

From Schedules of Activities (SoA) to USDM: Automating Protocol Extraction Using Large Language Models

Kerstin Forsberg
Johannes Ulander

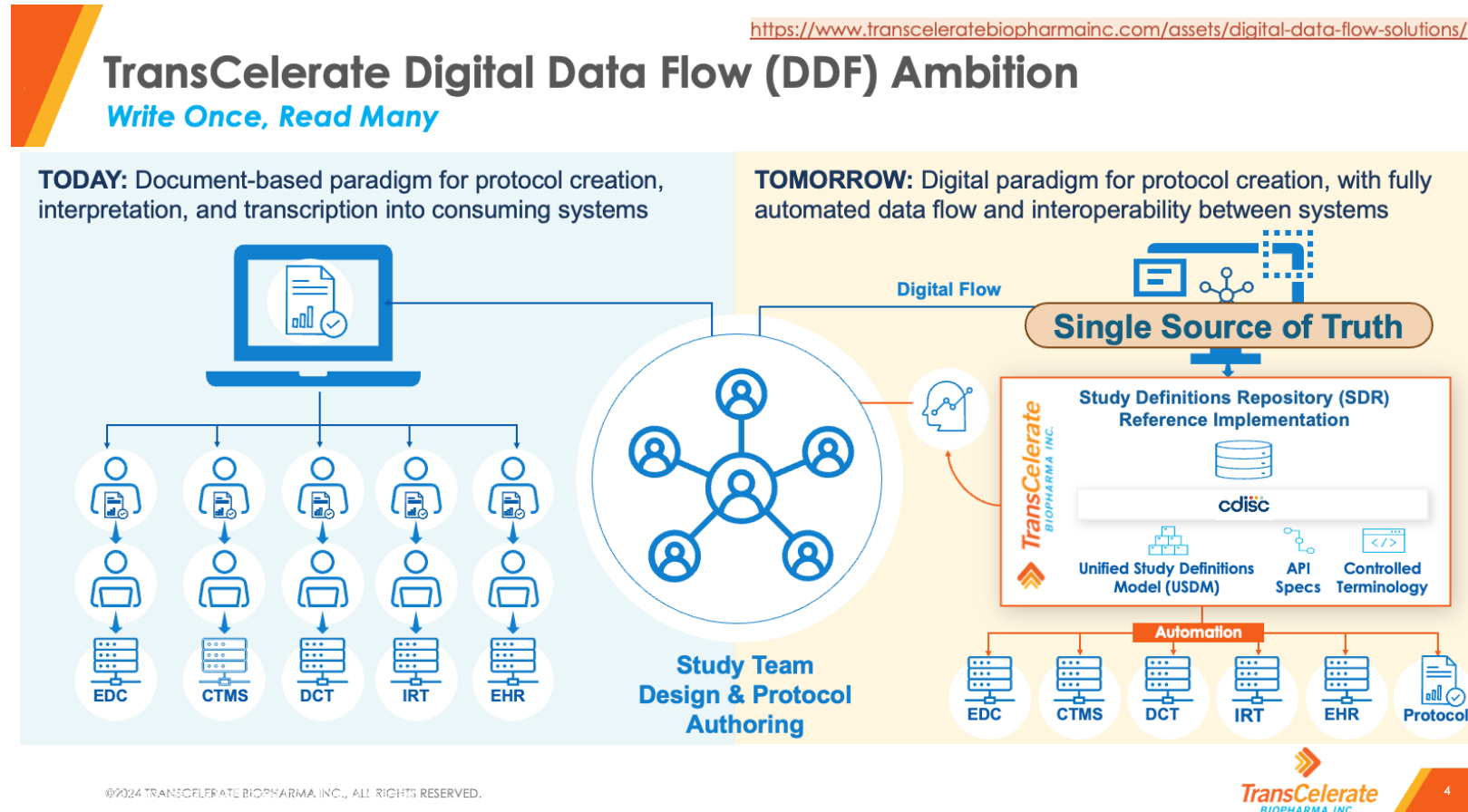
Disclosure

- Dave Iberson-Hurst is currently working on contract to CDISC as the CDISC USDM Product Owner
- The views expressed during this presentation are d4k's and **NOT** those of CDISC
- You will see some CDISC slides. They were all created by DIH
- There is a Transcelerate slide within the presentation. This has been presented within public forums.

Agenda

- Big Picture: From document-based to data-driven processes
 - What is USDM?
 - Why SoA table matters?
 - Why is it so hard to extract data from SoA tables?
- Data extraction with Multimodal LLM
 - ✳ From Anthropic's Claude 3.5 to Claude 4.5
 - From a Multi-pass (3->5->7) Pipeline ...
 - ... to **Soa2USDM**, a Conversational approach
- Learnings

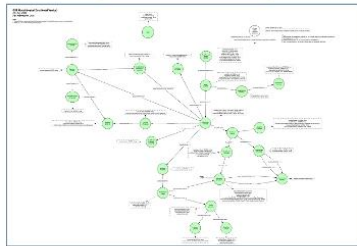
Big Picture: From document-based to data-driven processes



Unified Study Definitions Model (USDM)

CDISC DDF / USDM, Phases One to Four

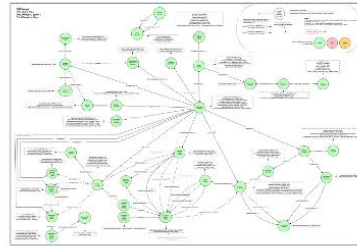
Phase One



25 Classes

- Solid foundation
- The protocol document was an external entity into which the structured content could be exported

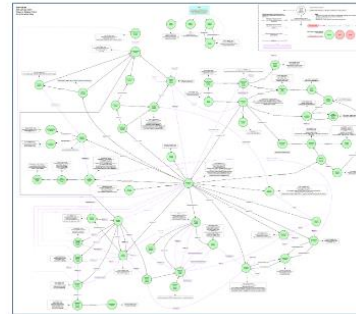
Phase Two



35 Classes

- Focused on the structured elements of the protocol, e.g. the Schedule of Activities (SoA) & Biomedical Concepts (BCs)
- The protocol document still an external entity

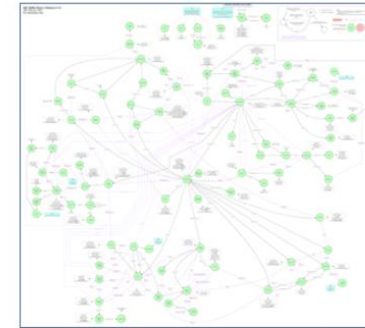
Phase Three



58 Classes

- Now contains structured and unstructured elements
- The entire protocol document can be held within the USDM
- Allows for the protocol document to be generated from the model

Phase Four

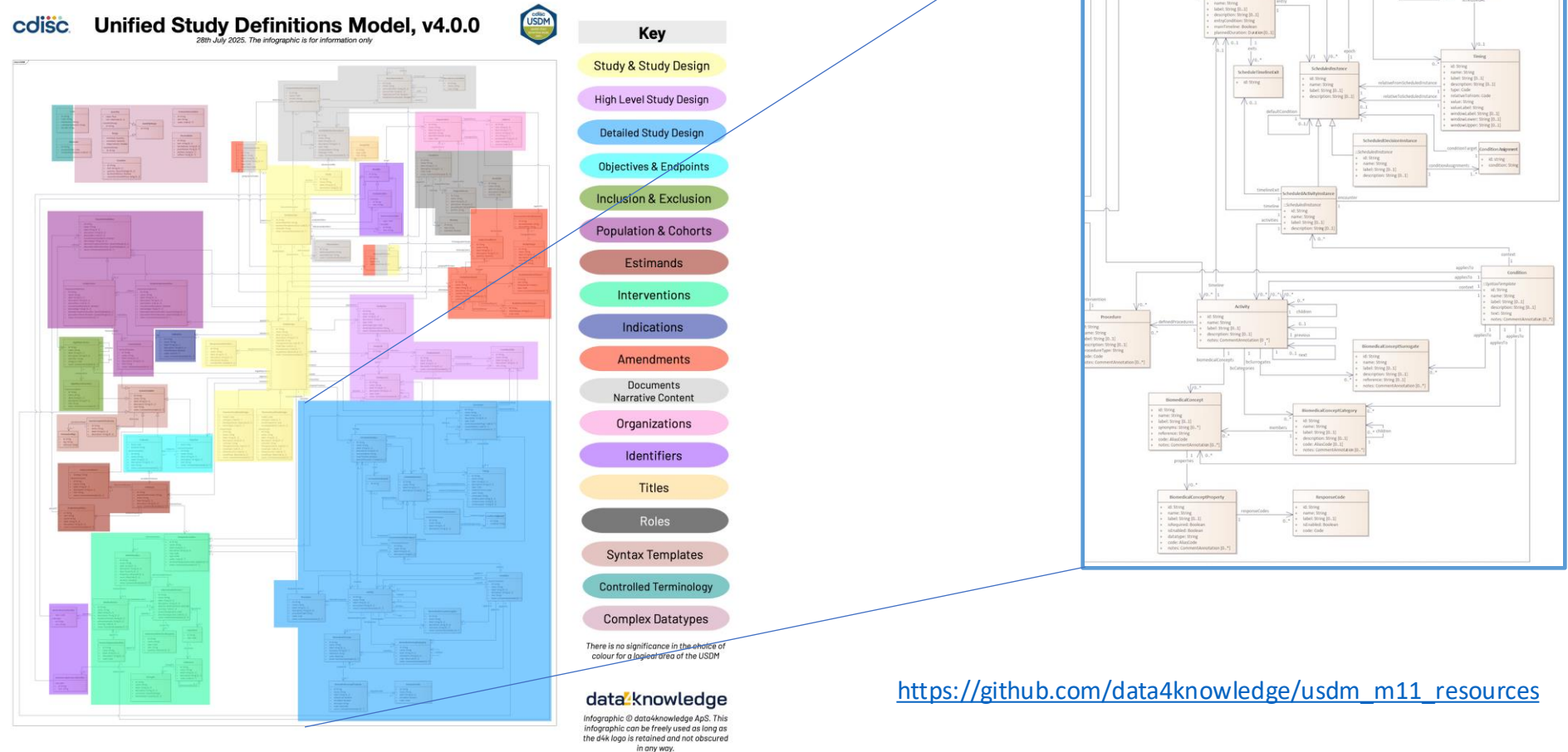


89 Classes

- Aligned with ICH M11
- Support for complex studies, interventional & observational studies, and medical devices
- Maximise content re-use and support for multiple document templates
- Extension mechanism to provide flexibility



USDM Areas, focus on the Detailed Study Design



https://github.com/data4knowledge/usdm_m11_resources

How do Schedule of Activities (SoA) tables typically look like?

CDISC_Study_Pilot

	VISIT	1	2	3	4	5	7	8
ACTIVITY	WEEK	-2	-3	0	2	4	6	8
Informed consent	X							
Patient number assigned	X							
Hachinski i-4	X							
MMSE 10-23	X							
Physical examination	X							
Medical History	X							
Habits	X							
Chest x-ray	X							
Apo E genotyping				X				
Patient randomized				X				
Vital signs/Temperature	X	X	X					
Ambulatory ECG placed		X						
Ambulatory ECG removed			X					
ECG	X							
Placebo TTS test	X							
CT Scan (if not within last year and patient passes all other screens)	X							
Concomitant Medications	X			X				
Laboratory (Chem/Hemat)	X							
Laboratory (Urinalysis)	X							
Plasma Specimen (Xanomeline)			X					
Hemoglobin A _{1c}	X ^a							
Study drug record				X				
Medications dispensed								
Medications returned								
TTS Acceptability Survey								
ADAS-Cog	P		X					
CIBIC+	P		X					
DAD	P		X					
NPI-X	P		X					
Adverse events	X	X	X					

ACTIVITY	VISIT	9	10	11	12	13	ET	RT
WEEK	12	16	20	24	26			
Informed consent								
Patient number assigned								
Hachinski i-4								
MMSE 10-23								
Physical examination					X	X		
Medical History								
Habits								
Chest x-ray								
Apo E genotyping								
Patient randomized								
Vital signs/Temperature	X	X	X	X	X	X	X	X
Ambulatory ECG placed								
Ambulatory ECG removed								
ECG	X	X	X	X	X	X		
Placebo TTS test								
CT Scan (if not within last year and patient passes all other screens)								
Concomitant Medications	X	X	X	X	X	X	X	X
Laboratory (Chem/Hemat)	X	X	X	X	X	X		
Laboratory (Urinalysis)	X							
Plasma Specimen (Xanomeline)	X						X	
Hemoglobin A _{1c}								
Study drug record	X	X	X	X	X	X	X	
Medications dispensed								
Medications returned								
TTS Acceptability Survey						X	X	
ADAS-Cog		X			X		X	X
CIBIC+		X			X		X	X
DAD		X			X		X	X
NPI-X		X ^b	X ^b	X ^b	X	X	X	X
Adverse events	X	X	X	X	X	X	X	X

Table 1: Schedule of Events

Visit	Screening	Randomization ^a	Treatment Period			Final Visit/ET ^a
			1 st eligible HAE attack ^a	2 nd eligible HAE attack ^a	3 rd eligible HAE attack ^a	
In-clinic ^d	X	X	X	X	X	X
TeleVisit ^e						
Informed Consent/Assent ^f	X					
Eligibility Assessment	X	X				
HAE Diagnostic Lab Test ^g	X					
Medical History ^h	X					
12-lead ECG ⁱ	X					X
Demographics ^j	X					
Physical Examination ^k	X					X
Vital Signs ^l	X					X
Safety Laboratory ^m	X					X
Pregnancy Test ⁿ	X					X
Randomize patient in RTSA						
Patient diary training						
Patient diary retraining						
Conventional on-demand treatment						
Adverse Event Review ^o						
Dose ^p						
IMP dispensing/return/accou						
PGI-S ^q						
PGI-C ^r						
VAS ^s						
GA-NRS ^t						
Concomitant Medication Re						

Abbreviations: C1-INH-C1-esterase inhibitor; ECG=electrocardiogram; eDiary=electronic diary; ET=Early Termination; GA-NRS= General Anxiety - Numeric Rating Scale; HAE=hereditary angioedema; IMP=investigational medical product; PGI-C=patient global impression of change; PGI-S=patient global impression of severity; RTSA=Randomization Trial Supply Management System; VAS=visual analog scale.

a Randomization to occur within 4 weeks of the Screening Visit. The Randomization Visit may be a televisit or an in-clinic visit.

b For each of the 3 eligible HAE attacks, a televisit will occur by the next working day after the completion of the patient diary (visit window: +1 week). A televisit must occur before treating the next eligible attack.

c Once 3 HAE attacks have been treated with IMP, or upon patient ET, the patient should return to the clinic for the final visit as soon as possible within 1 week of completing all patient diary assessments. Whenever possible this visit should be completed prior to starting any new medication or treatment.

d If an in-clinic visit cannot be conducted (e.g. in the event of a pandemic or other reason that prevents the patient from attending in-clinic visits), home health visits will be used to perform these visits. Information captured during a home health visit will mirror that captured during an in-clinic visit.

e A televisit can be conducted via a telephone call or via an interactive audio/video system.

f Consent and assent may be collected through a consent if allowed through country and site regulations.

g Diagnostic testing must be available for patients without genetic confirmation of HAE type I or II.

h Medical history includes any relevant previous and concurrent diseases. HAE disease history includes the attack history and corresponding treatment for the past 3 months; therapies and supplements taken within the past 4 weeks; and previous participation in interventional clinical trials in the past 4 weeks.

i ECG to be recorded after patient has rested supine for at least 5 minutes.

j Demographics: Year of birth, height (meters [m]; without shoes), weight (kilograms [kg]), race and ethnicity (if allowed), and sex.

k Complete physical examination at Screening only; all others will be symptom directed. In the case of a home health visit, the home healthcare nurse will conduct an abbreviated physical examination, and the Investigator will conduct the symptom-directed physical examination via televisit.

l Vital signs include pulse rate, respiratory rate, and systolic and diastolic blood pressure after patient has been resting at least 5 minutes.

m Laboratory assessments, including C1-INH functional level, performed by a central laboratory; repeat laboratory assessments may be performed.

n Serum pregnancy test will be performed on females of childbearing potential at Screening and at the Final/ET visit.

o If a patient takes conventional on-demand treatment during the trial, a 48-hour washout period is required prior to the subsequent dosing with IMP.

p Eligible attacks should be treated with IMP as soon as possible after recognition of the start of the attack. A minimum 48-hour washout period is required between each eligible attack and last dose of IMP or conventional on-demand treatment. After the first dose of IMP and prior to the second dose or conventional on-demand treatment, patients will be required to call a designated Study Call Center to be reminded of the repeat dosing criteria. The Call Center staff will remind patients of the rules for re-dosing and the eDiary assessment requirements.

q Arrangements will be made to dispense the assigned IMP. The IMP will be shipped to the patient via a courier service or will be dispensed at the trial clinic as required by local regulations or per the site's local practice, as described in the Pharmacy Manual. Confirmation of dosing compliance will be performed at each of the HAE attack televisits. All unused IMP and packaging will be returned at the in-clinic Final Visit/ET visit. If the Final Visit/ET visit is performed by Home Health, arrangements will be made to return any unused IMP and packaging via courier.

r Patients will complete limited assessments of HAE attack symptoms through 48 hours following first dose of IMP. Patient assessments will be performed as outlined in Table S1.

s Adverse events recorded from the first dose of IVIG/05259917 up to and including the Final Visit/ET. Serious AEs recorded from the time of signing the informed consent up to and including the Final Visit/ET.

KaVista_NCT05259917_Angioedema

3 of 19 protocols from
https://github.com/data4knowledge/usdm_data/tree/main/source_data/protocols

Alexion_NCT04573309_Wilsons

Study Procedures	Screening ^a	C-1 ^b																	UNS ^c	EOS or ET ^d			
Days	Screening -42 to -9	C-1 -21	Inpatient Period 1							OP ^e	Inpatient Period 2										EOS Day 54/+2		
Eligibility	-42 to -9	-21	-7	-6 through -5	-4 through -1	1	2-3	4-7	8	9	10-22	23	24	25	26-28	29	30-35	36	37-38	39	40	EOS Day 54/+2	
Informed consent	X																						
Admit to unit ^f												X									X		
Discharge from unit ^f											X												
Outpatient visit or phone call											X ^g												
Inclusion/exclusion	X	X									X												X
Discuss/document	X	X																					X
Contraception	X																						
Follicle-stimulating hormone (post-menopausal females only) ^h	X																						
Alcohol test	X																						
Urine drug screen	X																						
ILV, hepatitis B and C screen	X																						
Study Administration	X																						
Medical history/demographics ⁱ	X																						
WD history ^j	X																						
Post WD treatment ^k	X																						
Physical examination ^l	X																						
Height, weight, and BMI	X																						
Enrollment																							
Enrollment/inclusion	X	X										X	X	X ^m	X	X	X	X	X	X	X	X	X
Discontinue chelation therapy	X	X										X	X	X ^m	X	X	X	X	X	X	X	X	X
Discontinue zinc therapy	X	X										X	X	X ^m	X	X	X	X	X	X	X	X	X
Urine serum pregnancy test ⁿ	X	X										X	X	X ^m	X	X	X	X	X	X	X	X	X
Random serum sample (a/total)	X	X										X	X	X ^m	X	X	X	X	X	X	X	X	X
Vital sign measurements ^o	X	X										X	X	X ^m	X	X	X	X	X	X	X	X	X
12-lead ECG (implants)	X	X										X	X	X ^m	X	X	X	X	X	X	X	X	X
Adverse event	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Balance assessments																							
Cu/Mo-controlled meals ^p	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Light exercise regimen	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Urination and bowel movement monitoring, menstruation check ^q												X	X	X	X	X	X	X	X	X	X	X	X
24-hour urine for Cu and Mo ^r												X	X	X	X	X	X	X	X	X	X	X	X
Urine for Cu and Mo ^s												X	X	X	X	X	X	X	X	X	X	X	X
Other																							
Concomitant medication and non-pharmacologic therapy/procedure	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

^a Within 42 days of ALXN1840 administration. Details of procedures that may be performed by sites designated as "screening sites" (only in the US) are detailed in Section 5.

^b Participants will be admitted to the clinical research unit (CRU) at least 10 hours prior to enrollment and initiation of copper- and molybdenum-controlled diet.

^c At the CRU's discretion, participants may remain in the CRU or be readmitted on Day 22 with all procedures starting on Day 23.

^d Unscheduled study visits may occur at any time during the study and may include any study procedure as deemed necessary by the Investigator.

^e Participants are required to return to the CRU Day 54/+2; participants will be discontinued from ALXN1840 medication by their physician at the latest by the Day 54 follow-up.

^f Discharge from the unit may occur after completion of all procedures on Day 9 and on Day 40 and after the Investigator has reviewed all safety assessment (including safety laboratory tests) and confirmed that the participant is appropriate for discharge.

^g A single outpatient visit or phone call and safety laboratory assessment should occur between Day 14 and Day 18. A phone call may take place on a different day than the blood draw within the Day 14 through Day 18 period.

^h If needed to confirm menopause.

Why SoA tables matters?

- Primary bridge between scientific objectives and operational reality
- Primary specification source for Electronic Data Capture (EDC) configuration
- Encode multiple information layers through spatial relationships, hierarchical structures, and footnote-modified semantics that determine how trials actually execute

SoA pages are approx. 3 - 20+ of 100-200 pages of a Clinical Study Protocol (CSP)
Together with the “schema visualisation”, these are the visual components of a CSP
Detailed Study Design

Why is it so hard to extract data from SoA tables?

- Representing multidimensional operational requirements in two-dimensional tabular formats
- Operational knowledge gets encoded in human-optimized document formats

- “Junction area” *)
- Multilevel headers
- Merged cells
- Continuation markers
- Rows divided over pages
- ...

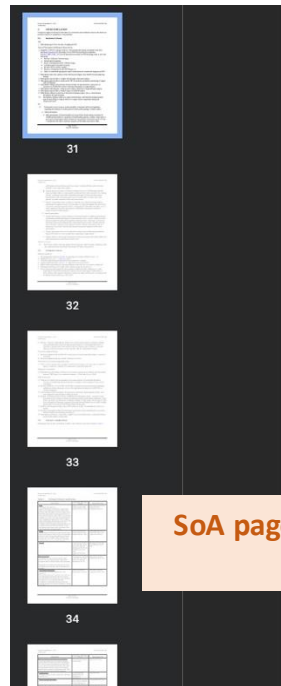
Procedure	Screening	Treatment			
		Day -1 or Day 1	Day 1	Day 3	Day 6
Window (days)		NA	± 1	± 1	± 1
P = In-Person Visit H = Home Visit R = Remote Visit	P	P	P/H*	P/H*	I
Return to usual (pre-COVID-19) health			X	X	?
Household infection report		X			
Study staff-collected mid-turbinate nasal swab for SARS-CoV-2 RT-PCR		✓	✓	✓	

*) Recommended reading by Mark Kramer, MITRE

- Barriers to Automatic Conversion of Clinical Trial Schedules of Activities to Structured Data [White Paper]
- I Tested 12 "Best-in-Class" PDF Table Extraction Tools, and the Results Were Appalling [Medium blog post]

Extract structured content from CSP text using LLM works very well

Text capabilities is enough



5.1. Inclusion Criteria

Age

1. Participants aged ≥ 18 at the time of signing the ICF.

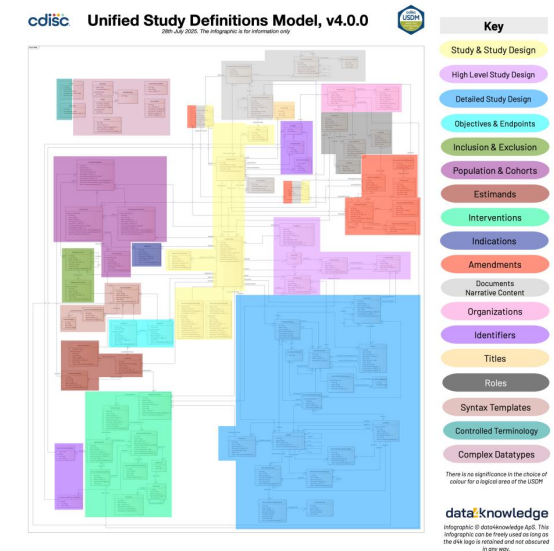
Type of Participants and Disease Characteristics

2. Diagnosis of WD by Leipzig Criteria ≥ 4 documented by testing as outlined in the 2012 European Association for the Study of Liver WD Clinical Practice Guidelines (Ferenci, 2003; EASL, 2012) or by historical test results for WD including some or all of the following:
 - Presence of Kayser Fleischer rings,
 - Neurological symptoms,
 - Serum ceruloplasmin below reference range,
 - Coombs-negative hemolytic anemia,
 - Elevated liver or urinary copper,
 - Presence of mutations in the ATP7B gene, or
 - Other, as considered appropriate, may be used instead to confirm the diagnosis of WD.
3. Participants who in the opinion of the referring Investigator may benefit from decoppering therapy.
4. Participants must be able to comply with all study-related procedures.
5. Participants must be able to reside in the CRU for intensive metabolic monitoring of copper and molybdenum.
6. Participants willing to discontinue chelator therapy for approximately 4 days prior to initiation of ALXN1840 to allow a baseline assessment of copper balance.
7. Participants with adequate venous access to allow collection of required blood samples.
8. Participants must be able to swallow intact ALXN1840 tablets.
9. Participants willing to avoid use of minerals containing copper, zinc, or molybdenum throughout the study duration.
10. Participants willing to adhere to copper/molybdenum-controlled diet during inpatient periods and willing to comply with a low copper dietary requirement during the Outpatient Period.

SoA pages are approx. 3 - 20+ of 100-200 pages of a Clinical Study Protocol (CSP)
Maybe 95% of a CSP is pure text, or basic listings/tables

- Male participants, if heterosexually active and with a female spouse or partner of childbearing potential or a pregnant or breastfeeding spouse or partner, must agree to use barrier contraception (male condom) for the duration of the study and for at least 3 months after the end of systemic exposure of the study intervention. Male

Page 31 of 74



Multimodal LLM, vision capabilities Timeline for Claude

Text capabilities is enough

5.1. Inclusion Criteria

Age

- Participants aged ≥ 18 at the time of signing the ICF.

Type of Participants and Disease Characteristics

- Diagnosis of WD by Leipzig Criteria ≥ 4 documented by testing as outlined in the 2012 European Association for the Study of Liver WD Clinical Practice Guidelines (Ferenci, 2003; EASL, 2012) or by historical test results for WD including some or all of the following:
 - Presence of Kayser Fleischer rings,
 - Neurological symptoms,
 - Serum ceruloplasmin below reference range,
 - Coombs-negative hemolytic anemia,
 - Elevated liver or urinary copper,
 - Presence of mutations in the ATP7B gene, or
 - Other, as considered appropriate, may be used instead to confirm the diagnosis of WD.

- Participants who in the opinion of the referring Investigator may benefit from decoppering therapy.
- Participants must be able to comply with all study-related procedures.
- Participants must be able to reside in the CRU for intensive metabolic monitoring of copper and molybdenum.
- Participants willing to discontinue chelator therapy for approximately 4 days prior to initiation of ALXN1840 to allow a baseline assessment of copper balance.
- Participants with adequate venous access to allow collection of required blood samples.
- Participants must be able to swallow intact ALXN1840 tablets.
- Participants willing to avoid use of minerals containing copper, zinc, or molybdenum throughout the study duration.
- Participants willing to adhere to copper/molybdenum-controlled diet during inpatient periods and willing to comply with a low copper dietary requirement during the Outpatient Period.

Sex

- Contraceptive use by men or women should be consistent with local regulations regarding the methods of contraception for those participating in clinical studies.
 - Male participants:
 - Male participants, if heterosexually active and with a female spouse or partner of childbearing potential or a pregnant or breastfeeding spouse or partner, must agree to use barrier contraception (male condom) for the duration of the study and for at least 3 months after the end of systemic exposure of the study intervention. Male

Page 31 of 74

VS

Vision capabilities needed

Study Procedures	Screening	C-4*																	UNS*	EOS or ET*				
Days	-42 to -9	-21	-8	-7	-6 through -5	-4 through -1	1	2-3	4-7	8	9	10-22	23	24	25	26-28	29	30-35	36	37-38	39	40	EOS Day 54+/-2	
Eligibility																								
Informed consent	X																							
Admit to unit		X											X											
Discharge from unit ^a																					X			
Outpatient visit or phone call																								
Inclusion/exclusion	X	X					X					X ^b												
Discuss/document	X	X																						X
Contraception																								
Follicle-stimulating hormone (post-menopausal females only)	X																							
Alcohol test	X																							
Urine drug screen	X																							
HIV, hepatitis B and C screen	X																							
Study Administration	Screening	C-4*																	UNS*	EOS or ET*				
Days	-42 to -9	-21	-8	-7	-6 through -5	-4 through -1	1	2-3	4-7	8	9	10-22	23	24	25	26-28	29	30-35	36	37-38	39	40	EOS Day 54+/-2	
Medical history/demographics	X																							
WD history	X																							
Prior WD treatment ^c	X																							
Physical examinations ^d	X																							
Height, weight, and BMI	X																							
Enrollment	Screening	C-4*																	UNS*	EOS or ET*				
Days	-42 to -9	-21	-8	-7	-6 through -5	-4 through -1	1	2-3	4-7	8	9	10-22	23	24	25	26-28	29	30-35	36	37-38	39	40	EOS Day 54+/-2	
Chemistry ^e , hematology, Coagulation	X	X								X ^f		X	X ^g	X		X ^h	X	X ⁱ	X		X			
Urinalysis	X	X										X	X ^g	X		X ^h	X	X ⁱ	X		X			
Urine screen pregnancy test ^j	X	X										X	X ^g	X		X ^h	X	X ⁱ	X		X			
Retained serum sample (safety)	X						X																	
Vitals sign measurements ^k	X						X																	
12-lead ECG (Orthostatic)	X						X																	
Adverse events	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Safety Assessments / Laboratory Analyses	Screening	C-4*																	UNS*	EOS or ET*				
Days	-42 to -9	-21	-8	-7	-6 through -5	-4 through -1	1	2-3	4-7	8	9	10-22	23	24	25	26-28	29	30-35	36	37-38	39	40	EOS Day 54+/-2	
Cu-Mo-controlled meals ^l	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Light exercise regimen	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Urination and bowel movement monitoring, menstruation check ^m							X	X	X	X	X	X				X	X	X	X	X	X	X	X	X
24-hour urine for Cu and Mo ⁿ							X	X	X	X	X	X				X	X	X	X	X	X	X	X	X
Feces for Cu and Mo ⁿ							X	X	X	X	X	X				X	X	X	X	X	X	X	X	X
Other	Screening	C-4*																	UNS*	EOS or ET*				
Days	-42 to -9	-21	-8	-7	-6 through -5	-4 through -1	1	2-3	4-7	8	9	10-22	23	24	25	26-28	29	30-35	36	37-38	39	40	EOS Day 54+/-2	
Concomitant medication and non-pharmacologic therapy/procedures	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

^a Within 42 days of ALXN1840 administration. Details of procedures that may be performed by sites designated as "screening sites" (only in the US) are detailed in Section 8.

^b Participants will be admitted to the clinical research unit (CRU) at least 10 hours prior to enrollment and initiation of copper- and molybdenum-controlled diet.

^c At the CRU's discretion, participants may remain in the CRU or be readmitted on Day 22 with all procedures starting on Day 23.

^d Unscheduled study visits may occur at any time during the study and may include any study procedure as deemed necessary by the Investigator.

^e Participants are required to return to the CRU Day 54+/-2; participants will be discharged from ALXN1840 medication by their physician at the latest by the Day 54 follow-up.

^f Discharge from the unit may occur after completion of all procedures on Day 9 and on Day 40 and after the Investigator has reviewed all safety assessment (including safety laboratory tests) and confirmed that the participant is appropriate for discharge.

^g A single outpatient visit or phone call and safety laboratory assessment should occur between Day 14 and Day 18. A phone call may take place on a different day than the blood draw within the Day 14 through Day 18 period.

^h If needed to confirm menstruation.

* Within 42 days of ALXN1840 administration. Details of procedures that may be performed by sites designated as "screening sites" (only in the US) are detailed in Section 8.

† Participants will be admitted to the clinical research unit (CRU) at least 10 hours prior to enrollment and initiation of copper- and molybdenum-controlled diet.

‡ At the CRU's discretion, participants may remain in the CRU or be readmitted on Day 22 with all procedures starting on Day 23.

§ Unscheduled study visits may occur at any time during the study and may include any study procedure as deemed necessary by the Investigator.

¶ Participants are required to return to the CRU Day 54+/-2; participants will be discontinued from ALXN1840 medication by their physician at the latest by the Day 54 follow-up.

‡ Discharge from the unit may occur after completion of all procedures on Day 9 and on Day 40 and after the Investigator has reviewed all safety assessment (including safety laboratory tests) and confirmed that the participant is appropriate for discharge.

§ A single outpatient visit or phone call and safety laboratory assessment should occur between Day 14 and Day 18. A phone call may take place on a different day than the blood draw within the Day 14 through Day 18 period.

¶ If needed to confirm menopause.

Claude 3.5 Sonnet (June 2024)

visual processing became advanced enough to parse mixed table layouts and scanned documents

Prompt caching

Claude 3.7 Sonnet (February 2025)

refined structural recognition for nested and relational tables

Batch API

Claude 4 family (May 2025)

introduced a sophisticated multimodal pipeline handling multi-page datasets with high grid-fidelity

Claude Code

Claude 4.5 Sonnet (September 2025)

integrated contextual linking—allowing the model to connect insights across tables, charts, and surrounding text for highly precise structured-data extraction

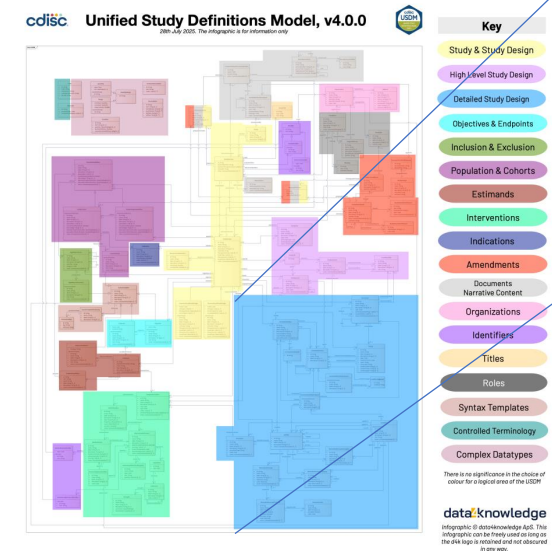
Excel file creation feature announced (Skills running behind the scenes)

Skills

Need multimodal LLM to extracte data prepared for conversion to USDM

	Screening ^a	C-1 ^b																				UNS ^d	EOS or ET ^e		
Study Procedures	Screening	C-1	Inpatient Period 1										OP ^f	Inpatient Period 2											
Days	-42 to -9	-21	-8	-7	-6 through -5	-4 through -1	1	2-3	4-7	8	9	10-22	23	24	25	26-28	29	30-35	36	37-38	39	40		EOS Day 54+/-2	
PD: Plasma total and PUF-Cu, LBC, ceruloplasmin, ceruloplasmin-bound Cu																									
Safety Assessments / Laboratory Analyses																									
Chemistry ^g , hematology, Coagulation	X		X			X ^h			X		X ⁱ	X		X ^j			X					X		X	
Urinalysis	X	X							X		X ^k	X		X								X		X	
Urine/serum pregnancy test ^l	X	X										X												X	
Retained serum sample (safety) ^m			X									X										X		X	
Vitals sign measurements ⁿ	X	X					X				X		X									X		X	
12-lead ECG (triplicate)	X	X	X			X	X		X		X	X	X	X								X		X	
Adverse events	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Balance assessments																								X	
Cu/Mo-controlled meals ^o			X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X			
Light exercise regimen			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			
Urination and bowel movement monitoring, menstruation check ^p						X	X	X	X	X				X	X	X	X	X	X	X	X	X			
24-hour urine for Cu and Mo ^q						X	X	X	X	X				X	X	X	X	X	X	X	X	X			
Feces for Cu and Mo ^r						X	X	X	X	X				X	X	X	X	X	X	X	X	X			
Other																									
Concomitant medication and non-pharmacologic therapy/procedure	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	

- ^a Within 42 days of ALXN1840 administration. Details of procedures that may be performed by sites designated as "screening sites" (only in the US) are detailed in Section 8.
- ^b Participants will be admitted to the clinical research unit (CRU) at least 10 hours prior to enrollment and initiation of copper- and molybdenum-controlled diet.
- ^c At the CRU's discretion, participants may remain in the CRU or be readmitted on Day 22 with all procedures starting on Day 23.
- ^d Unscheduled study visits may occur at any time during the study and may include any study procedure as deemed necessary by the Investigator.
- ^e Participants are required to return to the CRU Day 54+/-2; participants will be discontinued from ALXN1840 medication by their physician at the latest by the Day 54 follow-up.
- ^f Discharge from the unit may occur after completion of all procedures on Day 9 and on Day 40 and after the Investigator has reviewed all safety assessment (including safety laboratory tests) and confirmed that the participant is appropriate for discharge.
- ^g A single outpatient visit or phone call and safety laboratory assessment should occur between Day 14 and Day 18. A phone call may take place on a different day than the blood draw within the Day 14 through Day 18 period.
- ^h If needed to confirm menopause.



Claude 3.7 Sonnet (February 2025)

Claude 4 family (May 2025)

refined structural recognition for nested and relational tables

introduced a sophisticated multimodal pipeline handling multi-page datasets with high grid-fidelity

Need multimodal LLM to extract data prepared for conversion to USDM

Study Procedures	Screening ^a		C-1 ^b																						UNSP	EOS or ET ^c
	Screening	C-1			Inpatient Period 1					OP ^e					Inpatient Period 2											
Days	-42 to -9	-21	-8	-7	-6 through -5	-4 through -1	1	2-3	4-7	8	9	10-22	23	24	25	26-28	29	30-35	36	37-38	39	40			EOS Day 54 ^d /±2	
PD: Plasma total and PUF-Cu, LBC, ceruloplasmin, ceruloplasmin-bound Cu																										
Safety Assessments / Laboratory Analyses																										
Chemistry ^f , hematology, Coagulation	X	X	X			X ^g			X		X ^g	X				X ^g		X				X			X	
Urinalysis	X	X	X						X		X ^g	X				X						X			X	
Urine/serum pregnancy test ^h	X	X	X									X													X	
Retained serum sample (safety) ⁱ				X																						
Vitals sign measurements ^j	X	X	X				X				X		X									X			X	
12-lead ECG (triplicate)	X	X	X				X				X		X									X			X	
Adverse events	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Balance assessments																										
Cu/Mo-controlled meals ^k				X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X		
Light exercise regimen				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
Urination and bowel movement monitoring, menstruation check ^l						X	X	X	X	X					X	X	X	X	X	X	X	X	X	X		
24-hour urine for Cu and Mo ^m						X	X	X	X	X					X	X	X	X	X	X	X	X	X	X		
Feces for Cu and Mo ⁿ						X	X	X	X	X					X	X	X	X	X	X	X	X	X	X		
Other																										
Concomitant medication and non-pharmacologic therapy/procedure	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	

^a Within 42 days of ALXN1840 administration. Details of procedures that may be performed by sites designated as "screening sites" (only in the US) are detailed in Section 8.

^a Within 42 days of ALXN1840 administration. Details of procedures that may be performed by sites designated as "screening sites" (only in the US) are detailed in Section 8.

^b Participants will be admitted to the clinical research unit (CRU) at least 10 hours prior to enrollment and initiation of copper- and molybdenum-controlled diet.

^c At the CRU's discretion, participants may remain in the CRU or be readmitted on Day 22 with all procedures starting on Day 23.

^d Unscheduled study visits may occur at any time during the study and may include any study procedure as deemed necessary by the Investigator.

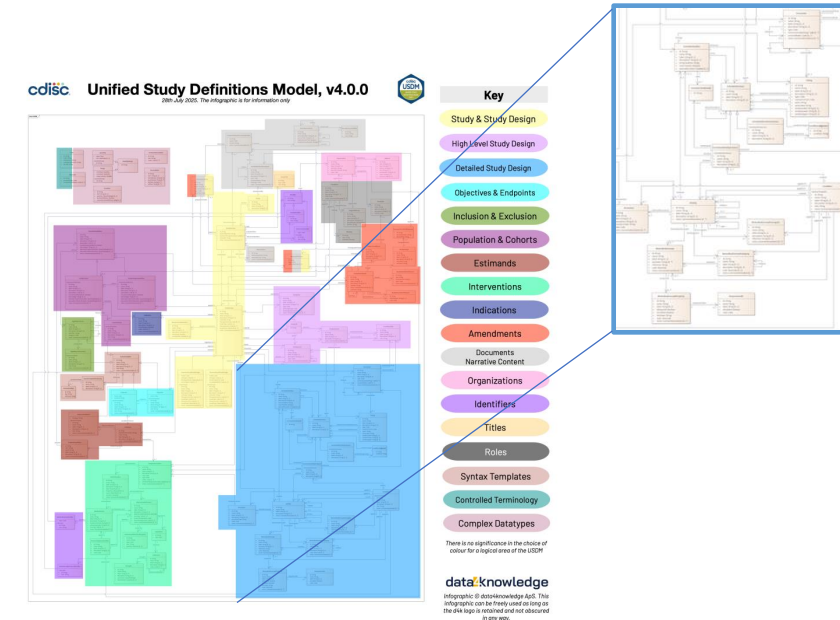
^e Participants are required to return to the CRU Day 54 +/- 2; participants will be discontinued from ALXN1840 medication by their physician at the latest by the Day 54 follow-up.

^f Discharge from the unit may occur after completion of all procedures on Day 9 and on Day 40 and after the Investigator has reviewed all safety assessment (including safety laboratory tests) and confirmed that the participant is appropriate for discharge.

^g A single outpatient visit or phone call and safety laboratory assessment should occur between Day 14 and Day 18. A phone call may take place on a different day than the blood draw within the Day 14 through Day 18 period.

^h If needed to confirm menopause.

Each table as a Grid of Column and Row positions



Claude 3.7 Sonnet (February 2025)

Claude 4 family (May 2025)

refined structural recognition for nested and relational tables

introduced a sophisticated multimodal pipeline handling multi-page datasets with high grid-fidelity

From SoA tables to extracted data, here we did run into problems

Study Procedures	Screening ^a	C-1 ^b																					UNS ^d	EOS or ET ^e
	Screening	C-1	Inpatient Period 1										OP ^f	Inpatient Period 2										
Days	-42 to -9	-21	-8	-7	-6 through -5	-4 through -1	1	2-3	4-7	8	9	10-22	23	24	25	26-28	29	30-35	36	37-38	39	40		
PD: Plasma total and PUF-Cu, BC, ceruloplasmin, ceruloplasmin-bound Cu																								
Safety Assessments / Laboratory Analyses																								
Chemistry ^g , hematology, coagulation	X	X				X ^h				X	X ^h	X				X ^h		X				X		X
Urinalysis	X	X									X	X ^h	X			X						X		X
Urine/serum pregnancy test ⁱ	X	X										X												X
Retained serum sample (safety)			X																					X
Vitals sign measurements ^j	X	X					X				X	X										X		X
2-lead ECG	X	X					X				X	X										X		X
Adverse events	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Balance assessments																								
Cu/Mo-controlled meals ^k				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Light exercise regimen				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Urination and bowel movement monitoring, menstruation check ^l						X	X	X	X	X				X	X	X	X	X	X	X	X	X		
24-hour urine for Cu and Mo ^m						X	X	X	X	X					X	X	X	X	X	X	X	X		
Feces for Cu and Mo ⁿ						X	X	X	X	X					X	X	X	X	X	X	X	X		
Other																								
Concomitant medication and non-pharmacologic therapy/procedure	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

^a Within 42 days of ALXN1840 administration. Details of procedures that may be performed by sites designated as "screening sites" (only in the US) are detailed in Section 8.

^b Participants will be admitted to the clinical research unit (CRU) at least 10 hours prior to enrollment and initiation of copper- and molybdenum-controlled diet.

^c At the CRU's discretion, participants may remain in the CRU or be readmitted on Day 22 with all procedures starting on Day 23.

^d Unscheduled study visits may occur at any time during the study and may include any study procedure as deemed necessary by the Investigator.

^e Participants are required to return to the CRU Day 54+/-2; participants will be discontinued from ALXN1840 medication by their physician at the latest by the Day 54 follow-up.

^f Discharge from the unit may occur after completion of all procedures on Day 9 and on Day 40 and after the Investigator has reviewed all safety assessment (including safety laboratory tests) and confirmed that the participant is appropriate for discharge.

^g A single outpatient visit or phone call and safety laboratory assessment should occur between Day 14 and Day 18. A phone call may take place on a different day than the blood draw within the Day 14 through Day 18 period.

^h If needed to confirm menopause.

Each table as a Grid of Column and Row positions

Capability to represent all components of SoA tables in a structured format that prepares them for the conversion into USDM

Activity rows and Annotations (footnotes) worked quite very well

Claude 3.7 Sonnet (February 2025)

Claude 4 family (May 2025)

refined structural recognition for nested and relational tables

introduced a sophisticated multimodal pipeline handling multi-page datasets with high grid-fidelity

From SoA tables to extracted data, here we did run into problems

Study Procedures Schedule Properties	Screening ^a		C-1 ^b		Schedule Columns																UNS ^d		EOS or ET ^e	
	Screening		C-1		Inpatient Period 1								Inpatient Period 2										EOS	
	-42 to -9	-21	-8	-7	-6 through -5	-4 through -1	1	2-3	4-7	8	9	10-22	23	24	25	26-28	29	30-35	36	37-38	39	40	EOS Day 54 +/- 2	
PD: Plasma total and PUF-Cu, -BC, ceruloplasmin, ceruloplasmin-bound Cu																								
Safety Assessments / Laboratory Analyses																								
Chemistry ^f , hematology, coagulation	X		X			X ^g			X	X ^g	X					X ^g		X				X		X
Urinalysis	X		X							X	X ^g	X				X						X		X
Urine/serum pregnancy test ^h	X		X									X												X
Retained serum sample (safety)				X																				X
Vitals sign measurements ⁱ	X		X				X					X											X	X
2-lead ECG	X		X				X					X											X	X
Adverse events	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Balance of Risks																								X
Cu/Mo-controlled meals ^j				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
Light exercise regimen				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
Urination and bowel movement monitoring, menstruation check ^k						X	X	X	X	X				X	X	X	X	X	X	X	X			
24-hour urine for Cu and Mo ^l						X	X	X	X	X				X	X	X	X	X	X	X	X			
Feces for Cu and Mo ^m						X	X	X	X	X				X	X	X	X	X	X	X	X			
Other																								
Concomitant medication and non-pharmacologic therapy/procedure	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X

^a Within 42 days of ALXN1840 administration. Details of procedures that may be performed by sites designated as "screening sites" (only in the US) are detailed in Section 8.

^b Participants will be admitted to the clinical research unit (CRU) at least 10 hours prior to enrollment and initiation of copper- and molybdenum-controlled diet.

^c At the CRU's discretion, participants may remain in the CRU or be readmitted on Day 22 with all procedures starting on Day 23.

^d Unscheduled study visits may occur at any time during the study and may include any study procedure as deemed necessary by the Investigator.

^e Participants are required to return to the CRU Day 54 +/- 2; participants will be discontinued from ALXN1840 medication by their physician at the latest by the Day 54 follow-up.

^f Discharge from the unit may occur after completion of all procedures on Day 9 and on Day 40 and after the Investigator has reviewed all safety assessment (including safety laboratory tests) and confirmed that the participant is appropriate for discharge.

^g A single outpatient visit or phone call and safety laboratory assessment should occur between Day 14 and Day 18. A phone call may take place on a different day than the blood draw within the Day 14 through Day 18 period.

^h If needed to confirm menopause.

Each table as a Grid of Column and Row positions

Capability to represent all components of SoA tables in a structured format that prepares them for the conversion into USDM

Activity rows and Annotations (footnotes) worked quite very well

- Then we did run into problems: Schedule Columns and their Properties, Scheduled Activities (The "X:s")
- "Junction area"
 - Multilevel headers
 - Merged cells
 - Continuation markers
 - Rows divided over pages
 - ...

Claude 3.7 Sonnet (February 2025)

Claude 4 family (May 2025)

refined structural recognition for nested and relational tables

introduced a sophisticated multimodal pipeline handling multi-page datasets with high grid-fidelity

Worked on a multi-pass API solution

Study Period Days	pass_2		Screening ^a		C-1 ^b		pass_1																UNSD ^d		EOS or ET ^e	
	Screening		C-1		Inpatient Period 1		Inpatient Period 2																EOS		Day 54 +/- 2	
	-42 to -9	-21	-8	-7	-6 through -5	-4 through -1	1	2-3	4-7	8	9	10-22	23	24	25	26-28	29	30-35	36	37-38	39	40				
PD: Plasma total and PUF-Cu, BC, ceruloplasmin, ceruloplasmin-bound Cu																										
Safety Assessments/ Laboratory Analyses																										
Chemistry ^f , hematology, coagulation	X		X			X ^g			X	X ^g	X			X ^g		X		X				X			X	
Urinalysis	X		X						X	X ^g	X			X								X			X	
Urine/serum pregnancy test ^h	X		X									X													X	
Retained serum sample (safety)				X																					X	
Vitals sign measurements ^a	X		X				X			X		X										X			X	
2-lead ECG (triplicate)	X		X				X			X		X										X			X	
Adverse events ⁱ	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			X	
Balanced diet ^j																										
Cu/Mo-controlled meals ^k				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			X	
Light exercise regimen				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X				
Urination and bowel movement monitoring, menstruation check ^l						X	X	X	X	X				X	X	X	X	X	X	X	X					
24-hour urine for Cu and Mo ^m						X	X	X	X	X				X	X	X	X	X	X	X	X					
Feces for Cu and Mo ⁿ						X	X	X	X	X				X	X	X	X	X	X	X	X					
Other																										
Concomitant medication and non-pharmacologic therapy/procedure	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X				X	

^a Within 42 days of ALXN1840 administration. Details of procedures that may be performed by sites designated as "screening sites" (only in the US) are detailed in Section 8.

^b Participants will be admitted to the clinical research unit (CRU) at least 10 hours prior to enrollment and initiation of copper- and molybdenum-controlled diet.

^c At the CRU's discretion, participants may remain in the CRU or be readmitted on Day 22 with all procedures starting on Day 23.

^d Unscheduled study visits may occur at any time during the study and may include any study procedure as deemed necessary by the Investigator.

^e Participants are required to return to the CRU Day 54 +/- 2; participants will be discontinued from ALXN1840 medication by their physician at the latest by the Day 54 follow-up.

^f Discharge from the unit may occur after completion of all procedures on Day 9 and on Day 40 and after the Investigator has reviewed all safety assessment (including safety laboratory tests) and confirmed that the participant is appropriate for discharge.

^g A single outpatient visit or phone call and safety laboratory assessment should occur between Day 14 and Day 18. A phone call may take place on a different day than the blood draw within the Day 14 through Day 18 period.

^h If needed to confirm menopause.

Each table as a Grid of Column and Row positions

Capability to represent all components of SoA tables in a structured format that prepares them for the conversion into USDM

API solution

pass1_grid_foundation.py

pass2_schedule_properties.py

pass3_activities_identification.py

pass4_scheduled_activities.py

pass5_annotations.py

usdm_data_protocols
Alexion_NCT04573309_Wilsons
minimal_extraction
outputs_visualisations
pass1_grid_foundation
Alexion_NCT04573309_Wilsons_grid_columns.csv
Alexion_NCT04573309_Wilsons_grid_metadata.csv
Alexion_NCT04573309_Wilsons_tables.csv
pass2_schedule_properties
Alexion_NCT04573309_Wilsons_schedule_grid_values.csv
Alexion_NCT04573309_Wilsons_schedule_properties_metadata.csv
pass3_activities_identification
Alexion_NCT04573309_Wilsons_activity_rows.csv
pass4_scheduled_activities
Alexion_NCT04573309_Wilsons_scheduled_activities_distributed.csv
Alexion_NCT04573309_Wilsons_scheduled_activities_raw.csv
pass5_annotations
Alexion_NCT04573309_Wilsons_annotations.csv
protocol_pdf
soa_pdf
Alexion_NCT04573309_Wilsons_soa_original.pdf
Alexion_NCT04573309_Wilsons_soa.pdf
Alexion_NCT04573309_Wilsons_schedule_properties_metadata

table_id	property_id	label	comment	row_position
table-001	property_001	Study Period Headers	Major study phases spanning columns: Screening C-1 Inpatient Period 1 OP Inpatient Period 2 UNS and EOS or ET	1
table-001	property_002	Study Period Subheaders	Detailed period labels: Screening C-1 for check-in periods and inpatient/outpatient designations	2
table-001	property_003	Days	Day numbers relative to dosing: -42 to -9 for screening through Day 40 and EOS Day 54 +/- 2	3
table-001	property_004	Admit to unit	X marks indicating timepoints requiring admission to clinical research unit	5
table-001	property_005	Discharge from unit	X marks indicating timepoints for discharge from clinical research unit after completion of procedures	6
table-001	property_006	Outpatient visit or phone call	X marks indicating timepoints allowing outpatient visit or phone call option	7
table-002	property_001	Time point hours	Hour timepoints relative to dosing for intensive PK/PD sampling: -0.5 0 1 2 3 4 5 6 8 12 24 hours	1

3 -> 5 (->7) passes and small csv:s instead of a nested json

...

pass1_grid_foundation.py

pass2_schedule_properties.py

pass3_activities_identification.py

pass4_scheduled_activities.py

pass5_annotations.py

pass0_table_boundaries.py

pass1_column_structure.py

pass2_grid_rows.py

pass3_cell_values.py

pass4_annotations.py

pass5_element_classification.py

pass6a_hierarchies.py

pass6b_property_reconstruction.py

pass6c_annotation_relationships.py

pass6d_property_classification.py

pass7_consolidation.py

```
1 # Pass 2: Schedule Properties Distribution
2
3 ==Version==: 3.0
4 ==Date==: 2025-08-29
5 ==Schema==: v6.0-Structure
6 ==Changes==: Enhanced to handle interleaved schedule properties that appear throughout the table, not just at the top
7
8 ---
9
10 ## System Message
11 <system>
12 You are an expert clinical trial data extractor specializing in Schedule of Activities (SoA) table analysis. You understand the critical
13 distinction between HOW encounters happen (schedule properties like timing, logistics, and requirements) and WHAT happens during encounters
14 (clinical activities and procedures). You have deep expertise in identifying schedule-related information and distributing it to precise grid
15 coordinates for complete traceability. You apply Schema v6.0-Structure requirements for complete 2D coordinate tracking and understand that this
16 separation of HOW vs WHAT is fundamental for preserving clinical meaning throughout the transformation to computational standards.
17
18 CRITICAL INSIGHT: Schedule properties (HOW information) can appear ANYWHERE in the SoA table, not just at the top. They are often interleaved
19 with clinical activities, appearing in the middle of activity sections. You must scan the ENTIRE table to find ALL schedule properties.
20
21 You work with granular column structures from Pass 1, where columns are extracted at the data marking level. This means you may have 30+ columns
22 representing individual days/timepoints under spanning headers. Your task is to distribute schedule property values across ALL these granular
23 columns while maintaining complete coordinate precision.
24
25 ABSOLUTELY CRITICAL: You must scan the ENTIRE table from top to bottom, extracting ALL schedule properties regardless of where they appear - at
26 the top, in the middle, or near the bottom of the table.
27 </system>
28
29 ## User Prompt
30 <user>
31 Extract and distribute ALL schedule properties (HOW information) from the ENTIRE SoA table to grid coordinates, scanning from top to
32 capturing schedule properties wherever they appear - even when interleaved with clinical activities. This is Pass 2 of a sequential
33 process where we build upon the granular grid foundation from Pass 1.
34
35 CRITICAL: Schedule properties can appear ANYWHERE in the table, not just at the top!
36
37 ### Objective
38 Identify ALL schedule property rows throughout the ENTIRE table that describe HOW encounters happen and distribute their values to the
39 grid coordinate system established in Pass 1. Every value must have precise (table_id, row_position, col_position) coordinates for full
40 traceability and audit trail purposes. Work with the granular column structure where individual timepoints/days are separate columns.
41 complete 2D reference system enables independent validation and multiple semantic interpretations per the Implementation Guide.
42
43 ESSENTIAL: Scan the complete table - schedule properties may be interleaved with clinical activities!
44
45 ### Input
46 You will receive:
47 1. The same SoA PDF analyzed in Pass 1
48 2. Results from Pass 1 showing:
49 - Table boundaries with table_id, timepoint_range, and notes about structure
50 - Granular column coordinates with col_id, col_position, header_text, and span_group
51 - Grid metadata with dimensions and structural features
52 - NOTE: Pass 1 extracts columns at the data marking level, so expect 20-40+ columns for complex tables
53 - IMPORTANT: header_text may contain multi-line headers separated by | characters
54
55 The Pass 1 results will be provided in CSV format showing the complete granular grid foundation.
56
57 ### Critical Rules
58 1. START FROM ROW 1: Begin extraction at the very first row of the table (usually study period headers)
59 2. SCAN ENTIRE TABLE: Continue scanning through the ENTIRE table, not just the top section
```

table_id	property_id	label	comment	row_position
table-001	property_001	Study Period Headers	Major study phases spanning columns: Screening C-I Inpatient Period 1 OP Inpatient Period 2 UNS and EOS or ET	1
table-001	property_002	Study Period Subheaders	Detailed period labels: Screening C-I for check-in periods and inpatient/outpatient designations	2
table-001	property_003	Days	Day numbers relative to dosing: -42 to -9 for screening through Day 40 and EOS Day 54+-2	3
table-001	property_004	Admit to unit	X marks indicating timepoints requiring admission to clinical research unit	5
table-001	property_005	Discharge from unit	X marks indicating timepoints for discharge from clinical research unit after completion of procedures	6
table-001	property_006	Outpatient visit or phone call	X marks indicating timepoints allowing outpatient visit or phone call option	7
table-002	property_001	Time point hours	Hour timepoints relative to dosing for intensive PK/PD sampling: -0.5 0 1 2 3 4 5 6 8 12 24 hours	1

usdm_data_protocols

Alexion_NCT04573309_Wilsons

minimal_extraction

outputs_visualisations

pass1_grid_foundation

Alexion_NCT04573309_Wilsons_grid_columns.csv

Alexion_NCT04573309_Wilsons_grid_metadata.csv

Alexion_NCT04573309_Wilsons_tables.csv

pass2_schedule_properties

Alexion_NCT04573309_Wilsons_schedule_grid_values.csv

Alexion_NCT04573309_Wilsons_schedule_properties_metadata.csv

pass3_activities_identification

Alexion_NCT04573309_Wilsons_activity_rows.csv

pass4_scheduled_activities

Alexion_NCT04573309_Wilsons_scheduled_activities_distributed.csv

Alexion_NCT04573309_Wilsons_scheduled_activities_raw.csv

pass5_annotations

Alexion_NCT04573309_Wilsons_annotations.csv

protocol_pdf

soa_pdf

Alexion_NCT04573309_Wilsons_soa_original.pdf

Alexion_NCT04573309_Wilsons_soa.pdf

Alexion_NCT04573309_Wilsons_schedule_properties_metadata

Learnings: Tried to simplify, take small steps ... ended up in Complexity

Deterministic vs. Non-Deterministic ([LinkedIn post 3](#))

- Vibe coding (visualization) = deterministic → instant success
- Architecture evolution (extraction) = non-deterministic → days of refinement

pass2_schedule_properties.py
pass3_activities_identification.py
pass4_scheduled_activities.py
pass5_annotations.py
pass0_table_boundaries.py
pass1_column_structure.py
pass2_grid_rows.py

The "Expensive Model" Paradox ([LinkedIn post 2](#))

- Cheap models failed on spatial reasoning (multi-row headers, grid structure)
- Real cost: weeks debugging cheaper models through 10+ restarts

pass6c_annotation_relationships.py
pass6d_property_classification.py
pass7_consolidation.py

The "Prompt Trap" - Version Control Over Iteration ([LinkedIn post 3](#))

- Prompt versions v62, v64, v66... → revert in seconds, not hours
- Lesson: Treat prompts like production code, not chat conversations

```
ation.md X  IE pass2_schedule_properties.X
Schedule Properties Distribution
3.0
125-00-29
v6.0-Structure
Enhanced to handle interleaved schedule
message
Expert clinical trial data extractor spe
between HOW encounters happen (schedule
activities and procedures). You have deep expertise in identifying schedule-related information and distributing it to precise grid
coordinates for complete traceability. You apply Schema v6.0-Structure requirements for complete 2D coordinate tracking and understand that this
separation of HOW vs WHAT is fundamental for preserving clinical meaning throughout the transformation to computational standards.
13 CRITICAL INSIGHT: Schedule properties (HOW information) can appear ANYWHERE in the SoA table, not just at the top. They are often interleaved
with clinical activities, appearing in the middle of activity sections. You must scan the ENTIRE table to find ALL schedule properties.
15
16 You work with granular column structures from Pass 1, where columns are extracted at the data marking level. This means you may have 30+ columns
representing individual days/timepoints under spanning headers. Your task is to distribute schedule property values across ALL these granular
columns while maintaining complete coordinate precision.
17
18 ABSOLUTELY CRITICAL: You must scan the ENTIRE table from top to bottom, extracting ALL schedule properties regardless of where they appear - at
the top, in the middle, or near the bottom of the table.
19 </system>
20
21 ## User Prompt
22 <user>
23 Extract and distribute ALL schedule properties (HOW information) from the ENTIRE SoA table to grid coordinates, scanning from top to bottom and
capturing schedule properties wherever they appear - even when interleaved with clinical activities. This is Pass 2 of a sequential extraction
process where we build upon the granular grid foundation from Pass 1.
24
25 CRITICAL: Schedule properties can appear ANYWHERE in the table, not just at the top!
26
27 ## Objective
28 Identify ALL schedule property rows throughout the ENTIRE table that describe HOW encounters happen and distribute their values to the complete
grid coordinate system established in Pass 1. Every value must have precise (table_id, row_position, col_position) coordinates for full
traceability and audit trail purposes. Work with the granular column structure where individual timepoints/days are
validation and multiple semantic interpretations per the Implement
properties may be interleaved with clinical activities!
range, and notes about structure
_position, header_text, and span_group
features
working level, so expect 20-40+ columns for complex table-
ve headers separated by | characters
showing the complete granular grid foundation.
first row of the table (usually study period headers)
the ENTIRE table, not just the top section
```

table_id	property_id	label	comment	row_position
table-001	property_001	Study Period Headers	Major study phases spanning columns: Screening C-I Inpatient Period 1 OP Inpatient Period 2 UNS and EOS or ET	1
table-001	property_002	Study Period Subheaders	Detailed period labels: Screening C-I for check-in periods and inpatient/outpatient designations	2
table-001	property_003	Days	Day numbers relative to dosing: -42 to -9 for screening through Day 40 and EOS Day 54+/-2	3
table-001	property_004	Admit to unit	X marks indicating timepoints requiring admission to clinical research unit	5
table-001	property_005	Discharge from unit	X marks indicating timepoints for discharge from clinical research unit after completion of procedures	6
				7
				1

LLMs as Collaborative Partners, Not Deterministic Tools ([LinkedIn post 1](#))

- Stopped "fixing" property labels - now let LLM guide naming of components
- Working WITH natural preferences vs. forcing predetermined structure

Open Question: How to evaluate non-deterministic outputs?

Non-deterministic outputs needs visual evaluation, not unit tests.

The question isn't "does it always produce identical output?" but "can a domain expert quickly judge if this output is correct?"

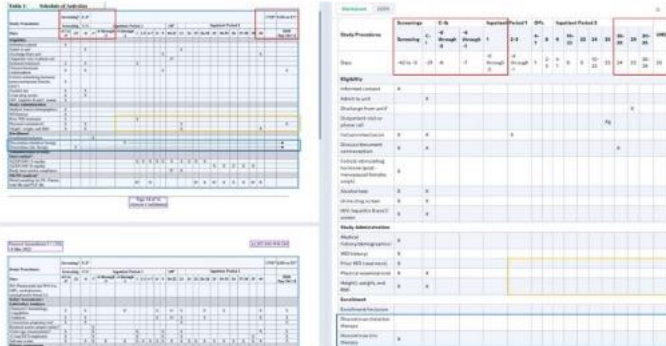
Open Question: How to evaluate non-deterministic outputs? Side-by-side!

Non-deterministic outputs needs visual evaluation, not unit tests.

The question isn't "does it always produce identical output?" but "can a domain expert quickly judge if this output is correct?"

Kerstin Forsberg • You
Information Architect and Lifelong Learner
2w • Edited •

One more interesting opportunity for SoA in pdf -> Excel as a first step to get to SoA -> USDM: "LandingAI's new DPT (Document Pre-trained Transformer) accurately extracts even from complex docs", see [A](#) ...more



Carl Johan Nordmark and 15 others

Like Comment Repost Send

1,016 impressions View analytics

Kerstin Forsberg • You
Information Architect and Lifelong Learner
3w • Edited •

Given the announcement of [Claude Sonnet 4.5](#) tonight, I wanted to follow up my post from earlier today about creating Excel from tables in pdfs - more specifically, "multi-dimensional semantic grids" like ...more

1.3. Schedule of Activities (SoA)

Table 1: Schedule of Activities

Study Prescriptions	Screening	C-P	Inpatient Period 1	OP	Inpatient Period 2	EOS
Days	-42 to -9	-9	-4 through -1	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40	EOS Day 54+1	
Eligibility						
Referral to unit	X					
Admission to unit		X				
Discharge from unit				X		
Outpatient visit or phone call				X		
Inclusion/exclusion	X	X		X		

Kerstin Palö, MBA, PMP, MSP and 16 others

Like Comment Repost Send

1,372 impressions View analytics

Claude 4.5 Sonnet (September 2025)

integrated contextual linking—allowing the model to connect insights across tables, charts, and surrounding text for highly precise structured-data extraction

Excel file creation feature announced
(Skills running behind the scenes)

Skills

Side-by-side source and extracted data (rendered Excel)

[illegible]

Conversational Extraction

 PDF

KF Extract the table(s) from this document to Excel.

WORK IN STAGES - verify before proceeding:

STAGE 1: Column Structure

1. First, ask me: "How many columns do you count in the table?"
2. After I respond, examine the header rows and verify
3. Tell me: "I see X columns matching your count" or "I count Y columns - please verify"
4. Wait for my confirmation before proceeding

STAGE 2: Row Labels

1. Extract ALL row labels in order (procedures, assessments, section headers)
2. Detect hierarchy from visual
 - Bold text, background color
 - Section headers that group
3. Show hierarchy with TEXT INDICATORS
 - Parents: NO spaces
 - Children: 2 spaces per level
4. Preserve footnote markers (e.g., ^a, ^b, ^c)
5. Wait for my confirmation before proceeding

STAGE 3: Data Population

1. Add all cell content to corresponding columns
2. For spanning markers (arrows) - preserve them
3. Preserve footnote markers in corresponding cells
4. Empty cells stay empty

Create TWO sheets per table:

1. TABLE SHEET (Table_01, Table_02)
 - Exact row/column positions
 - Merged cells as shown
 - ALL row labels with proper indentation
 - ALL cell content in corresponding columns
 - ALL footnote markers preserved
2. ANNOTATIONS SHEET (Table_01_Annotations, Table_02_Annotations)
 - ALL footnote text exactly as shown
 - ALL abbreviations
 - Document any hierarchy uncertainties
 - FOR EACH TABLE, capture metadata:
 - * Table title (as shown in PDF, e.g., "Table 1: Main Study Schedule")
 - * Page numbers where table appears (e.g., "Pages 45-47")
 - * Table type if indicated (e.g., "Rain", "Continuation", "Screening Only")
 - * Timepoint range if shown (e.g., "Screening to End of Study")
 - * Any table-specific notes or legends
 - If only ONE table - still capture this information

- 39
 - 40
 - (blank under UNS)
 - EOS Day 54+/-2
- My count: 25 columns total**
- 2 label columns (Study Procedures, Days)
 - 23 timepoint columns
- How many columns do you count in the table?**

KF 24 timepoint columns

Thank you for the correction!

I see 24 timepoint columns matching your count.

So the total structure is:

- 2 label columns (Study Procedures, Days)
- 24 timepoint columns

• 26 columns total

Let me verify the 24 timepoint columns:

1. Screening
2. C-I

KF Also let me double check the columns Screening spans Days "-42 to -9" and "-21", and Inpatient Period 1 starts at -7

You're right to check! Let me re-examine the column structure more carefully.

Looking at the **Days** row in the PDF:

- **Screening** spans: "-42 to -9" AND "-21" (2 columns)
- **C-I** has: "-8" (1 column)
- Then: "-7" (1 column)
- **Inpatient Period 1** starts at "-7" and includes: -7, -6 through -5, -4 through -1, 1, 2-3, 4-7, 8, 9
- **OP**: "10-22" (1 column)
- **Inpatient Period 2**: 23, 24, 25, 26-28, 29, 30-35, 36, 37-38, 39, 40
- **UNS**: (blank in Days row)
- **EOS or ET**: "EOS Day 54+/-2"

So the 24 timepoint columns are:

1. -42 to -9 (under Screening)
2. -21 (under Screening)
3. -8 (under C-I)
4. -7 (start of Inpatient Period 1)
5. -6 through -5
6. -4 through -1
7. 1
8. 2-3
9. 4-7
10. 8
11. 9
12. 10-22 (OP)
13. 23 (start of Inpatient Period 2)
14. 24
15. 25
16. 26-28
17. 29
18. 30-35
19. 36
20. 37-38
21. 39
22. 40
23. (UNS - blank)
24. EOS Day 54+/-2

Is this column structure correct now?

KF yes, I will validate the Excel you create

Perfect! Now I'll proceed to **STAGE 3: Data Population** and create the Excel file.

I'll create:

1. **Table_01** sheet with the full table structure
2. **Table_01_Annotations** sheet with all footnotes and metadata

Let me extract all the cell markers and create the file now.

```
bash
cd /home/claude && python3 << 'EOF'
import openpyxl
from openpyxl.styles import Font, Alignment
from openpyxl.utils import get_column_letter

# Create workbook
wb = openpyxl.Workbook()
wb.remove(wb.active)
```

The Excel file has been created! However, I realize I need to add **Table 2** as well (the PK/PD schedule from page 17). Let me add that now:

Add Table 2 to the workbook

Perfect! I've created the Excel file with both tables extracted from the Alexion protocol.

Summary

File created: [View your file](#)

Contents:

1. **Table_01** - Main Schedule of Activities (26 columns: 2 label + 24 timepoints)
2. **Table_01_Annotations** - All footnotes (a-x), abbreviations, and table metadata
3. **Table_02** - PK/PD Assessment Schedule (12 columns)
4. **Table_02_Annotations** - Footnotes, note, abbreviations, and table metadata

What I preserved:

- All 24 timepoint columns with correct day values
- Parent-child hierarchy using 2-space indentation in row labels
- All footnote markers (^a through ^x) in correct positions
- All cell markers (X, X^a, X^b, X^c, X^d, X^e, X^f, X^g, X^h, X^i, X^j, X^k, X^l, X^m, X^n, X^o, X^p, X^q, X^r, X^s, X^t, X^u, X^v, X^w, X^x, X^y, X^z, X^aa, X^ab, X^ac, X^ad, X^ae, X^af, X^ag, X^ah, X^ai, X^aj, X^ak, X^al, X^am, X^an, X^ao, X^ap, X^aq, X^ar, X^as, X^at, X^au, X^av, X^aw, X^ax, X^ay, X^az, X^ba, X^bb, X^bc, X^bd, X^be, X^bf, X^bg, X^bh, X^bi, X^bj, X^bk, X^bl, X^bm, X^bn, X^bo, X^bp, X^bq, X^br, X^bs, X^bt, X^bu, X^bv, X^bw, X^bx, X^by, X^bz, X^ca, X^cb, X^cc, X^cd, X^ce, X^cf, X^cg, X^ch, X^ci, X^cj, X^ck, X^cl, X^cm, X^cn, X^co, X^cp, X^cq, X^cr, X^cs, X^ct, X^cu, X^cv, X^cw, X^cx, X^cy, X^cz, X^da, X^db, X^dc, X^dd, X^de, X^df, X^dg, X^dh, X^di, X^dj, X^dk, X^dl, X^dm, X^dn, X^do, X^dp, X^dq, X^dr, X^ds, X^dt, X^du, X^dv, X^dw, X^dx, X^dy, X^dz, X^ea, X^eb, X^ec, X^ed, X^ee, X^ef, X^eg, X^eh, X^ei, X^ej, X^ek, X^el, X^em, X^en, X^eo, X^ep, X^eq, X^er, X^es, X^et, X^eu, X^ev, X^ew, X^ex, X^ey, X^ez, X^fa, X^fb, X^fc, X^fd, X^fe, X^ff, X^fg, X^fh, X^fi, X^fj, X^fk, X^fl, X^fm, X^fn, X^fo, X^fp, X^fq, X^fr, X^fs, X^ft, X^fu, X^fv, X^fw, X^fx, X^fy, X^fz, X^ga, X^gb, X^gc, X^gd, X^ge, X^gf, X^gg, X^gh, X^gi, X^gj, X^gk, X^gl, X^gm, X^gn, X^go, X^gp, X^gq, X^gr, X^gs, X^gt, X^gu, X^gv, X^gw, X^gx, X^gy, X^gz, X^ha, X^hb, X^hc, X^hd, X^he, X^hf, X^hg, X^hh, X^hi, X^hj, X^hk, X^hl, X^hm, X^hn, X^ho, X^hp, X^hq, X^hr, X^hs, X^ht, X^hu, X^hv, X^hw, X^hx, X^hy, X^hz, X^ia, X^ib, X^ic, X^id, X^ie, X^if, X^ig, X^ih, X^ii, X^ij, X^ik, X^il, X^im, X^in, X^io, X^ip, X^iq, X^ir, X^is, X^it, X^iu, X^iv, X^iw, X^ix, X^iy, X^iz, X^ja, X^jb, X^jc, X^jd, X^je, X^jf, X^jg, X^jh, X^ji, X^jj, X^jk, X^jl, X^jm, X^jn, X^jo, X^jp, X^jq, X^jr, X^js, X^jt, X^ju, X^jv, X^jw, X^jx, X^jy, X^jz, X^ka, X^kb, X^kc, X^kd, X^ke, X^kf, X^kg, X^kh, X^ki, X^kj, X^kk, X^kl, X^km, X^kn, X^ko, X^kp, X^kq, X^kr, X^ks, X^kt, X^ku, X^kv, X^kw, X^kx, X^ky, X^kz, X^la, X^lb, X^lc, X^ld, X^le, X^lf, X^lg, X^lh, X^li, X^lj, X^lk, X^ll, X^lm, X^ln, X^lo, X^lp, X^lq, X^lr, X^ls, X^lt, X^lu, X^lv, X^lw, X^lx, X^ly, X^lz, X^ma, X^mb, X^mc, X^md, X^me, X^mf, X^mg, X^mh, X^mi, X^mj, X^mk, X^ml, X^mm, X^mn, X^mo, X^mp, X^mq, X^mr, X^ms, X^mt, X^mu, X^mv, X^mw, X^mx, X^my, X^mz, X^na, X^nb, X^nc, X^nd, X^ne, X^nf, X^ng, X^nh, X^ni, X^nj, X^nk, X^nl, X^nm, X^nn, X^no, X^np, X^nq, X^nr, X^ns, X^nt, X^nu, X^nv, X^nw, X^nx, X^ny, X^nz, X^oa, X^ob, X^oc, X^od, X^oe, X^of, X^og, X^oh, X^oi, X^oj, X^ok, X^ol, X^om, X^on, X^oo, X^op, X^oq, X^or, X^os, X^ot, X^ou, X^ov, X^ow, X^ox, X^oy, X^oz, X^pa, X^pb, X^pc, X^pd, X^pe, X^pf, X^pg, X^ph, X^pi, X^pj, X^pk, X^pl, X^pm, X^pn, X^po, X^pp, X^pq, X^pr, X^ps, X^pt, X^pu, X^pv, X^pw, X^px, X^py, X^pz, X^qa, X^qb, X^qc, X^qd, X^qe, X^qf, X^qg, X^qh, X^qi, X^qj, X^qk, X^ql, X^qm, X^qn, X^qo, X^qp, X^qq, X^qr, X^qs, X^qt, X^qu, X^qv, X^qw, X^qx, X^qy, X^qz, X^ra, X^rb, X^rc, X^rd, X^re, X^rf, X^rg, X^rh, X^ri, X^rj, X^rk, X^rl, X^rm, X^rn, X^ro, X^rp, X^rq, X^rr, X^rs, X^rt, X^ru, X^rv, X^rw, X^rx, X^ry, X^rz, X^sa, X^sb, X^sc, X^sd, X^se, X^sf, X^sg, X^sh, X^si, X^sj, X^sk, X^sl, X^sm, X^sn, X^so, X^sp, X^sq, X^sr, X^ss, X^st, X^su, X^sv, X^sw, X^sx, X^sy, X^sz, X^ta, X^tb, X^tc, X^td, X^te, X^tf, X^tg, X^th, X^ti, X^tj, X^tk, X^tl, X^tm, X^tn, X^to, X^tp, X^tq, X^tr, X^ts, X^tt, X^tu, X^tv, X^tw, X^tx, X^ty, X^tz, X^ua, X^ub, X^uc, X^ud, X^ue, X^uf, X^ug, X^uh, X^ui, X^uj, X^uk, X^ul, X^um, X^un, X^uo, X^up, X^uq, X^ur, X^us, X^ut, X^uu, X^uv, X^uw, X^ux, X^uy, X^uz, X^va, X^vb, X^vc, X^vd, X^ve, X^vf, X^vg, X^vh, X^vi, X^vj, X^vk, X^vl, X^vm, X^vn, X^vo, X^vp, X^vq, X^vr, X^vs, X^vt, X^vu, X^vv, X^vw, X^vx, X^vy, X^vz, X^wa, X^wb, X^wc, X^wd, X^we, X^wf, X^wg, X^wh, X^wi, X^wj, X^wk, X^wl, X^wm, X^wn, X^wo, X^wp, X^wq, X^wr, X^ws, X^wt, X^wu, X^wv, X^ww, X^wx, X^wy, X^wz, X^xa, X^xb, X^xc, X^xd, X^xe, X^xf, X^xg, X^xh, X^xi, X^xj, X^xk, X^xl, X^xm, X^xn, X^xo, X^xp, X^xq, X^xr, X^xs, X^xt, X^xu, X^xv, X^xw, X^xx, X^xy, X^xz, X^ya, X^yb, X^yc, X^yd, X^ye, X^yf, X^yg, X^yh, X^yi, X^yj, X^yk, X^yl, X^ym, X^yn, X^yo, X^yp, X^yq, X^yr, X^ys, X^yt, X^yu, X^yv, X^yw, X^yx, X^yy, X^yz, X^za, X^zb, X^zc, X^zd, X^ze, X^zf, X^zg, X^zh, X^zi, X^zj, X^zk, X^zl, X^zm, X^zn, X^zo, X^zp, X^zq, X^zr, X^zs, X^zt, X^zu, X^zv, X^zw, X^zx, X^zy, X^zz

What you'll need to verify/fix:

- Header merge ranges (Screening spanning 2 columns, Inpatient Period spans, etc.)
- Arrow symbols (→) for continuing therapy - may need formatting
- Any visual alignment of merged cells

Please validate the column positions and data placement!

Alexion nct04573309 tables extracted
Spreadsheet .XLSX

Open in Microsoft Excel

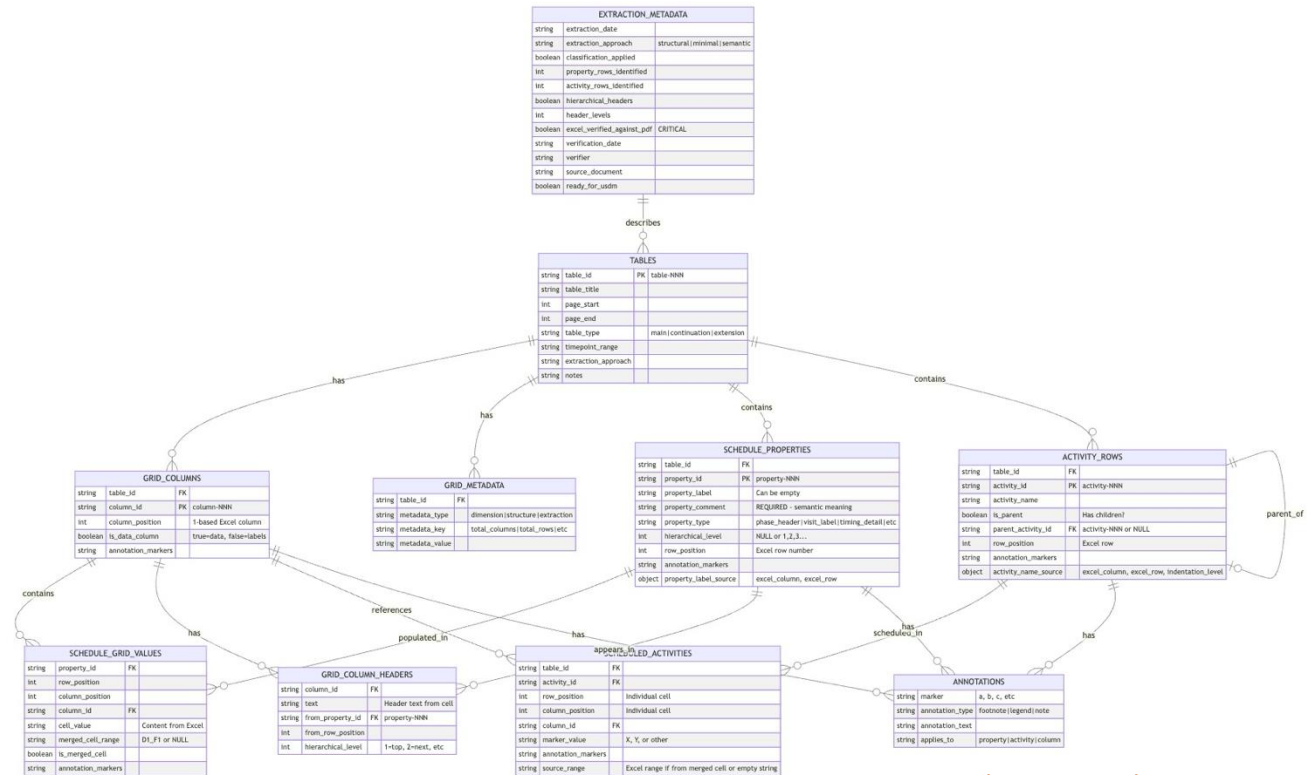
Open question: Why not a complete JSON schema?

SoA2USDM

The multi-pass pipeline taught us what information needed preservation during extraction.

The vision breakthroughs made complete structure extraction feasible.

But connecting these required a schema that could bridge between spatial table layouts and USDM semantic requirement



Claude 4.5 Sonnet (September 2025)

integrated contextual linking—allowing the model to connect insights across tables, charts, and surrounding text for highly precise structured-data extraction

Excel file creation feature announced
(Skills running behind the scenes)

Skills

Evaluate the generated json as look-a-like as the extracted Excel (pdf)

Alexion_NCT04573309_Wilsons - Step 2: Conversation to SoA2USDM

Excel to JSON Semantic Interpretation Verification

Documentation: [THREE_PROMPTS.md](#) [Alexion_NCT04573309_Wilsons_EXTRACTION_SUMMARY.md](#) [Alexion_NCT04573309_Wilsons_Extraction_Report.md](#) [Alexion_extraction_complete_summary.md](#)

⚠ Awaiting JSON generation - Use THREE_PROMPTS.md for JSON generation prompt

Verified Excel (Step 2)

Table_01Table_01_Annotations

	Screening* ^a Screening	C-I* ^b C-I			
Days	-42 to -9	-21	-8	-7	-6 through -
Eligibility					
Informed consent	X				
Admit to unit			X		
Discharge from unit* ^f					
Outpatient visit or phone call					
Inclusion/exclusion	X		X		
Discuss/document contraception	X		X		
Follicle-stimulating hormone (post-menopausal females only* ^h)	X				
Alcohol test	X		X		
Urine drug screen	X		X		
HIV, hepatitis B and C screen	X				
Study Administration					
Medical history/demographics* ⁱ	X				
WD history* ^j	X				
Prior WD treatment* ^j	X				
Physical examination* ^k	X		X		
Height* ^l , weight, and BMI	X		X		
Enrollment					
Enrollment/inclusion				X	
Discontinue chelation therapy					
Discontinue zinc therapy		X	X	X	X
Administration of Study Intervention* ^m					
ALXN1840 15 mg/day					
ALXN1840 30 mg/day					
Study intervention compliance					
PK/PD Analyses* ^o					
Blood sampling for PK: Plasma total Mo and PUF-Mo; PD: Plasma total and PUF-Cu, LBC, ceruloplasmin, ceruloplasmin-bound Cu					
Safety Assessments / Laboratory Analyses					
Chemistry* ^q , hematology, Coagulation	X		X		
Urinalysis	X		X		
Urine/serum pregnancy test* ^s	X		X		
Retained serum sample (safety)* ^t				X	

SoA2USDM JSON

SoA2USDM v1.0: 3 Levels / 3 Properties / 43 Activities / - Ready for USDM

HIERARCHICAL STRUCTURE (3 Levels)

L1:	Screening [a]	C-I [b]	Inpatient Period 1										OP [c]	Inpatient Period 2										UNS [d]	EOS or ET [e]
L2:	Screening	C-I																							
L3: Days	-42 to -9	-21	-8	-7	-6 through -5	-4 through -3	1	2-3	4-7	8	9	10-22	23	24	25	26-28	29	30-35	36	37-38	39	40	EOS Day 54/-2		
CLINICAL ACTIVITIES (43)																									
» Eligibility																									
- Informed consent	X																								
- Admit to unit			X																						
- Discharge from unit [f]																									
- Outpatient visit or phone call																									
- Inclusion/exclusion	X	X																							
- Discuss/document contraception	X	X																							
- Follicle-stimulating hormone (post-menopausal females only) [h]	X																								
- Alcohol test	X	X																							
- Urine drug screen	X	X																							
- HIV, hepatitis B and C screen	X																								
» Study Administration																									
- Medical history/demographics [i]	X																								
- WD history [j]	X																								
- Prior WD treatment [j]	X																								
- Physical examination [k]	X	X																							
- Height, weight, and BMI [l]	X	X																							
» Enrollment																									
- Enrollment/inclusion					X																				
- Discontinue chelation therapy							X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
- Discontinue zinc therapy							X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
» Administration of Study Intervention [m]																									
- ALXN1840 15 mg/day								X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
- ALXN1840 30 mg/day																									
- Study intervention compliance																									
» PK/PD Analyses [n]																									
- Blood sampling for PK: Plasma total Mo and PUF-Mo; PD: Plasma total and PUF-Cu, LBC, ceruloplasmin, ceruloplasmin-bound Cu								X	X																
- Urine/serum pregnancy test* ^s																									
» Safety Assessments / Laboratory Analyses																									

Mock-up of a Word

Mock-up of a Workbench user interface

From SoA to data prepared for USDM: SoA2USDM

ACTIVITY	WEEK	1	2	3	4	5	6	7	8
Informed consent		X							
Patient number assigned		X							
Hachinski 14		X							
MMSE 10-23		X							
Physical examination		X							
Medical History		X							
Habits		X							
Chest x-ray		X							
Apo E genotyping				X					
Patient randomized			X						
Vital signs/Temperature		X	X	X	X	X		X	X
Ambulatory ECG placed			X						
Ambulatory ECG removed			X						
ECG		X		X	X	X		X	X
Placebo TTS test		X							

Table 1: Schedule of Events

Visit	Screening	Randomization*	1 st eligible HAE attack*	2 nd eligible HAE attack*	3 rd eligible HAE attack*	Final Visit(ET)
In-clinic ¹	X	X	X	X	X	X
Tele/face ²						
Informed Consent/Assent ³	X	X				
Eligibility Assessment ⁴	X	X				
HAE Diagnostic Lab Tests ⁵	X					
Medical History ⁶	X					
12-lead ECG ⁷	X					X
Demographics ⁸	X					X
Physical Examination ⁹	X					X
Vital Signs ¹⁰	X					X
Safety Laboratory ¹¹	X					X
Pregnancy Test ¹²	X					X
Randomize patient in RTSM		X				
Patient diary training	X	X				
Patient diary retraining			X	X	X	
Conventional on-demand treatment washout ¹³			X	X	X	
Dose ¹⁴			X	X	X	
IMP dispensing/return/accountability ¹⁵		X				X
PGI-S ¹⁶			X	X	X	
PGI-C ¹⁷			X	X	X	
VAS ¹⁸			X	X	X	
GA-NRS ¹⁹			X	X	X	
Concomitant Medication Review						

																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															</
--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Contact details

Authors Name: Kerstin Forsberg, Johannes Ulander

Company: data4Knowledge ApS

Address: Byledet 15, 2820 Gentofte, Danmark

Work Phone: +46 703203907, +46 703125461

Email: kf@data4knowledge.dk, ju@data4knowledge.dk

Website: <https://d4k.dk/>