

R: (...) Science Community ist in den *[Teildisziplin1]*, *[Teildisziplin1]*, schon ein relativ fester Bestandteil. Und, auch sehr, sehr, alles ist „community-driven“, „community-based“, alle Leute arbeiten zusammen und das ist ein offenes Miteinander und dadurch wird es auch gut vorangetrieben. Und, die Leute, die diesen Standard hauptsächlich mit entwickeln, die sind auch daran interessiert das zu verteilen durch eben Workshops und eben in Publikationen. Oder, dass das auch von Initiativen quasi, dass die da auch mit reingezogen werden. Dass zum Beispiel verschiedene Journals oder verschiedene große Gesellschaften über die Entwicklung Bescheid wissen und auch die Plattform geben können darüber zu reden. Und, ich glaube, dieser Community-Aspekt von den Leuten, die sich da zusammenfinden, und wie das implementiert ist, und versucht wird zu verteilen und zu verbessern, etwaige Strukturen gibt es in den anderen Bereichen nicht. (-) Das ist, wie ich das wahrnehme. Das andere ist natürlich, dass die (-) *[Teildisziplin1]* per se, eine höhere Notwendigkeit von Programmieren und IT und solchen Geschichten einher, als das jetzt vielleicht in einem kognitiven Bereich der Fall ist. Je nachdem, wie man es angeht. Um Gottes Willen. Und, ich glaube, einfach daher ist man sowieso eher in so Sachen drinnen. Und, im Programmieren oder halt JSON und diese Geschichten, was in anderen Bereichen nicht der Fall ist. (-) Also, dass ist so das, was wir immer mitkriegen. (-) Mhm, ja. #00:02:05#

Q: (...) #00:02:17#

R: Also, bei uns ist oder, was ich jetzt gemerkt habe, in den letzten drei Jahren, in denen ich relativ viele Workshops zu dem Thema gegeben habe, war, dass (-) man hat immer in der Wissenschaft so, so ein Spektrum. Von Student bis PI (...). Je niedriger man ist im Spektrum, je niedriger in der Nahrungskette ist, desto eher ist man natürlich bereit über Dinge zu reden. Noch nicht zwanzig Jahre eine bestimmte Sache, auf einen bestimmten Weg, Weg und Aspekt verfolgt hat. Das heißt, wenn ich jetzt diese Workshops mache, wenn man jetzt diesen Standard *[Dokumentationsstandard1]* vorstellt, wird es, da kann man auch, glaube ich, eine schöne Präregistrierung machen und Prädiktion dazu, weiß nicht, sagen wir mal fünfzig bis sechzig Prozent werden sagen „Warum? Ich sehe den Mehrwert nicht. Was bringt mir das?“. Was ich die letzten paar Jahre gemacht habe, wie auch, wie lange auch immer das ist, das hat funktioniert. Warum soll ich das jetzt ändern? Und dann ist es dieses klassische Mehrwert-, Wiedernutzbarkeits- Gedankenexperiment. Das heißt, wenn man dann sagt, wenn ihr jetzt, ihr könnt natürlich so weitermachen wie bisher, aber, wenn ihr jetzt Teil davon seid und diese Sachen nehmt und vielleicht sogar mit weiterentwickelt, helft ihr euch vor allen Dingen selber und auch den ganzen anderen Leuten, weil alle Daten eine höhere Wiederverwendbarkeit haben. Das heißt, ich muss nicht für jedes Experiment wieder komplett neu alles erheben. Ich kann

all diese Daten nutzen, die da sind, um wirklich adäquate Poweranalysen zu rechnen. (...) Und, ich kann vielleicht auch, je nachdem was da genau passieren soll, kann ich auch sowas wie *Analyse-Pipelines* und Sachen eher standardisieren.

(...)Weil bei [*Dokumentationsstandard2*] ist ja zum Beispiel soweit, ich setzte mich an das [*Messmethode*] (...) Gerät, nehme die Daten auf, komme am nächsten Tag wieder, der Datensatz ist gewandelt in Struktur, ist qualitätskontrolliert, vorverarbeitet und hat schon mal die ersten (paar t und F-Tests gerechnet, ohne dass ich da überhaupt irgendwas anfassen muss. Und, aufgrund dessen, dass das abläuft in bestimmten Virtualisierungsumgebungen der Berechnung kann ich das genau immer wieder so durchführen. Jedes Mal wieder und kann genau gucken, wo geht was schief, zum Beispiel jetzt hier einen anderen *Threshold* weil ich einen *Outlier* habe. Das kann ich alles wieder so machen und kann genau zeigen wo das herkommt. Also, dieses Vermitteln von eher nutzungsorientiert und nicht so mit dem (...). Und, das ist das, was ich am Anfang immer so ein bisschen gemacht habe, das funktioniert natürlich nicht so gut. (lacht) #00:05:38#

Q: (...) #00:05:40#

R: Nee, nee. Leider nicht. (lacht)

Q: (...) #00:05:47#

R: Okay. #00:05:47#

Q: (...) #00:05:50#

R: Ja, ja. Genau, ich will jetzt auch nicht (-) #00:05:52#

Q: (...) #00:05:54#

R: Ihre Zeit da (...) #00:05:54#

Q: (...) #00:05:59#

R: Gut, (...) je nachdem, was man macht, wenn es jetzt natürlich eher um behaviorale Sachen geht, wo vielleicht die zeitliche Auflösung nicht ein großes Problem darstellt, muss es auf jeden Fall die Software sein, die Version der Software, und, es tut nicht weh und ist nicht viel, auch vielleicht das

Betriebssystem, und soweit ich mich erinnern kann, die meisten kommerziellen brauchen sowieso Windows (-), also, aber nur für den Fall, falls vielleicht mal jemand was Anderes nimmt, dass man das dazu schreibt. (-) Punkt 14 und 15, nur für den Fall, dass ich es nicht richtig verstehe, da geht es jetzt schon um visuelle Stimuli? #00:07:01#

Q: (...) #00:07:02#

R: Mhm (zustimmend). #00:07:04#

Q: (...) #00:07:22#

R: Also, was (-) was ich jetzt dazu sagen kann, weil, aus (-) ich komme aus dem [Teildisziplin2], die, die vergleichbaren Punkte (-) wie Lautstärke und Mono / Stereo und auf jeden Fall wie (...) (-) genau (-) und für, wie sie schon sagten, Frequenz, die, die Darbietungsfrequenz. #00:08:06#

Q: (...) #00:08:17#

R. Ja, man...- Ja, es ist schwierig. Ist die Frage, ob man jetzt quasi nochmal ein Extrapunkt da für auditorisch aufmacht oder, ob man versucht übergeordnete Begriffe zu finden, die dann unterspezifiziert sind. #00:08:32#

Q: (...) #00:08:33#

R: Ja. #00:08:34#

Q: (...) #00:08:47#

R: Ja. #00:08:47#

Q: (...) #00:08:50#

R: Ja. Man könnte machen (-) Gute Frage. Stimulus-Eigenschaftenspezifikation. #00:08:59#

Q: (...) #00:09:13#

R: Ja. #00:09:14#

Q: (...) #00:09:15#

R: Weil zum Beispiel, sobald es multimodal wird, dann, dann braucht man es ja. (...) jetzt lese ich auch erst mal weiter. (lacht) #00:09:30# Okay, das, die ganzen Spezifikationen beziehen sich auf Rohdaten? #00:10:25#

Q: (...) #00:11:23#

R: Vergleichbares haben wir ja auch im [Dokumentationsstandard2]. Also, die Rohdaten und die Derivate folgen den gleichen Regeln. Wir gehen noch eine Stufe weiter und haben quasi noch ein Source-Datenlevel. Das heißt, bevor die Daten zu [Dokumentationsstandard2] kommen, dann [Dokumentationsstandard2]-Daten im Rohformat, Daten in Derivaten. Und, alle Aspekte sind auf all diese Stufen angelegt und tendenziell so gestaltet, auch zwischen den Schritten nachverfolgen kann, was passiert. Das wäre nämlich die Frage, jetzt zum Beispiel bei Punkt 7, wenn ich, habe ein paar *missing data* und die Leute stellen mir den Rohdatensatz zur Verfügung, aber primär den Derivatsdatensatz, ob zum Beispiel, ob *missing data* (...) #00:12:39#

Q: (...) #00:12:32#

R: Oh, Entschuldigung. Bei, bei mir wird jetzt, da die Frage bei Punkt 7(-), wenn die Leute jetzt primär oder hauptsächlich einen *derivated data set* zur Verfügung stellen, ob eine Interpolation passiert ist (-) Und, wenn ja, wie Und was. Vergleichbar verhält es sich mit dem *Scale*, ob zum Beispiel hier Transformationen passiert sind. Eine z-Transformation, Log, diese Geschichten. #00:13:17#

Q: (...) #00:13:18#

R: Das wäre dann die Frage. (...) Rohdaten und Derivaten #00:13:20#

Q: (...) #00:14:07#

R: Okay. #00:14:07#

Q: (...) #00:14:15#

R: Bin nicht sicher, ob ich es vielleicht überlesen habe, aber der Punkt einer Qualitäts-
Qualitätskontrolle? #00:14:37#

Q: (...) #00:14:43#

R: Das wäre ja auch die Sache, es kann gleich passieren nach der Datenerhebung, bevor ich es (...) in
ein Format übertrage. Oder von den Rohdaten zu den ersten Derivaten. Wenn ich zum Beispiel mir
meine Verteilung anschau und sehe, warum auch immer, eine Person ist ganz weit weg. Und, da ist
für mich klar ein *Outlier* und was mein, was mein Vorgehen ist, oder das Vorgehen in der Studie,
führe ich eine Qualitätskontrolle. Und, nach welchen Kriterien Daten ausgeschlossen werden. Wenn
man jetzt zum Beispiel sagt, eine Person, sobald jemand drei Standardabweichungen unter dem
Mean ist oder drüber, das ist klar ein *Outlier*. Dann wurde die Aufgabe nicht verstanden, oder da war
irgendein Fehler im Skript. Das aufgrund dieser Kriterien die Leute ausgeschlossen werden oder
dieser Datenteil. #00:15:44#

Q: (...) #00:15:54#

R: Ja. #00:15:54#

Q: (...) #00:16:01#

R: Mhm (zustimmend). #00:16:02#

Q: (...) #00:16:11#

R: Also, mir ging es tatsächlich auch eher um, um eine Beschreibung. Quasi ein, ein neuen
„*identifizier*“ „Qualitätskontrolle“ oder „Datenausschluss“ (-) ob da (-) das könnte (-) kann ja, oder ist
auf jeden Fall sehr wichtig. #00:16:31#

Q: (...) #00:16:34#

R: Je nachdem, an welcher Stelle die Leute jetzt den Standard ansetzen. #00:16:36#

Q: (...) #00:17:12#

R: Ja. Also, auch hier wieder, ich kann aufgrund meiner Vorerfahrung da in der Sache nicht so sehr reingehen. Ich habe natürlich im Bachelor irgendwann mal SPSS geklickt, aber dann auch nie wieder.
#00:17:27#

Q: (lacht) #00:17:29#

R: (...) Wenn man jetzt auf den (lacht) *Procedure/Syntax*, wäre auf jeden Fall Qualitätskontrolle...
#00:17:40#

Q: (...) #00:17:42#

R: Ja. Also, Entschuldigung, ich will nur noch sagen, ich möchte Sie nicht unterbrechen. Ich höre mich nur immer, ich weiß aber nur nicht, wann ich sprechen soll, weil ich die Feedbacks habe. #00:17:52#

Q: (...) #00:17:56#

R: Also wenn das passiert, ist nicht böse gemeint. #00:17:58#

Q: (...) (lacht) #00:17:59#

R: Tut mir leid. (-) Wenn man sagt, die Qualitätskontrolle findet statt, dann müsste die in dem *Procedure / Syntax* Teil enthalten sein. (...) Qualitätskontrolle ist, der zu einem bestimmten Skript verlinkt. Und dann aber vielleicht auch klarmacht, oder mit in einem Kommentar hat, was das Kriterium war. Ich weiß nicht, ob das aus dieser SPSS Syntax ersichtlich ist. #00:18:34#

Q: (...) #00:18:47#

R: Okay. #00:18:47#

Q: (...) #00:19:22#

R: Okay. Genau, das wäre auch noch so ein Punkt gewesen. Wenn man den *Code* bereitstellt, ist es natürlich gut. Aber, wenn der nicht dokumentiert ist, bringt es für die meisten Leute relativ wenig.
#00:19:35#

Q: (...) #00:19:53#

R: (-) Okay. (-) Wenn es unabhängig von der Sprache nachbaubar sein soll, bedarf es natürlich einer gewissen Nomenklatur. (...) #00:20:12#

Q: (...) #00:20:21#

R: Weil viele auch unterschiedliche Worte für das Gleiche nehmen. Oder für, oder andersrum.
#00:20:29#

Q: (...) #00:21:45#

R: Ja. #00:21:46#

Q. (...) #00:21:49#

R: Nee, ja auf jeden Fall. #00:21:51#

Q: (...) #00:21:55#

R: Mhm (zustimmend). (-) Also, ja. Nee klar. Wie gesagt, ich kann das mit dem SPSS nicht einschätzen. Aber natürlich alles, was manuell passiert, ist im Sinne einer Wiederverwendbarkeit, Reproduzierbarkeit nicht schön. #00:22:13#

Q: (...) #00:22:15#

R: Weil es schwieriger nachzuvollziehen ist. Bezüglich des *Textfiles*, das ist eine Idee (-) da hat man eine Art Protokoll, wo Schritt für Schritt... #00:22:32#

Q: (...) #00:22:32#

R: In der Syntax? Mit dem, es ist... #00:22:34#

Q: (...) #00:22:34#

R: „...provided“ eine *Textfile*. #00:22:37#

Q: (...) #00:23:22#

R: Mhm (zustimmend). #00:23:22#

Q: (...) #00:23:41#

R: Okay. (-). Was vielleicht so ein kleines Gimmick-Ding wäre, dass man die Variablen, die anderen Aspekte vielleicht auch noch verlinkt. Dass die Leute schneller darauf zugreifen können, um zu wissen, „okay, was war das *Scaling* hier und was war das nochmal genau“. #00:23:58#

Q: (...) #00:24:03#

R: Mhm (zustimmend). Genau. #00:24:03#

Q: (...) #00:24:05#

R: Einen, einen Aspekt, ich weiß nicht, ob Sie sich das schon mal angeschaut haben. Es gibt die Seite *[Plattform1]* #00:24:14#

Q: (...) #00:24:16#

R: Und, *[Plattform1]* macht im Prinzip (-) das. Aber quasi auch mit (-) Man schreibt ein Protokoll, was passiert ist mit den Daten, schickt das hin und kann aber verlinken, und kann das auch in *cloud computing* Sachen wieder ausführen. Und, genau, die verlinken halt verschiedene *Frameworks*. Also, Daten hier, *Code* hier (-) Analyse da oder so was. (...) könnte es vielleicht noch so ein *Pointer* geben. #00:24:59#

Q: (...) #00:25:15#

R: Mhm (zustimmend). #00:25:15#

Q: (...) #00:25:23#

R: Ja. Dann genau (-) das ist mir noch eingefallen, hier zu dem Teil „Design & Hypothesen“. Vor allen Dingen Design. Ja, genau. Da gibt es, natürlich auch aus dem [Teildisziplin1] kommend, aber da gibt es (...) genauso eine (...) #00:25:49#

Q: (...) #00:25:51#

R: Okay. Es gibt (...), die so eine Checkliste haben dafür, auf die sich die Community halbwegs geeinigt hat. Und auf diese Webseite geht man einfach drauf und geht diese Fragen durch. Dann wird man all diese Fragen gefragt, um sicher zu gehen, dass all diese Informationen drinnen sind. (...) einen automatisierten Text raus. (-) Und, unter einer „zero-Lizenz“ einfach für das Paper nehmen kann. Für den Teil mit dem Design und den (-) Hypothese und der Poweranalyse gibt es auch Tools dafür wo man das online eingeben kann. (...) Daten, die die Effektstärke berechnet aber man kriegt dann auch wieder so ein Text raus. Oder es wird berechnet, was das bestmögliche Design ist, zum Beispiel. Und das ist natürlich, vor allen Dingen, wenn man daran interessiert ist, dass das die Leute das wieder nehmen und dann ne hohe Wiederverwendung hat. Das sind natürlich schöne Tools, die man den Leuten, den Leuten zeigen kann. #00:27:19#

Q: (...) #00:27:28#

R: Ehm, das kommt drauf an. Es sind natürlich Effektstärken relativ, ganz wichtig, Verteilung von vergleichbaren Altdatensätzen und im [Teildisziplin1] natürlich die, die physiologischen Marker. #00:27:51#

Q: (...) #00:27:

R: Ehm, das... #00:27:57#

Q: (...) #00:28:27#

R: Es (-) Genau. Es kommt natürlich ganz auf den, auf den Inhalt der Studie an. Aber, wenn man jetzt im behavioralen Bereich bleibt, wo Leute (-) es gibt verschiedene Trials, es gibt verschiedene Fragen und Items haben eine gewisse (...) haben einen gewissen (-) Wahrscheinlichkeit aufzutauchen und all diese Sachen, die kann man halt relativ gut berechnen. *Counterbalancing*, beste Anzahl an

Repetitionen. Also, das, das je nach, wie gesagt, je nach Studiendesign und was man genau untersuchen möchte, dann funktioniert das relativ gut. #00:29:08#

Q: (...) #00:29:13#

R: Und, dann habe ich mal aber auch die Erfahrung gemacht, dass, wenn man Leuten, die so ein bisschen, naja, vielleicht nicht ganz so überzeugt sind, kann man da auch sagen „hier, gibt es, basierend auf diesem Standard, diese, was damit integriert, viele Tools, die einfach helfen können“. Weil im Endeffekt wäre ja, niemand will irgendwie den zwanzigsten Methodenteil lesen, wo man erst wieder hin- und herschauen muss und was immer passiert, sondern, wenn man auch diese, die Reports davon standardisiert, kann man natürlich viel, viel leichter die Informationen herausbekommen. Man kann dann am Ende zum Beispiel auch viel leichter Meta-Analysen machen mit automatischen Abfragen (...). Wenn man jetzt zum Beispiel denkt an die klassische G-Power-Anwendung, wo man den Studenten sagt, Studierenden sagt „mach da jetzt mal was damit“, dann fangen die da an, sagen „joa, (...) Effektstärke 0.8“ und dann war es das so ein bisschen, ne. Also, (...) aufgrund der nicht grade bestehenden offenen Datenteilung und –Struktur, ist es natürlich auch sehr schwierig informierte Power-Analysen zu machen. (...) Daten nutzen, die Sachen durchgehen und schauen was wie rauskommt. Alleine das Berechnen von Effektstärken, das wird ja zum Glück ein bisschen mehr und mehr, würde ja da schon viel helfen. #00:31:02#

Q: (...) #00:31:04#

R: Okay, aber, ich glaube, jetzt bin ich auch erst mal (lacht) Ich befürchte... es tut mir leid. (lacht) #00:31:15#

Q: (...) #00:31:19#

R: Genau. Dann habe ich doch jetzt aber auch (-) das soweit durchgelesen. #00:31:34# Und jetzt würde ich auf „Rohdaten“ klicken und mir das ganze anschauen. #00:31:59#

Q: (...) #00:32:14#

R: Ja. Ja, ich versuche gerade das Menü zu verstehen. #00:32:18#

Q: (...) #00:32:20#

R: Welche. Ah, okay. Okay. Ja. Entschuldigung. (lacht) Ich dachte jetzt, das, was ich eben gelesen hätte, wäre... #00:32:28#

Q: (...) #00:32:39#

R: Ja. #00:32:40#

Q: (...) #00:32:45#

R: Okay. Ich dachte jetzt, das Manual würde quasi hier nochmal die Menüpunkte einzeln aufzählen. #00:32:51#

Q: (...) #00:32:53#

R: Aus dem Manual. #00:32:53#

Q: (...) #00:32:59#

R: Ja. Nee. Genau. Deswegen habe ich das jetzt vorhin missverstanden. Okay. Jetzt habe ich es. Entschuldigung. #00:33:06# Und, nur zur Sicherheit, wenn es jetzt um die Forschungsfrage geht, geht es schon da drum, dass so objektiv wie möglich hier zu berichten und nicht was ich denke, was untersucht wird, gegeben dem Design und den, der Hypothese untenstehend. #00:35:08#

Q: (...) #00:35:18#

R: Mhm (zustimmend). Okay. Und, Sie meinten, *copy paste* ist okay. #00:35:22#

Q: (...) #00:35:24#

R: Okay. (lacht) #00:35:26#

Q: (...) #00:35:47#

R: Ah, jetzt. #00:36:55#

Q: (...) #00:36:56#

R: Glaube ich. Hat nur ein bisschen gedauert. #00:36:58#

Q: (...) #00:37:12#

R: Ist schon okay. #00:37:16# Mhm. Ah, jetzt ah, okay. (...) #00:38:00#

Q: (...) #00:42:57#

R: Ich versuche es nur so sicher wie möglich zu beantworten. (-) lacht #00:43:06# Die Frage bezieht sich jetzt rein auf die Syntax und nicht auf diese Punkte im Design. #00:44:40#

Q: (...) #00:44:42#

R: Also, wenn ich jetzt sagen wollen würde, ich werde, meiner Meinung nach oder für die Beantwortung der Frage bräuchte man schon nochmal so zehn andere Variablen. Das wäre jetzt hier nicht miteingeschlossen. #00:44:59#

Q: (...) #00:45:04#

R: Okay. (lacht) Sondern rein, was in dem SPSM, SPSS passiert ist. #00:45:12#

Q: (...) #00:54:19#

R: (lacht) Okay. #00:45:21#

Q: (...) #00:45:26#

R: Oh, Entschuldigung. Hätte ich das jetzt da machen sollen? #00:45:54#

Q: (...) #00:45:59#

R: Ah, okay. #00:45:59#

Q: (...) #00:46:02#

R: (lacht) Ich bin, ich komme aus der Richtung, wo ich immer versuche zu schauen, was ist im Sinne eines Modells möglich. Und, da gibt es natürlich das Spektrum, ich mache eine klassische parametrische Testung im Sinne wie die hier stattgefunden haben. Oder ich gehe vielleicht in Modelle, die ein bisschen komplexer sind, die mir vielleicht mehr Aufschluss geben könnten, aber natürlich auch die Gefahr haben schnell zu „overflippen“. Und, für den Fall hier hätte ich auch genauso eine Frage gehabt, nämlich, ob man nicht hätte ein Modell nehmen können, was vielleicht nicht, was alle Variablen gleichzeitig hat, aber doch vielleicht ein paar mehr sich in einem Modell anschaut. Und nicht zum Beispiel die Sachen alle einzeln durchtestet hinsichtlich der Kontrollvariablen, sondern vielleicht ein Modell, wie gesagt baut, was mehrere miteinfließt. Wäre dann vielleicht die Varianzaufklärung anzuschauen oder das Kovarianz-Muster. Das hätte ich in dem Fall noch getan. #00:47:40#

Q: (...) #00:47:43#

R: Ist das so in Ordnung so? #00:47:45#

Q: (...) #00:47:46#

R: Okay. #00:47:47#

Q: (...) #00:47:49#

R: (...) #00:47:51#

Q: (...) #00:48:28#

R: Das wäre jetzt die Frage gewesen. Ja. Sehr gut. #00:48:32# Für diese Frage jetzt (-) Also, wirklich, wenn ich nur dieses *Codebook* hätte, und auch zum Beispiel mir nicht diese, diese Studie durchlesen könnte, sondern nur die Sache, was jetzt hier ist, ob ich es dann wiederherstellen könnte. #00:53:44#

Q: (...) #00:53:47#

R: Muss nur kurz überlegen, wie ich es formuliere. #00:56:12# Okay, ist eine gute Bewertung. (lacht) Sehr, sehr schön. #01:00:52#

Q: (...) #01:01:10#

R: Mhm (zustimmend). #01:01:11#

Q: (...) #01:01:16#

R: Also, bis jetzt, ich weiß auch nicht, ob ich es richtig oder falsch verstanden habe, aber für mich wäre jetzt, das zum Beispiel würde für mich bedeuten, es gibt, es folgt diese Bedingung in der Ausprägung hier, in dem Fall *Mindfulness*, aber, da ich keine Informationen über Durchführung, Anfang, Ende und sowas habe, ist der Indikator für „Jetzt kommt diese Bedingung“ und dann folgen diese anderen Teile innerhalb der Bedingung. #01:02:02#

Q: (...) #01:01:03#

R: Wieder dann Zeiten haben aber ich (...). Deswegen meinte ich ist es möglich, aber es bedarf eben eines genauen Studiums. Es ist nicht so (-) wie sagt man das #01:02:24#

Q: (...) #01:02:25#

R: Ja, genau. Nicht so intuitiv. #01:02:17#

Q: (...) #01:02:30#

R: „*straightforward*“, wie es vielleicht sein könnte (...) #01:02:33#

Q: (...) #01:02:57#

R: Ja, ich denke, das wäre ein guter Weg. Wenn man auch sich die *Procedure* anguckt. Hier habe ich ja auch gewisse (-) Hier ist das ja drinnen, das wäre jetzt meine *Mindfulness* oder *Mindwandering*. Dann bekomme ich mein Paradigma und er-, erhebe direkt so (...). Und, vergleichbar (-) Nee, jetzt habe ich mich wieder verklickt. Entschuldigung. (lacht) Hier. *Codebook*. Jetzt für mich (...) würde, glaube ich, sowas helfen. Wie Sie meinten, es wird dieses *Nesting*, dass man jetzt zum Beispiel hätte

(...) Durchläufe (...) behavioraler Daten (...) Dann mache ich das auf und sehe da drunter, okay, dann kommen diese ganzen Punkte und die Ausprägungen. Und, dann hier mit dem weiter. Wenn man jetzt zum Beispiel sich mehr damit (...) beschäftigt und auskennt, ist das okay. Aber jetzt zum Beispiel jemand wie ich. Ich habe nie in diesem Bereich gearbeitet. Ich kann das jetzt methodisch, verstehe ich das ungefähr schon. Ich weiß auch, glaube ich, ungefähr worum es gehen würde in der Studie. Jetzt nicht, ganz einfach das nachvollziehen. (...) okay (...) die Instruktionen an. Habe da eins, zwei, drei und (...) was angenehm wäre... #01:03:48#

Q: (...) #01:04:53#

R: Für mich zum Beispiel, wenn es (-) Ist, ist es wieder okay? #01:04:58#

Q: (...) #01:05:08#

R: Ja, dann gut. Gut. Ich habe danach auch gleich aufgehört. (-) Wenn das von der *Procedure* vielleicht so ähnlich aufgebaut wäre wie die Syntax wie dieser Reiter, dass man das *Procedure* hätte. Einem Art grafischen Repräsentation (...) des Ablaufs und dann auf die einzelnen Sachen des *Codebooks*. Oder vielleicht mit das farblich markieren könnte. Oder eben diese ausklappbaren Reiter. #01:05:55#

Q: (...) #01:05:56#

R: Ich glaube, das wäre jetzt, wenn man, wie gesagt wie ich von der Sache überhaupt keine Ahnung hat, es nie gelesen hat, würde ich mir jetzt wünschen, ich hätte die Präsentation des Paradigmas, man kann da drunter schauen, und kann, also, die relativ simplifiziert ist und kann drunter schauen und drauf klicken und kriegt mehr Informationen in einer genesteten Struktur, was diese einzelnen Sachen sind. Und die dann auch wieder angenehm verlinken auf, zum Beispiel, die Variable verlinkt wieder auf die Syntax und andersrum. #01:06:36#

Q: (...) #01:07:45#

R: Ich (-) Bei mir ist jetzt das Problem, ich komme natürlich aus einer Richtung, wo relativ wenig mit *[Forschungsdesign]* gearbeitet wird. Dahingehend würde ich jetzt sagen, für mich persönlich würde das sehr viel Sinn ergeben. Schon allein, wenn ich diese Sachen sehe und ich sehe nur die Daten, die subst-, also, nur die *substantial data*, dann, wenn ich jetzt von einem *Data Science* Standpunkt

rangehe, müsste das so gesetzt sein, und mir so viel Informationen geben, dass ich nur damit arbeiten kann. Und, es nicht einer weiteren Information bedarf, wie das erhoben wurde, was die genauen Besonderheiten sind, sondern dieser Klassiker „die Daten sollten für sich sprechen“.

#01:08:50#

Q: (...) #00:01:08:51#

R: Und dann (-) Genau, ja. (lacht) Anders wird es schwierig. Ja auf jeden Fall (...) Aber, ich glaube, die Teilung, die Sie beschrieben haben, würde Sinn ergeben. Weil jetzt momentan hätte ich das Gefühl persönlich, dass das *Procedure* ein bisschen zu schlank ist. Da in dem *Procedure* selber würde ich mir wünschen, dass ich vielleicht, also, ja, gut, das ist auch hilfreich hier als Grafik, zum Beispiel, dass ich auch wie so eine Art Flowchart hätte. Und, wenn ich auf, jetzt zum Beispiel diesen Teil hier draufklicke von dem Flowchart, kriege ich dann die Informationen mit den Instruktionen, was da genau passiert, wie der genaue, da drinnen nochmal der Ablauf ist. Weil hier...#01:09:46#

Q: (...) #01:09:46#

R: Weil das ist ja in dem *Codebook* hier sehe ich das jetzt als Eins. Es gibt *Mindfulness* oder *Mindwandering*. Und... #01:09:55

Q: (...) #01:09:57#

R: Ja. Und in dem *Codebook* (-) nee, Moment. In der *Procedure* (-) Doch im *Codebook*. Ah. So. da würde ich ja dann sehen, okay, es geht los, es kommt die eine oder das andere, das sind die Instruktionen, halt es gibt eine *Practice* und dann wird das eigentlich ausgeführt. Und, genau das hätte ich, jetzt für mich gerne, in dem *Procedure* drinnen. #01:10:28#

Q: (...) #01:10:35#

R: Als Flowchart oder, dass ich vielleicht eine kleinere Version des *Codebooks*, wenn ich auf, da draufklicke, dann das kommt. #01:10:44#

Q: (...) #01:10:48#

R: Ehm, weil in dem Fall könnte das *Codebook* als Term auch ein bisschen nicht verwirrend sein, aber vielleicht auch nicht genau das beschreiben, was man erreichen möchte. #01:11:05#

Q: (...) #01:11:07#

R: (...) Das... #01:11:10#

Q: (...) #01:11:12#

R: Das *Codebook* für mich wäre jetzt, wäre eine, die, eine Beschreibung der Variablen. Also, ich habe die Variable, das bezeichne das, das ist die Skala, das ist die Ausprägung und das ist der Wert, den ich für, den ich bekommen habe. Und, die anderen Sachen sind ja die (-), die Schritte, die dem vorausgehen. Also, ich brauche diese ganzen Schritte, um dann am Ende meine Variable zu bekommen. #01:11:57#

Q: (...) #01:11:59#

R: Ja. #01:12:00#

Q: (...) #01:12:22#

R: Ja. #01:12:23#

Q: (...) #01:12:25#

R: Ich glaube, dafür könnte dieses, das... #01:12:27#

Q: (...) #01:12:27#

R: „Protocol.io“ vergleichbarer Tools einen schönen Ansatzpunkt bieten. (...) vergleichbar. #01:12:40#

Q: (...) #01:12:42#

R: Und, hier, Entschuldigung, hier muss, weil ich das bewusst auch von mir nie gesehen habe. Das ist, die Sachen würde ich jetzt (-) Ich jetzt öffnen und hier rauskopieren. #01:13:10#

Q: (...) #01:13:15#

R: Die Frage, nur kurz für mich zum Verständnis. #01:13:18#

Q: (...) #01:13:18#

R: Und ich, es wäre jetzt so, ich hätte den Datensatz als (-) und würde den mit SPSS öffnen.

#01:13:33#

Q: (...) #01:13:40#

R: Und, in SPSS (...) #01:13:43#

Q: (...) #01:13:48#

R: Ja, das Echo dauert auch länger. #01:13:53#

Q: (...) #01:14:00#

R: Würde ich diesen Code hier kopieren. #01:14:07#

Q: (...) #01:14:09#

R: SPSS öffnen und dort einfügen (...) und mal durchlaufen lassen. #01:14:16# Okay. #01:14:20#

Q: (...) #01:14:42#

R: Mhm (zustimmend). (lacht) #01:14:44#

Q: (...) #01:14:52#

R: Ja. Das stimmt schon. Wäre vielleicht eine Art (-) Ich weiß nicht, wie ich es beschreiben soll. Eine Art Hilfestellung oder Ähnliches, oder eine ungefähre Vorgabe für die statistischen... #01:15:20#

Q: (...) #01:15:35#

R: Bin ich jetzt wieder da? #01:16:09#

Q: (...) #01:16:11#

R: Okay. (lacht) Dieses Internet. Ich weiß nicht, wieviel Sie noch gehört haben. Meine Frage war, ob man die statistische Terminologie vielleicht so ein bisschen auch, in gewissen Teilen zumindest, standardisieren könnte. Sozusagen das sind die Operationen, die abgelaufen sind und hier ist jetzt damit das Modell beschrieben auf eine Art und Weise, die sich übertragen lässt. #01:16:47#

Q: (...) #01:17:12#

R: Mhm. #01:17:12#

Q: (...) #01:17:31#

R: Also (-) Wenn ich mir das jetzt so anschau (-) Ich meine, ich könnte es jetzt zum Beispiel mit R oder Python nachbauen. Nicht alles. Weil ich jetzt zum Beispiel nicht wüsste, was das jetzt, was „SSType“ beschreibt. #01:17:53#

Q: (...) #01:17:58#

R: Ja. Ja. (lacht) (--) Oder, ja oder, vielleicht so ein Grundset an, an Vari-, an Modellspezifikationen, was unabhängig der Software beschrieben werden kann. Zum Beispiel ANOVA, das mal das, korrigiert für, mit der und der Korrektur hinterher. #01:18:46#

Q: (...) #01:18:47#

R: Irgendwie so... #01:18:49#

Q: (...) #01:18:58#

R: Ja. (-) Und die andere Frage, ich weiß nicht, ob das dann schon „out of scope“ wäre, wäre sowas wie Modelldiagnostik. #01:19:13#

Q: (...) #01:19:15#

R: Also, ...#01:19:17#

Q: (...) #01:19:17#

R: (lacht) sowas wie R Square und solche Geschichten. Oder, dass man sagt, okay, bitte auch den, die Effektstärke mit angeben und sowas. #01:19:32#

Q: (...) #01:19:38#

R: Mhm, das würde, wie gesagt, wenn man jetzt noch eine Zeit bei G-Power bleibt, auch, also ich gehe jetzt davon aus, ich würde das jetzt gerne nochmal vergleichbar machen oder vielleicht einen anderen Aspekt untersuchen in einem vergleichbaren Setting, dann würde mir ja die Effektstärke, die hier gefunden wurde für ein bestimmtes Modell, schon mal weiterhelfen hinsichtlich meiner Planung. #01:20:10#

Q: (...) #01:20:16#

R: Genau. #01:20:18#

Q: (...) #01:20:26#

R: Ja (lacht) #01:20:27#

Q: (...) #01:20:44#

R: Also, wie gesagt, wenn man jetzt davon ausgeht, dass, okay das einfachste wäre, die Leute schreiben es selber hin. Das andere wäre, wenn man jetzt wüsste es werden die Daten werden immer geteilt, und sind immer wieder verwendbar, dann könnte ich die Effektstärke auch selber herausfinden. Oder bei meiner Fragestellung als die Variable miterhoben wurde, einfach durchführen. (-) Wenn man jetzt geht, noch so im Allgemeinen, nach den *FAIR Principles* für *Research Data*, ich, dann würde ich meinen, und das ist natürlich auch eine *biased opinion*, dass das auf jeden Fall auf einem guten Weg ist, aber noch einige Aspekte fehlen. (-) Wo das jetzt genau angreift und,

ob das jetzt für alle Daten und Vorgehen möglich ist überhaupt, zumindest momentan, wage ich gerade keine Aussage drüber zu treffen. #01:22:09#

Q: (...) #01:22:17#

R: Ja. #01:22:17#

Q: (...) #01:22:50#

R: Ja. #01:22:50#

Q: (...) #01:23:26#

R: (lacht) #01:23:26#

Q: (...) #01:23:39#

R: Ja. #01:23:39#

Q: (...) #01:23:53#

R: Nee, also, ich glaube auch, wenn man hier den ganzen, das ganze Feld sich anschaut, es ist einfach sehr, sehr divers. Und, sobald man eben kommt, in sowas wie Interviews, das was wir gerade machen, das ist ja, würde ich sagen, momentan nahezu überhaupt nicht standardisierbar oder quasi fair aufbereitbar. Das ist einfach der Methodik geschuldet. #01:24:28#

Q: (...) #01:24:29#

R: Und, für die Sachen hier, meine wirklich das ist ein guter Weg und ich, wie Sie gesagt haben, bewegt sich auf einem Level, was für viele Leute möglich sein sollte, ohne jetzt zu viel aufzuerlegen. Es geht natürlich damit einher, dass für das andere Ende, also, von FAIR jetzt ausgehend, der, der Endnutzer der Daten vielleicht ein bisschen mehr Arbeit investieren muss. #01:25:06#

Q: (...) #01:25:07#

R: Aber (-) besser als gar nichts. Auf jeden Fall viel, viel besser als das, was gerade stattfindet.

#01:25:14#

Q: (...) #01:25:40#

R: Ja. #01:25:41#

Q: (...) #01:25:54#

R: Mhm (zustimmend). #01:25:54#

Q: (...) #01:26:05#

R: Ja. Nee, klar. Das, das geht ja sofort damit einher. Wenn jetzt Leute auf ihre Universitätswebsite oder Labwebsite oder pers-, oder eigene Homepage diese Sachen stellen, ist das auch schon mal gut und viel, viel besser als nichts. Aber in den Kreisen, wo ich mich bewege, sagt man dann schon, wenn die das so machen, also, quasi Daten hochladen und hier sind die Skripts und das war es, dann ist es genauso gut wie gar nichts, weil ich nichts damit anfangen kann. #01:26:46#

Q: (...) #01:26:47#

R: Und, genau dieser Punkt hier mit der Dokumentation und Verlinkung der einzelnen Teile, das ist auf jeden Fall schon sehr hilfreich. Und der Weg, den man braucht. #01:27:00#

Q: (...) #01:27:01#

R: Was man noch überlegen könnte vielleicht, je nachdem, wie das dann da reinpasst, man könnte vielleicht so eine kleine *Section* haben, die auf gewisse Tools oder Ressourcen verweist als Idee. Weil ich habe auch immer wieder festgestellt, dass viele Leute auch nicht die, überhaupt nicht den Überblick haben, was, was möglich ist. Dass es zum Beispiel gezielt *data repositories* gibt. #01:27:39#

Q: (...) #01:27:49#

R: Das... #01:27:49#

Q: (...) #01:28:08#

R: Ja. #01:28:08#

Q: (...) #01:28:15#

R: Ja. Also, ich denke, [Plattform2] ist auf jeden Fall eine gute Idee. Das ist schon weiter bekannt und erlaubt. Ja, komplett freie Standards zur Strukturierung. Für die [Teildisziplin1] Sachen, was ich jetzt sagen kann, wurden halt andere Wege gefunden, die halt spezifisch funktionieren. Weil zum Beispiel die Datensätze, die wir haben sehr, sehr vieles größer sind im Sinne von dem Speicherplatz als viele Verhaltensdaten, was nochmal spezielle Anforderungen braucht. Aber, wie gesagt, die Gefahr ist immer, wenn man das komplett dezentralisiert und sagt, jeder, jedes Lab, jede Uni, das ist ja momentan der Fall in Deutschland und auch anderswo, hat ihr eigenes *Data Repository* und alle laden das dahin. Dann ist das natürlich auch okay, wenn das entsprechend in der Publikation so vermerkt ist. Aber, wenn man jetzt zum Beispiel die Möglichkeit hat zum Beispiel [Plattform2] und die protocols.io jetzt als zusätzliches *Framework* zu nehmen, worauf die Strukturen hier, die Sie hier erarbeitet haben, schon sehr gut passen, ist das, glaube ich, auch dann um einiges leichter. Wenn man sagen kann unter der Haube nimmt man einfach diese Sachen und das ist ja durchaus möglich, dass man bestimmte für Institute, oder für Gesellschaften bestimmte Varianten von [Plattform2] bekommen kann. Und, dass man das (...) dann kann man nämlich, dann hat es diese schöne Teilung von, hier sind die Daten, hier sind die Skripte und hier ist die Dokumentation, die erklärt, was wie wo wann warum passiert ist. #01:30:24#

Q: (...) #01:30:25#

R: Und, ich glaube, vieles andere, was gerade so hochkommt und „*state of the art*“ wird, ist, glaube ich, für das Publikum an, um das es hier hauptsächlich gehen soll, ein bisschen zu, zu viel IT.
#01:30:49#

Q: (...) #01:31:10#

R: (lacht) #01:31:10#

Q: (...) #01:31:17#

R: Ja. #01:31:17#

Q: (...) #01:31:22#

R: Nee, das, ja, das ist der Punkt. *[Skriptsprache]* ist, wenn man sich die Zeit nimmt und reinguckt, sehr, sehr schön. Weil es eben diese *[Attribut]* Strukturen erlaubt und eben auch von Menschen und Maschinen lesbar ist. Aber, wenn man jetzt natürlich primär, ich sag das jetzt wertungsfrei, SPSS *Files* oder Excel-Dateien rumschickt, ist das nochmal, ist das eine andere Hausnummer. Und, ähnlich verhält es sich, jetzt vor allen Dingen bezogen auf die Sachen hier mit den Daten und den Ableitungen davon und den Modellen, da weiß ich nicht, ob Sie da schon mal drüber gestolpert sind. Es gibt dann die Sache, die nennt sich *[Projekt]*. #01:32:16#

Q: (...) #01:32:20#

R: *[Projekt]*. Das kommt von, hauptsächlich von *[Name]* aus *[Stadt1]*. #01:32:29#

Q: (...) #01:32:30#

R: Und Kollegen. Und, das ist im Prinzip eine Versionskontrolle für Daten. Und damit ist es möglich, jedes, jede einzelne Operation, die passiert ist auf ein Datenset quasi zu verlinken. Und, was man damit machen kann, ist, also, wenn ich jetzt zum Beispiel ein Datenset bekomme und schauen möchte was passiert ist und die Leute, die die Verarbeitung gemacht haben, kann ich mir dieses *data set* nehmen und in dem *data set* selber ist eine, ist ein Log drinnen, ein *log File* mit den einzelnen Befehlen, die gelaufen sind. Und, die kann ich mit bestimmten Voraussetzungen hinsichtlich der Rechenstruktur, kann ich genauso wieder ausführen. #01:33:23#

Q: (...) #01:33:24#

R: Wie die das gemacht haben. Das heißt, ich kann von einem Rohdatensatz bis zu den finalen Visualisierungen gehen. Und jeden einzelnen Schritt wieder genauso ausführen (...), weil es dazu verlinkt ist. Und das ist das, was grade so passiert. Aber, wenn man das den Leuten sagt, von Anfängen mit Versionskontrolle, ist es meistens sowieso schon vorbei. (-) Und, das ist ja auch nicht die, die Schuld der Leute. Es ist ja einfach, ich weiß nicht, wie Sie das erlebt haben, aber ich, bei mir im Studium, bei vielen von meinen Bekannten war es so, dass für die Sachen, die wir hier machen,

quasi die ersten ein, zwei Jahre vom Doktor nochmal so ein Informatikstudium waren quasi.

#01:34:14#

Q: (...) #01:34:15#

R: Um diese ganzen Sachen drauf zu bekommen. Weil das natürlich auch im Psychologiestudium und in vielen anderen empirischen Wissenschaften auch nicht der Platz da ist oder nicht eingeräumt wird.

Von daher (-) #01:34:34#

Q: (...) #01:34:38#

R: Ja. #01:34:38#

Q: (...) #01:34:45#

R: (lacht) #01:34:45#

Q: (...) #01:35:02#

R: Ja. #01:35:02#

Q: (...) #01:35:25#

R: Ja. #01:35:25#

Q: (...) #01:35:27#

R: Ja. Stimmt. #01:35:27#

Q: (...) #01:35:40#

R: Es ist nicht einfach, ja. (lacht) Das stimmt. (-) Je nachdem, was jetzt bei Ihnen in dem Projekt, oder vielleicht auch darüber hinaus möglich ist, was gerade bei uns in den *[Teildisziplin1]* passiert, ist genau versuchen, diese Hürde so ein bisschen anzugehen. Nämlich da drüber all diese Strukturen zur Verfügung zu stellen, das heißt, virtualisierte Rechenumgebung, Versionskontrolle der Daten und

vielleicht auch komplexe Analysen, aber jetzt ohne die Notwendigkeit selber, selber zu programmieren. Das heißt, man stellt Plattformen zur Verfügung, die genau darauf beruhen, das Ausführen im Hintergrund, aber die Leute ihre gewohnten Analysen weitermachen. #01:36:34#

Q: (...) #01:36:36#

R: Man aber dann durch die zugrundeliegende, oder unterliegende Struktur, diese Sache miteinfließen lassen kann, um eben FAIRness zu erhöhen. #01:36:48#

Q: (...) #01:36:57#

R: Mhm (zustimmend). #01:36:57#

Q: (...) #01:37:22#

R: Ja. Nee, ja genau. Genau. Sobald ich den Leuten im Seminar erzähle wir machen jetzt alles in *[web application1]* und darin macht ihr das, ist die Freude meistens gering. #01:37:35#

Q: (...) #01:37:37#

R: Und, ehm, aber da ist noch ein Punkt, eine schöne andere Geschichte, die auch das alles integriert. Das nennt sich, ist eine Weiterentwicklung von *[web application1]*. Und, das nennt sich *[web application2]*. Und, die Idee da ist, die *[web application1]* Idee auszubauen. Und, es würde ja auch hier für den Fall, vor allen Dingen in diesem Syntaxfeld, relativ gut passen. Und, die Idee da ist eben, dass man *[web application1]* noch eine Stufe höher nimmt. Das heißt, man hat wirklich ein, eine interaktive wieder ausführbare Publikation, wo man eine Website hat mit verschiedenen Bereichen/Unterpunkten, die dann einen Text haben im Sinne der Dokumentation, die das beschreiben. Und da drunter aber die, die Ausführung über *[web application1]* Zellen. Und, dass man da wirklich den Leuten so ein, so ein „walk through“ bereitstellen kann. #01:38:50#

Q: (...) #01:38:51#

R: Und dann kann man sich diese, ehm (-) Also, ich kann das ja auch hier einmal auf... Ist das okay, wenn ich das einmal hier aufmache? #01:39:01#

Q: (...) #01:39:01#

R: Nur, dass ich halt nicht (-) Das ist von einer Kollegin von mir hier ein Beispiel dafür. Und, das würde eben so aussehen. Man hat hier die Infos dazu mit Inhaltsangaben. Das jetzt zum Beispiel ist ein Tutorial, was sie geschrieben hat. Und, man kann jetzt hier die einzelnen Aspekte durchgehen, man hat hier die Dokumentation dazu mit Links, mit Infos und die *Code* (...). #01:39:35#

Q: (...) #01:39:36#

R: Aber auch den Output. Und kann dann auch weitermachen und kann dann hier zum Beispiel auch, das ist alles interaktiv. Man kann hier zum Beispiel auch dann eine Verteilung von Daten plotten. Und das wirklich coole kommt dann, man, klar hier Fullscreen, aber man kann sich das runterladen als *[Output1]*, als *[Output2]*, als *[Output3]*. Oder, es gibt ein paar neue schöne Geschichten. Entweder personalisierte *[Output4]*, also, Rechenumgebungen spezifisch für *[web application1]* und ähnliche Geschichten basierend auf dem *[Framework]*. Das ist eine, eine Plattform, die kostenfrei, online Rechenkapazität zur Verfügung stellt. Und, was das macht, es nimmt eine Virtualisierungssoftware, man nimmt die (...) *[web application1]* da, die man hat, stellt die zur Verfügung und dann sagt man „bitte jetzt starten“, und dann stellt das alles für einen selber bereit. Wenn ich jetzt zum Beispiel hier draufklicke, werde ich weitergeleitet auf diese Seite. Jetzt bildet sich im Hintergrund, wenn man sich das anschaut dieses, die Rechenstruktur und je nachdem, weil es ein freier Service ist, dauert es halt manchmal ein bisschen (lacht). Und, was jetzt aber gleich passiert ist, das *[web application1]*, was ich vorher hatte, wird jetzt hier interaktiv aufgemacht und ich kann es genauso ausführen, wie es da passiert ist. Und, wenn ich diese *[web application1]* Implementation habe, kann ich das quasi auch genau da drinnen gleich so machen. Und das dann... #01:41:30#

Q: (...) #01:41:50#

R: Ganz genau. Wenn es eine Rechenkapazität gibt da, ist das eine sehr gute Idee. Das Ganze kommt hier bei uns aus dem *[Projekt2]*. Hier von uns aus in *[Stadt2]*, wo auch die Kollegin von eben, die das gemacht hat und ein paar andere Leute, sehr viel in der Entwicklungsarbeit machen und mal gucken. Ich versuche es grade aufzumachen. Ah, da bin ich, okay. (lacht) Und, hier ist nämlich die Idee von *[Forschungsidee]*. Und das ist quasi wie so eine kleine eigene Submission, wo man, je nachdem, wie das mit der Hauptsubmission abläuft quasi den Zusatz „*submitten*“ kann. Oder, wenn es um spezifischere Projekte geht, man kann auch das ganze Paper dahin „*submitten*“. Und, (...) das ist alles

halt noch in, grade im Entstehen. Also, das ist, das ist noch nicht „*state of the art*“ (...), das kommt grade so deswegen, ist hier noch nichts weiter da. Aber...#01:43:10#

Q: (...) #01:43:10#

R: Mal gucken, ob (...) hier mal irgendwo ist, es gibt die Dokumentation noch dazu. #01:43:17#

Q: (...) #01:43:23#

R: Genau, die kann man nutzen. Hier komme ich natürlich jetzt gerade aufgrund der Internetverbindung nicht durch. #01:43:29#

Q: (...) #01:43:31#

R: Aber da würde das hier genau passieren. Und, da braucht es auch gar nicht viel. Und meine *[framework]*-Seite sieht so aus. Und man müsste eigentlich nur, man gibt hier den Link an zu dem *[Repository]*. Wie man sieht mache ich das häufiger (lacht) Und, zum Beispiel, was ich mache für alle meine Workshops. Erstelle ich diese Sachen für Leute, die nicht die Rechenkapazität zur Verfügung haben. Und dann mache ich den Link rein, klicke einfach auf (...) und dann muss ich kurz warten bis das passiert. Momentan scheint wieder ein bisschen mehr los zu sein. #01:44:17#

Q: (...) #01:44:18#

R: Deswegen dauert es. (-) Ja. #01:44:27#

Q: (...) #01:44:42#

R: Nee, also, es ist halt auch, wenn man es sich überlegt, dieser Wahnsinn jeden Datensatz wieder neu zu erheben. (-) Ist ja auch nicht angenehm. #01:44:56#

Q: (...) #01:44:59#

R: Es ist sehr mühselig. Und so viel Geld, was auch weggeht. Und, (...) Also, wir haben jetzt hier bei uns am Institut, also am *[Institut]*, *[Institutsleitlinie]*. 01:45:26#

Q: (...) #01:45:26#

R: Wir nutzen einfach offene Daten. #01:45:29#

Q: (...) #01:45:32#

R: Mhm. Und für die funktioniert das sehr gut bisher. #01:45:35#

Q: (...) #01:45:58#

R: Ja. #01:45:58#

Q: (...) #01:46:17#

R: Mhm. Ja, da wäre halt die Hoffnung, dass je mehr Leute das machen, dass das vielleicht auch später leichter möglich ist. Ich weiß es nicht. Dafür bin ich einfach zu weit weg vom Feld. Was ich halt als gutes Beispiel sehe, ist die *[Datenbank]*. Da, also, kennen Sie die? #01:46:41#

Q: (...) #01:46:41#

R: *[Datenbank]* ist eine riesig angelegte Studie aus *[Land]*. Und da ist alles drinnen von *[Teildisziplin1]*, im Prinzip alle klassischen Fragebögen, über *[Teildisziplin2]*. Da ist genetische, Genetikinformation drin, da ist Emotion drin, Entwicklung, alle möglichen anderen Diagnosen. Und, von *[Stichprobengröße]*. Wenn man sich da, ja, wenn man sich da bewirbt drauf, kriegt man einen bestimmten Zugang. Da gibt es natürlich gewissen Rechte, wie man das Nutzen darf und für was. Aber, wenn man da den Zugang hat, dann kann man, glaube ich, erstmal gut ein paar Sachen machen. #01:47:39#

Q: (...) #01:48:08#

R: Ja. #01:48:09#

Q: (...) #01:48:23#

R: Nee, also, wirklich die *[Datenbank]* ist momentan, glaube ich, am meisten mit vorne. #01:48:30#

Q: (...) #01:48:38#

R: Für die Sachen, die ich brauche, ja. Ich weiß natürlich nicht, wie es ist (...) ich habe mir nicht alle Daten angeguckt oder Variablen, aber die haben schon eine sehr gute Dokumentation. Man kann auch diesen *data showcase* von denen nutzen und nach zum Beispiel, zum Beispiel einfach mal suchen. Was es hier so gibt. Und da gibt's halt eben diese verschiedenen Subbereiche. Zum Beispiel [*Datenbankabfrage1*], ist halt, soweit ich weiß, glaube ich noch nicht alles da. Bin mir aber auch nicht sicher. Aber genau, zum Beispiel, (...) [*Datenbankabfrage2*] #01:49:48#

Q: (...) #01:49:49#

R: Und so ein paar Sachen. Mhm. Also, da kann man sich durchgucken und dann kriegt man dazu die Erklärungen, und das Feld und alles. #01:50:00#

Q: (...) #01:50:03#

R: Und jetzt hat es auch hier eins fertig geschafft. (lacht) So, und das ist, das ist ein Notebook. Das ist wie es auch sonst aussieht. Und ich kann das jetzt auch einfach durchführen mit Daten zur Verfügung gestellt und kann jetzt hier durchgehen. Und, jetzt lade ich mir grade hier noch ein anderes Dataset runter. Kann das aber alles ausführen. Und, all das hier passiert aber dezentral auf Servern, die ich nicht bedienen muss. Das heißt, all diese ganze Rechenkapazität, das alles passiert nicht auf meinem Rechner. Das, was ich mache, ist das Aufmachen und Starten. Und, das, ja... #01:50:46#

Q: (...) #01:50:50#

R: Genau. Das wäre der letzte Punkt, dann muss ich tatsächlich auch leider los. Und ich will Sie jetzt hier auch nicht zu lange vom Feierabend abhalten. #01:51:01#

Q: (...) #01:51:02#

R: Hier bin ich natürlich jetzt ein bisschen weiter weg von den ganzen [*Gesetzestext*] Sachen. Da muss ich mich grade nicht so mit beschäftigen, aber ich habe es noch so ein bisschen gemacht. Wir haben jetzt grade einen Preprint rausgebracht über [*Teildisziplin1*] „*data sharing*“ unter [*Gesetzestext*]. Das war die letzten, naja, [*Zeitraum*], seitdem das rausgekommen ist, haben wir halt

versucht alles rauszufinden über *[Gesetzestext]* und die Sachen, die ganzen *Legal Terms*, alles was drinnen steht, und haben dann so eine Art Guideline geschrieben. Und, mal schauen, ob ich es grade (-) Genau, wir haben da so eine Seite dafür. Und, was die Idee ist, wir haben hier so eine *[Gesetzestext]* Edition, wo wir Sachen erklären mit so bestimmten ganzheitlichen Problemen, was alles so, was man beachten muss und wie man das machen kann. Und, was wir dann haben ist so eine Art, so ein Template, die unter freier Lizenz ist. Und, das könnte man sich einfach kopieren und in den Platzhaltern ihre Sachen einsetzen. Und die ganze Idee davon ist eben, das steht, glaube ich, auf der Eingangs-Website (-) Ja, *[Überschrift]*“. Sondern, dass man halt dann als Researcher schon diese ganzen Infos hat und Dokumente, wo das alles kontrolliert drinnen ist und das nehmen kann, ersetzt die spezifischen Inhalte, was notwendig ist, und reicht das dann ein. Das heißt, wir haben zum Beispiel jetzt auch diese *[Inhalt]* Geschichten und sowas alles drinnen. Und, haben das übersetzt halt in ein paar Sprachen. #01:53:18#

Q: (...) #01:53:19#

R: Und dann gibt es noch so ein *[Vorlage]*, was wir halt basierend gemacht haben auf den Erfahrungen von uns allen, die da mitgemacht haben und die Daten geteilt haben. Und, das wäre halt so ein Punkt, wenn man das anstrebt, was Sie vorhaben, was nicht, was man nicht ignorieren kann. Ne, also, diesen, diesen Aspekt von *data sharing*. #01:53:57#

Q: (...) #01:54:07#

R: Ja. Das ist eben das. Man kann sich am Ende noch so viel Mühe geben. Wenn dieser erste Punkt fehlt, alles vorbei. Und, was wir herausgefunden haben ist, dass unter *[Gesetzestext]*, das ist eine lustige Geschichte. Die EU sagt ja mit Horizon 2020, dass Daten wieder, Daten offen machen, alles wiederverwenden, *Open Data*, und wiedernutzen, was natürlich top ist, aber gleichzeitig haben sie *[Gesetzestext]* gemacht, was keine Datenteilung erlaubt eigentlich. Und, bei uns sind die Sachen, *[Teildisziplin1]* ist immer personalisierte Daten, personenbezogenen Daten, und die sind quasi nie anonymisierbar, daher auch quasi nie teilbar und da wird es halt schnell haarig. Und, bisher ist der Fall, wir haben das gemacht und viele Leute nehmen das, manche aber auch nicht. So hat man teilweise in der gleichen Uni, aber verschiedenen Departments, unterschiedliche *[Beruf]*, die unterschiedliche Sachen sagen. Das heißt, der eine sagt, eine Person sagt „Ja, ist okay“ und die andere „auf gar keinen Fall“. Und, wir sind mittlerweile der Meinung, es muss leider anscheinend erstmal irgendwer verklagt werden, damit wir wissen, wie es funktioniert. #01:55:51#

Q: (...) #01:55:56#

R: Also, wir haben für diese Sachen tatsächlich mit [Land] Anwälten und so gesprochen und auch die konnten keine spezifischen Sachen sagen, dass das immer funktioniert und das nicht. #01:56:11#

Q: (...) #01:56:17#

R: Und... #01:56:17#

Q: (...) #01:56:19#

R: Ganz genau. Und da ist auch der Punkt, das was bei uns zum Beispiel wichtig war, wir brauchen so ein *Data User Agreement*. Das ist hiermit verlinkt. Weil selbst, wenn man die Daten offen zur Verfügung stellt, es muss immer auf Servern innerhalb der EU passieren und es braucht immer ein *Data User Agreement*, dass, wenn man die Daten nimmt, dass die Daten nicht irgendwo anders quasi dann später gelagert werden. Weil im Prinzip dürfen die Daten auch nie außerhalb, aus der EU raus. #01:57:04#

Q: (...) #01:57:10#

R: Und, deswegen ist dieses *Data User Agreement*. Es ist halt nicht bindend, auf keinen Fall. #01:57:15#

Q: (...) #01:57:16#

R: Aber, wenn man sagt, die Leute, wenn die die Daten nehmen von uns, von unseren europäischen Servern und verarbeiten, legen sie die gleichen Nutzungsbedingungen wieder drauf. Das heißt, man gibt quasi die, die böse Karte ab so ungefähr, ne, an die (lacht) #01:57:40#

Q: (...) #01:57:42#

R: Ja. #01:57:42#

Q: (...) #01:57:44#

R: Mhm. Und dann weiß ich halt nicht, wie das bei Ihrem *Repository* ist? #01:57:50#

Q: (...) #01:57:53#

R: Genau so. Super. Dann funktioniert das ja schon mal an der Stelle (lacht) #01:57:56#

Q: (...) #01:58:18#

R: Ja. #01:58:20#

Q: (...) #01:58:27#

R: Ganz genau. Das ist (...), was wir gerade machen. #01:58:29#

Q: (...) #01:58:49#

R: Ja. (lacht) #01:58:52#

Q: (...) #01:58:54#

R: Ja, das ist (-) alles so ein bisschen schwierig. #01:58:58#

Q: (...) #01:58:59#

R: Leider so. #01:59:00#

Q: (...) #01:59:09#

R: Oh ja, doch so. Gut. #01:59:13#

Q: (...) #01:59:17#

R: Ja. Ja. #01:59:20#

Q: (...) #01:59:20#

R: Ja. #01:59:21#

Q: (...) #01:59:42#

R: Doch ja, ich finde das gut, weil es von dem Theoretischen zum Praktischen geht. Und, wenn die Leute das hoffentlich nutzen, auch die notwendigen Informationen enthalten soll. #02:00:00#

Q: (...) #02:00:13#

R: Wenn ich jetzt einmal noch draufklicken kann? #02:00:16#

Q: (...) #02:00:20#

R: Okay. (lacht) Okay. Der Rohdatensatz wäre jetzt eine jede Versuchsperson. #02:00:45#

Q: (...) #02:00:46#

R: (...) Nachdem, man sich jetzt das *Codebook* ein paar Mal angeguckt hat, ist das, denke ich, okay. #02:01:17#

Q: (...) #02:01:18#

R: Also, man kann auf jeden Fall (-) Ich denke, man müsste beides zusammen aufhaben, parallel halt natürlich, um nochmal dann genau nachzugucken, was, was ist. Aber das, (...) alles in eins reinschreiben funktioniert ja nicht. #02:01:39#

Q: (...) #02:01:52#

R: Ja. #02:01:53#

Q: (...) #02:02:11#

R: Nee. #02:02:11#

Q: (...) #02:02:13#

R: (lacht) Genau. Genau. (...) Ja, aber ich denke, mit drei ist man, hat man einen guten, eine gute Implementation (...) #02:02:22#

Q: (...) #02:02:48#

R: Ja. Ich versuche es jetzt mal aus der Perspektive einer primär psychologischen Forschungsrichtung, im Sinne von ich würde jetzt nicht [*Dokumentationsstandard 2*] nutzen und [*Teildisziplin1*] machen. Ich denke, wenn man eine Dokumentation dafür bereitstellt und je nachdem, wie der das vielleicht möglich ist bei Ihnen einer, ein Tutorial, vielleicht auch so, so ein *Video Lectures* dazu hätte, wo man das dann so einen paar da... #02:03:30#

Q: (...) #02:03:30#

R: Ah, okay. Wo man an so ein paar Datensätzen das erklärt und durchgeht, fände ich das sehr nützlich und würde das auf jeden Fall natürlich vorziehen über da drüber überhaupt keinen Standard zu haben. Ich finde, das ist gut. Es ist möglich umzusetzen für Leute die sich damit noch nicht beschäftigt haben und denen dann ein Einblick zu geben, oder denen einen Anfangspunkt zu geben. #02:04:05#

Q: (...) #02:04:06#

R: Und, ich denke, wenn das über Ihre Institut / Gesellschaft funktioniert, ist das, glaube ich, auch ein gutes, ein guter Weg darzustellen, dass das was ist aus der, aus der Community. Wo viele Leute dabei sind und dass viele Leute das unterstützen und adaptieren. Halt damit geht ja auch wieder der Nutzen einher. Und ich denke dafür ist es sehr, sehr gut. Und die, die Problematik besteht halt wie bei jeder dieser Initiativen darin, die Leute, die nicht wirklich daran interessiert sind und nicht die große Lust darauf haben oder vielleicht auch nicht nachvollziehen können, warum das nützlich ist, die in das Boot zu holen mit. Und, ich denke, da ist eben diese Tutorials und über E-Mail-Verteiler, oder falls mal irgendwann wieder möglich (...) Tagungen, das, das entsprechend an die Leute ran zu bringen. #02:05:26#

Q: (...) #02:05:33#

R: Ja. #02:05:35#

Q: (...) #02:05:40#

R: Ja. #02:05:41#

Q: (...) #02:05:59#

R: Ja. Nee, das auf jeden Fall. Ein sehr wichtiger Punkt. Weil sonst bekommt man immer diese Argumentationsgeschichten rein, wo man selber sagt „das ist nützlich und gut“, aus all diesen Gründen. Aber diese Gründe sind für die Leute nur schwierig nachzuvollziehen. Wenn man aber sagt „pass mal auf, der Standard hier, so und so viel Prozent bessere Wiedernutzbarkeit“... #02:06:28#

Q: #02:06:28#

R: ...oder Zustimmung oder eine andere gute Variable, dann kann man die Leute auf jeden Fall auch hoffentlich gut mit Daten überzeugen. #02:06:39#

Q: (...) #02:06:52#

R: (lacht) Ja. #02:06:53#

Q: (...) #02:06:55#

R: Oh, ja. #02:06:55#

Q: (...) #02:07:13#

R: Die, ich würde sie auf jeden Fall drinnen lassen. Ich finde es immer angenehm, wenn man dann nochmal so eine grafische Repräsentation hat. Ich meine, meine Idee wäre halt, mein Wunsch (lacht) was ich halt vorher erwähnt hatte, nämlich, dass man das vielleicht so ein bisschen mehr *interoperable* macht. Mit dem *Codebooks* so ein bisschen mehr zusammenbringt und so ein bisschen mehr Infos (-) Genau. #02:07:48#

Q: (...) #02:07:57#

R: Genau, genau. Ja. #02:07:57#

Q: (...) #02:08:07#

R: (...) #02:08:07#

Q: (...) #02:08:10#

R: Nee, vielen Dank Ihnen für die Anfrage! #02:08:13#

Q: (...) #02:08:15#

R: Tut mir leid, dass ich jetzt so viel rum, rumerzählt habe. (lacht) #02:08:17#

Q: (...) #02:08:22#

R: Okay. #02:08:24#

Q: (...) #02:08:37#

R: Okay. Dann ist gut. #02:08:37#

Q: (...) #02:08:50#

R: Ja. #02:08:50#

Q: (...) #02:08:52#

R: Nee, das ist wirklich eins der Probleme, dass, dass viele Leute haben vergleichbare Ideen, aber anstatt zu versuchen sich so zusammenzufinden oder zu gucken, was schon los ist, macht jeder nochmal so, nochmal so das Eigene. Und, dann hat man diesen Urwald an Möglichkeiten und weiß überhaupt nicht, wo man anfangen soll. #02:09:13#

