensor serial number - manufacturer's serial number of seismometer if known - text	6.85% 5	47.95% 35	42.47% 31	2.74% 2	73
Galvanometer Free Period - real in seconds nn.nnn	28.77% 21	47.95% 35	23.29% 17	0.00% 0	73
Galvanometer Damping Constant - the instrument's damping constant - real dimensionless 0.nnn	26.03% 19	52.05% 38	21.92% 16	0.00% 0	73
Horizontal 1 - the dip/azimuth of the first horizontal - real in SEED convention	41.10% 30	47.95% 35	10.96% 8	0.00% 0	73
Horizontal 2 - the dip/azimuth of the second horizontal - real in SEED convention	41.10% 30	46.58% 34	12.33% 9	0.00% 0	73
Vertical Dip/Azimuth - the dip/azimuth of the vertical channel - real SEED convention	36.99% 27	50.68% 37	12.33% 9	0.00% 0	73
Nature of Instrument - Mechanical (e.g., Wiechert) or Electromagnetic (e.g., Golitsyn) - Mechanical (enter T, V, epsilon) or electromechanical (enter Tp, hp, Tg, hg, mu, Vmax)	34.25% 25	43.84% 32	21.92% 16	0.00% 0	73

Q10 Recording System Metadata

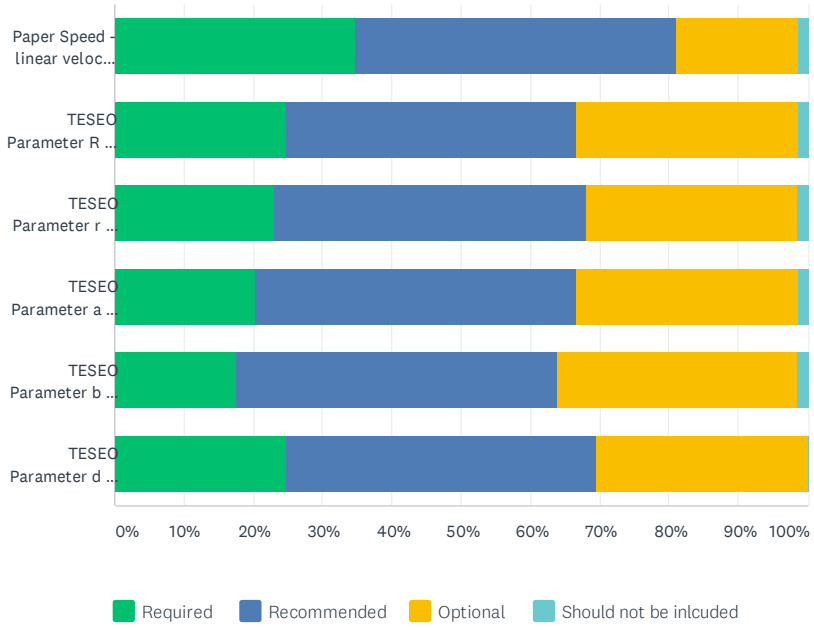
Answered: 72 Skipped: 3



	REQUIRED	RECOMMENDED	OPTIONAL	SHOULD NOT BE INCLUDED	TOTAL
Type of recording system - type of recording system (e.g., Teledyne helicorder) - text	40.28% 29	47.22% 34	9.72% 7	2.78% 2	72
Recording system serial number - manufacturers serial number if known - text	4.17% 3	44.44% 32	45.83% 33	5.56% 4	72
Scale/gain/amplification - scale or gain factor (scalar) - real dimensionless	50.00% 36	44.44% 32	5.56% 4	0.00% 0	72
Period of scale/gain -period at which the gain is valid - real in seconds	40.28% 29	54.17% 39	5.56% 4	0.00% 0	72
arm length (Teseo) - writing arm length in meters	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0
arm-cylinder distance (Teseo) - distance from the rotating arm axis to the driving cylinder axis in meters (can be estimated from other parameters)	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0
Radius of Cylinder - radius of the drive cylinder in meters	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0
paper velocity/speed (Teseo) - shift of the arm axis (meters) to base line (a "best fit" can be estimated)	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0
linear velocity of cylinder - meters/minute, meters/sec - linear velocity, can be derived	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0
lateral velocity (Teseo) - meters/hour - can be derived, not required	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0
scan rotation (Teseo) - misorientation of the paper on the scanner in degrees (positive means clockwise rotation). Not required	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0
helical distortion - units???	0.00% 0	0.00% 0	0.00% 0	0.00% 0	0

Q11 TESEO - Parameters for helicorder drum recorders

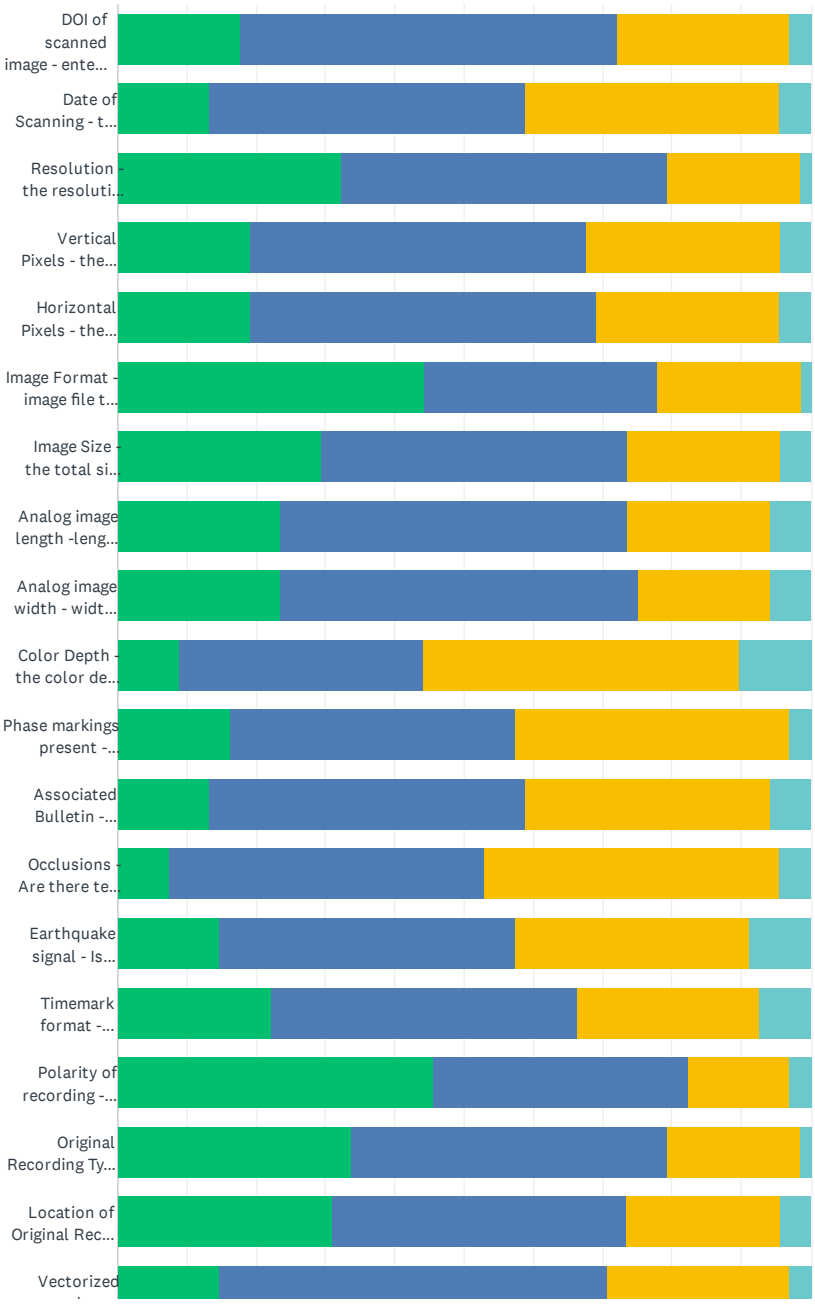
Answered: 69 Skipped: 6



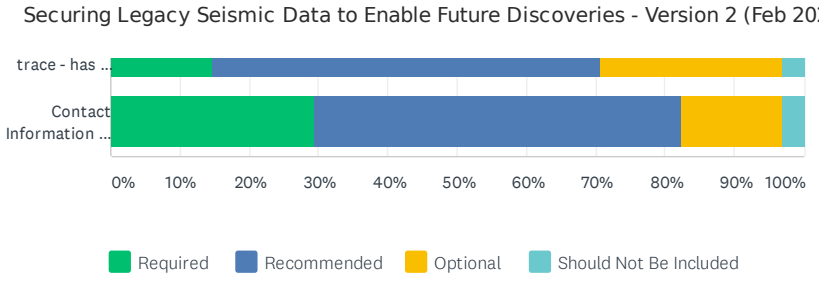
	REQUIRED	RECOMMENDED	OPTIONAL	SHOULD NOT BE INLCUDED	TOTAL
Paper Speed - linear velocity of paper - real in mm/min	34.78% 24	46.38% 32	17.39% 12	1.45% 1	69
TESEO Parameter R - length of the writing arm, from its rotating axis to the tip of the needle - real in cm	24.64% 17	42.03% 29	31.88% 22	1.45% 1	69
TESEO Parameter r - radius of the drive cylinder bearing the smoked paper - real in cm	23.19% 16	44.93% 31	30.43% 21	1.45% 1	69
TESEO Parameter a - distance from the rotating arm axis to the driving cylinder axis - real in cm	20.29% 14	46.38% 32	31.88% 22	1.45% 1	69
TESEO Parameter b - shift of the arm axis to the base line on the smoked paper - real in mm	17.39% 12	46.38% 32	34.78% 24	1.45% 1	69
TESEO Parameter d - length of 1 minute on paper -real in mm	24.64% 17	44.93% 31	30.43% 21	0.00% 0	69

Q12 Image File Details

Answered: 68 Skipped: 7



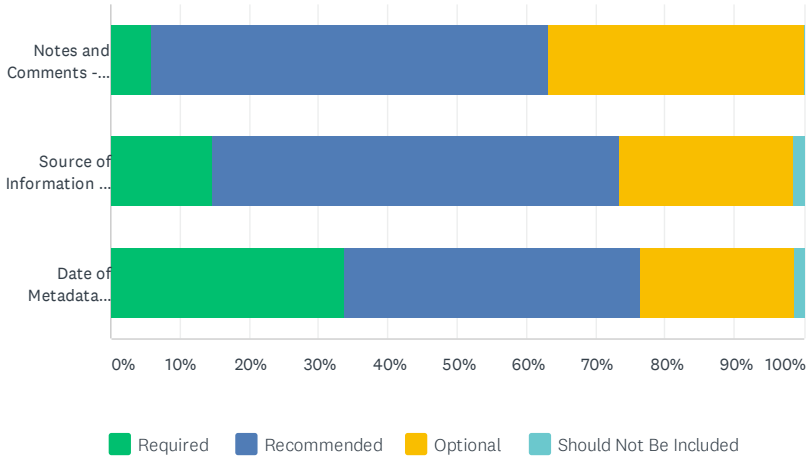
1 / 3


Securing Legacy Seismic Data to Enable Future Discoveries - Version 2 (Feb 2021)
trace - has ...
Contact Information ...
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Required
Recommended
Optional
Should Not Be Included

	REQUIRED	RECOMMENDED	OPTIONAL	SHOULD NOT BE INCLUDED	TOTAL
DOI of scanned image - enter the DOI of the scanned image	17.65% 12	54.41% 37	25.00% 17	2.94% 2	68
Date of Scanning - the date the image was scanned in YYYY-MM-DD-T00:00:00Z format	13.24% 9	45.59% 31	36.76% 25	4.41% 3	68
Resolution - the resolution of the scanned image - pixels per inch	32.35% 22	47.06% 32	19.12% 13	1.47% 1	68
Vertical Pixels - the number of pixels in the vertical dimension - integer	19.12% 13	48.53% 33	27.94% 19	4.41% 3	68
Horizontal Pixels - the number of pixels in the horizontal dimension - integer	19.12% 13	50.00% 34	26.47% 18	4.41% 3	68
Image Format - image file type - (e.g., jpeg, jpeg-2000, openEXR, pdf, png, tiff) - text	44.12% 30	33.82% 23	20.59% 14	1.47% 1	68
Image Size - the total size of the image in bytes - integer	29.41% 20	44.12% 30	22.06% 15	4.41% 3	68
Analog image length -length of the original document - real in meters	23.53% 16	50.00% 34	20.59% 14	5.88% 4	68
Analog image width - width of the original document - real in meters	23.53% 16	51.47% 35	19.12% 13	5.88% 4	68
Color Depth - the color depth scanned at, if applicable - integer	8.82% 6	35.29% 24	45.59% 31	10.29% 7	68
Phase markings present - Indicate if phase notations were placed in the image (T or F)	16.18% 11	41.18% 28	39.71% 27	2.94% 2	68
Associated Bulletin - Earthquake phases are present on image and phases were reported to a bulletin or otherwise published. (T or F). If T, provide Bulletin Name or PID of the Bulletin	13.24% 9	45.59% 31	35.29% 24	5.88% 4	68
Occlusions - Are there tears or other flaws that obscure trace data? (T or F)	7.35% 5	45.59% 31	42.65% 29	4.41% 3	68
Earthquake signal - Is there an earthquake signal present? (T or F)	14.71% 10	42.65% 29	33.82% 23	8.82% 6	68
Timemark format - positive real to indicate length of vertically offset timemarks, negative real to indicate length of gapped timemarks, null to indicate no timemarks - text	22.06% 15	44.12% 30	26.47% 18	7.35% 5	68
Polarity of recording - Ground motion up = up on paper or down = down on paper - text	45.59% 31	36.76% 25	14.71% 10	2.94% 2	68
Original Recording Type - e.g., Photographic paper vs. drum recordings (smoke, hot stylus, ink) - text	33.82% 23	45.59% 31	19.12% 13	1.47% 1	68
Location of Original Record - country, state or province, city, institution, room of original analog document when scanned - text	30.88% 21	42.65% 29	22.06% 15	4.41% 3	68
Vectorized trace - has the trace in this record been vectorized? PID of the vectorized trace	14.71% 10	55.88% 38	26.47% 18	2.94% 2	68
Contact Information of owner - contact information for the original owner of the data - text	29.41% 20	52.94% 36	14.71% 10	2.94% 2	68

Q13 Miscellaneous metadata

Answered: 68 Skipped: 7



	REQUIRED	RECOMMENDED	OPTIONAL	SHOULD NOT BE INCLUDED	TOTAL
Notes and Comments - optional Notes and/or comments - text	5.88% 4	57.35% 39	36.76% 25	0.00% 0	68
Source of Information - information about source of metadata entered - e.g., lat/lon adopted from a published source, or response assumed based on "X" information or publication. - freeform text	14.71% 10	58.82% 40	25.00% 17	1.47% 1	68
Date of Metadata Creation - date and time when the metadata was created or last updated - In YYYY-MM-DD-T00:00Z	33.82% 23	42.65% 29	22.06% 15	1.47% 1	68

Appendix C – New Metadata Elements Recommended in Latest Survey

Survey takers were offered the opportunity to suggest new categories of metadata (e.g., sensor, recording system) as well as specific metadata elements in each category. Appendix C shows their suggested new elements for consideration. There was a total of 19 additional elements suggested.

	Category	Metadata Element	Description	Possible answers	Units
1	Helical Drum Recorder	Shrinking Factor	records on photo have shrunk due to photo development or by time	Real number	none
2	Station and Channel	Calibration Pulse Present	Is there one (or several ???) calibration pulses present in the seismogram.	T/F	NA
3	Pertinent to Explosion or Earthquake	Nuclear test explosion signal	Characterization of the source	Scan	none
4	Sensor	Instrument Response	Provide either the analytic transfer function or a poles and zeros file, for instrumental response correction	Text	units of amplitude after correction and time-series type (displacement / velocity)
5	Sensor	Channel Transfer function	Left blank	Poles zeroes/gain	m/s to mm on recording
6	MMI Felt Report	seismogram associated with MMI report collection or magnitude	mapped MMI response (topographic) & reports	Y/N , text, URL address	MMI Scale
7	Time of Data	Time Zone	Which time zone the reported times and those time labels on the seismograms are in? If reported in local time, please indicate the time zone relative to GMT+0.	e.g., GMT+5, text	none
8	Left Blank	Left-to-right ordering of long seismograms scanned in parts	I've received images covering parts of seismograms labelled by L/R, a/aa/aaa, or a/b/c, etc. We can define a format like L1, L2, L3, etc.	e.g., L1, L2, L3 etc. to indicate 1,2,3 parts from the left, text	None

9	Image Files	Overlaps between seismograms scanned in parts	Describe an indicator of overlap between partial scans. Sometimes it is obvious when parts of the title plate appear in all scan parts, but other times it can be tricky for the user to spot the correct overlaps. I've received before really carefully labeled red circles on the scanned images, but the exact indicator can vary.	Text	none
10	Image Files	Was the image scanned in one piece?	Sometimes the seismograms are longer than the available scanner and requires scanning in parts. Mosaicing the scanned parts can be tricky as it can require careful image rotation and the soft paper may not lay completely flat on the scanner. The data provider can opt to mosaic themselves if comfortable, but it is good to indicate that clearly.	T/F	none
11	Additional Metadata	Notes on record interruptions	Sometimes a record was clearly interrupted due to needle dislocation or human pushing the drum one way to allow space for large amplitude surface waves to be recorded. These features are sometimes recorded in the station logbooks.	Text	none
12	Time of Data	Sampling Rate	Left blank	Real	ms.
13	Station and Channel	Exploration seismic section info	A description of the location, geometry, and history of an exploration seismic section, with all known details including scanned images of header blocks on plots.	Text and images	None
14	Image Additional Category	Storage conditions	General description of the conditions under which the original data were stored	Text	none
15	Recording System	Instrument Response	As complete a description of the instrument response as possible	image, frequency-amplitude-phase,	depends on response type (but specify whether response

				RESP file, etc.	corrects to acceleration, velocity, or displacement)
16	Sensor	daq	'daq' column is the one-sided full-scale input of the datalogger (rel#23), bit-weight, one-sided full-scale assuming 24bit	20.0 V, 0.3 uV, 2.5V	V, uV, V
17	Sensor	sensor	sensor' column is the sensor sensitivity (rel#42), one- sided sensor max output (rel#44)	1.25 v/g, 2.00 g	v/g, g
18	Station and Channel	Address	Identifies the resource on the network (Current 2 digit letter Station and Channel codes can't scale to provide enough unique addresses to unique identify growing number of number of sensors)	similar to a URI	Any sequence of unique characters
19	Image Archival procedure – agency contact info provenance	Publication author citing a (seismograph station) data	tracking living sources for anecdotal information	Y/N , text, URL address	Left blank
END					