

# Generative AI for Research Data Processing: Lessons Learnt From Three Use Cases



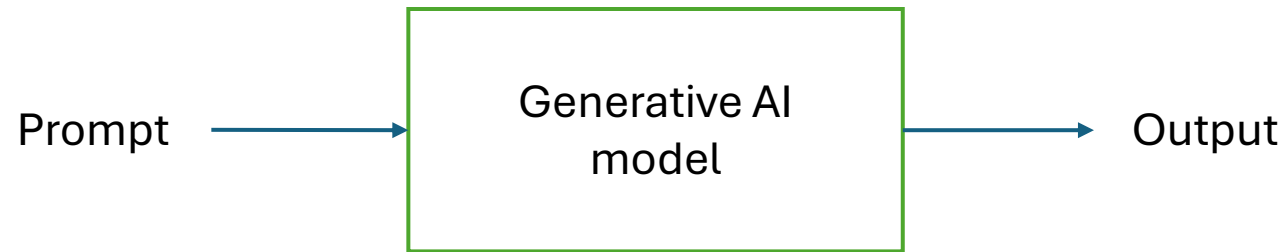
**Modhurita Mitra**

Research Engineering Team  
Research and Data Management Services  
Utrecht University  
The Netherlands




Work done jointly with Martine de Vos,  
Nicola Cortinovis, and Dawa Ometto

# Generative AI : General-purpose technology



Prompts in natural language → Low barrier to entry for domain researchers

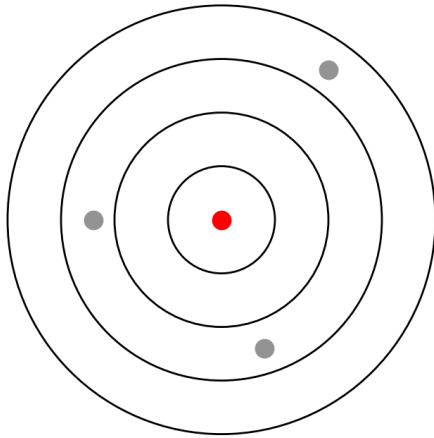


**Andrej Karpathy**  @karpathy · Jan 24, 2023

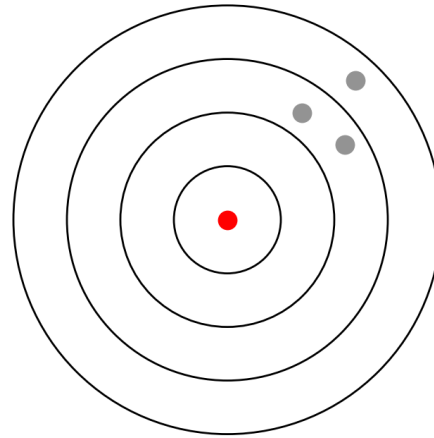
The hottest new programming language is English

# Goals of the study

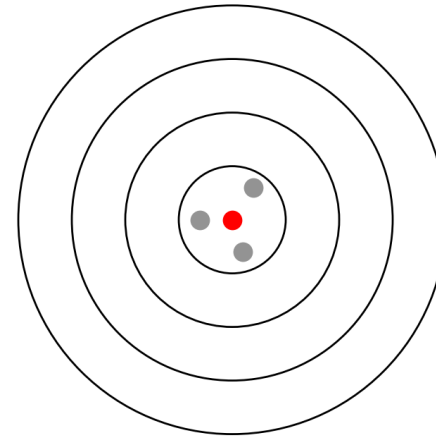
- Appropriateness
- Accuracy and consistency



Neither Accurate nor Consistent



Not Accurate but Consistent



Both Accurate and Consistent

# Use cases

- **Seedlists** (Botanical Gardens)
- **Health Technology Assessment (HTA) documents** (Pharmaceutical Sciences)
- **Kickstarter** (Economic Geography)

# Use case: Seedlists

- 7 -

Salvia austriaca L.	Scilla amoena Horn.	Serratula coronata L.	Silene vesperitia Retz.
— cicutaria L.	— italica L.	— heterophylla Desf.	— vicia Pers.
— dormans Lam.	— verna Huds.	— quicquid M. Bieb.	— zaidensis Herb.
— forskohlii Vahl.	Scirpus atrovirens Willd.	— radiata M. B.	— pubescens Lam.
— limosa Jacq.	Scirpus amurensis Schrad.	— scirpoides Ledeb.	— siamensis Lam.
— hispanica L.	Scorpioides alba Lam.	— gumifera Sm.	— brachycarpus Jacq.
— linaria C. A. M.	Scorpioides austriaca Willd.	— pilosella Benth.	— Camptocarpus F. et M.
— micrantha Vahl.	— humilis L.	— sphaerocarpa Ardl.	— hirsutum Lam.
— multifida Sibth.	— villosa Scop.	— tenuifolia Schrad.	— incanum Benth.
— pteris R. Br.	Scrophularia bilobata Horn.	— Sideris Benth.	— hirsutum Lam.
— pseudocrocata Jacq.	— canalis L.	— aurita Willd.	— junceum M. B.
— rhomboides R. et P.	— Scutellaria alpina L.	— crispula Lam.	— millefolium Ait.
— rubra Ehrh.	— hastata L.	— glauca Cav.	— perenne Spr.
— sylvatica Benth.	— peregrina L.	— mollissima Willd.	— polycephala Lam.
— verbasifolia M. Bieb.	— scorpioides Fisch.	— thymifolia Lam.	— strictissimum L.
— Verbena L.	— Selena congesta Schrad.	— Sellowiana Klotzsch.	— Zea Spreng.
— virgata Ait.	— Sclerum alba L.	— spinosa L.	— Sideriophora Bernadina L.
Sanguisorba media L.	— viridis L.	— strumosa L.	— iridifolia Hamb.
— tenuifolia Fisch.	— anagallis L.	— stratum L.	— vicia Pers.
— Sanicula maritima L.	— cephalanthus Cand.	— ramosa L.	— tenuifolia Hamb.
— Santolina incana Lam.	— coccineum Vahl.	— scorpioides L.	— Zuccopisium Dun.
— leucantha Benth.	— cruciatum Desf.	— taurica M. Bieb.	— Soldanella alpina L.
— pinnata Vitr.	— elegans Lej.	— Sideriophora Bernadina L.	— minima Hopp.
— aquatica Willd.	— Erythra Ledeb.	— Sideriophora Bernadina L.	— montana Willd.
— viridis Willd.	— Palustris Koch.	— Sideriophora Bernadina L.	— Solidago media Benth.
Sparganium angustifolium C. A. Meyer.	— Fortisimum Sm.	— Sideriophora Bernadina L.	— Solanum longifolium Trev.
— glaucum M. Bieb.	— globosum Mart.	— Sideriophora Bernadina L.	— tamariscus L.
— ocyroides L.	— hybridum L.	— Sideriophora Bernadina L.	— nigricans M. B.
— porteri Lam.	— involucreum M. B.	— Sideriophora Bernadina L.	— Sparganium angustifolium C. A. Meyer.
— vicia Pers.	— ochroleucum Sm.	— Sideriophora Bernadina L.	— Specularia perfoliata Jacq.
Sarracenia coccinea R. et P.	— oppositifolium Sims.	— Sideriophora Bernadina L.	— pentagona Cand.
— Sarracenia Schrad.	— pallidum M. B.	— Sideriophora Bernadina L.	— perfoliata A. DC.
Satureja cuneifolia Ten.	— populifolium Pall.	— Sideriophora Bernadina L.	— Spermocarpus articulata L. fil.
— intermedia C. A. M.	— pulchrum Michx.	— Sideriophora Bernadina L.	— hirsuta L.
— Thymus L.	— scordium M. Bieb.	— Sideriophora Bernadina L.	— stricta L. fil.
Saussurea alpina Cand.	— Sieboldii Mark.	— Sideriophora Bernadina L.	— tenuis L.
— elongata Cand.	— spiraea M. B.	— Sideriophora Bernadina L.	— Sphacelium campanulata Benth.
— pulchella Fisch.	— spuria L.	— Sideriophora Bernadina L.	— Sphacelium Chamaejasme F. et M.
— salicifolia Cand.	— Semperivum barbatum Sm.	— Sideriophora Bernadina L.	— cheiranthifolium F. et M.
Saxifraga aizoides L.	— githium Ait.	— Sideriophora Bernadina L.	— hirsuta F. et M.
— austriaca Lam.	— Smithii Sims.	— Sideriophora Bernadina L.	— latifolia Lam.
— bulbifera L.	— spathulatum Horn.	— Sideriophora Bernadina L.	— longifolia Ehrh.
— cernua L.	— tabulaforme Hux.	— Sideriophora Bernadina L.	— Muscivora Lam.
— controversa Sternb.	— tortuosum Ait.	— Sideriophora Bernadina L.	— obtusifolia Willd.
— cuneifolia L.	— Sutherlandia incana Cand.	— Sideriophora Bernadina L.	— alpina Scop.
— exarata Vill.	— Scutellaria glauca Gaertn.	— Sideriophora Bernadina L.	— arctica L.
— graminifolia L.	— Senecio abrotanifolia L.	— Sideriophora Bernadina L.	— campestris Cand.
— irrigua M. B.	— alpina Scop.	— Sideriophora Bernadina L.	— delphinifolia Vahl.
— leptophylla Pers.	— arctica L.	— Sideriophora Bernadina L.	— foveolatus Ten.
— muscivora Willd.	— campetris Cand.	— Sideriophora Bernadina L.	— gallica Vill.
— mutata L.	— delphinifolia Vahl.	— Sideriophora Bernadina L.	— leucanthemifolia Cand.
— alvula L.	— foveolatus Ten.	— Sideriophora Bernadina L.	— lychnifolia Reicheb.
— pedatifida Ehrh.	— gallica Vill.	— Sideriophora Bernadina L.	— purpurea Lam.
— peruviana Lam.	— leucanthemifolia Cand.	— Sideriophora Bernadina L.	— reclinata L. fil.
— planifolia Lapp.	— lychnifolia Reicheb.	— Sideriophora Bernadina L.	— rigida L.
— alvula L.	— purpurea Lam.	— Sideriophora Bernadina L.	— tritona Lam.
— stellata L.	— reclinata L. fil.	— Sideriophora Bernadina L.	— umbrosa W. et K.
— tomentosa Willd.	— rigida L.	— Sideriophora Bernadina L.	— venusta Ait.
— trifurcata Schrad.	— tritona Lam.	— Sideriophora Bernadina L.	— verbasifolia Jacq.
— Scabiosa grammata L.	— umbrosa W. et K.	— Sideriophora Bernadina L.	— bacifera C. A. M.
— Hylaeus Lam.	— venusta Ait.	— Sideriophora Bernadina L.	— corniculata C. A. Meyer.
— holosericea Benth.	— verbasifolia Jacq.	— Sideriophora Bernadina L.	
— nigrifolia Desf.	— glauca Tin.	— Sideriophora Bernadina L.	
— Schizanthus Walkeri Sims.	— Serratula caerulea Lam.	— Sideriophora Bernadina L.	
— Scholertia alba Ledeb.		— Sideriophora Bernadina L.	
— bacifera C. A. M.		— Sideriophora Bernadina L.	
— corniculata C. A. Meyer.		— Sideriophora Bernadina L.	

Göttingen, Germany, 1841

Caryopteris mastochanthus Schum.	Corchoropsis ornata Sieb. et Zucc.	Euphorbia speciosa L.
Cassia marylandica L.	Corchorus leucostylis Hort.	Forsythia viridis DC.
— americana L.	Coreopsis grandiflora Nutt.	Fernia communis L.
— leucostylis Hort.	Coriandrum sativum L.	Ficifolia floridana Moq.
— sinensis Juss.	Coronilla montana Scop.	Glycyrrhiza echinata L.
— salmantica L.	Crataegus crus-galli L.	Glandelia lanceolata Nutt.
— macrocephala Mous.	Crotalaria incana L.	Guizotia chlorifera DC.
— salmantica L.	Cucumis prophetarum L.	Gunnera scabra Retz.
— melanosperma Braun.	Cucurbita aurantia Willd.	Guthriea viscosa Regel.
— Dalea lagopus Willd.	— melanosperma Braun.	Helianthemum apenninum (L.) Mill.
— Darwinia brachyloba DC.	— rigens Jacq.	— var. roseum Grosser
— strata Catula L.	— Zuccopisium Dun.	— nummularium (L.) Mill.
— Delphinium Staphisagria	— minima Hopp.	— Tubularia guttata (L.) Fourr.
— Desmodium marylandicum Boott.	— montana Willd.	
— Dianthus monosperma Jacq.	— Solidago media Benth.	
— Digitalis lutea L.	— tamariscus L.	
— Dioscorea divaricata L.	— nigricans M. B.	
— Dolichos labialis L.	— Sparganium angustifolium C. A. Meyer.	
— Echinops spinosus L.	— Specularia perfoliata Jacq.	
— Enelia calva Gray.	— pentagona Cand.	
— Epilobium adnatum Griseb.	— perfoliata A. DC.	
— Erigeron speciosus DC.	— Spermocarpus articulata L. fil.	
— Eryngium planum L.	— hirsuta L.	
— Euphorbia Crista-galli L.	— stricta L. fil.	
— Euphorbia aromaticum L.	— tenuis L.	
— foefoliatum L.	— Sphacelium campanulata Benth.	
— Preslii Juss.	— Sphacelium Chamaejasme F. et M.	
	— cheiranthifolium F. et M.	
	— hirsuta F. et M.	
	— latifolia Lam.	
	— longifolia Ehrh.	
	— Muscivora Lam.	
	— obtusifolia Willd.	
	— alpina Scop.	
	— arctica L.	
	— campestris Cand.	
	— delphinifolia Vahl.	
	— foveolatus Ten.	
	— gallica Vill.	
	— leucanthemifolia Cand.	
	— lychnifolia Reicheb.	
	— purpurea Lam.	
	— reclinata L. fil.	
	— rigida L.	
	— tritona Lam.	
	— umbrosa W. et K.	
	— venusta Ait.	
	— verbasifolia Jacq.	
	— bacifera C. A. M.	
	— corniculata C. A. Meyer.	

Bordeaux, France, 1921

167	Chenopodium capitatum (L.) Aschers.
168	— ficifolium Sm.
169	— foliosum Aschers.
170	— opulifolium Schrad.
171	— schradonianum Schult.
172	— urticum L.
173	Corispermum leptopterum (Asch.) Iljin
174	Kochia scoparia (L.) Schrad.
<u>Cistaceae</u>	
175	Cistus crispus L.
176	— hirsutus Lam.
177	— monspeliensis L.
178	Helianthemum apenninum (L.) Mill.
179	— var. roseum Grosser
180	— nummularium (L.) Mill.
181	Tuberaria guttata (L.) Fourr.
<u>Cneoraceae</u>	
182	Cneorum tricoccon L.
<u>Commelinaceae</u>	
183	Ancilema papuanum Warburg
184	Commelina communis L.
185	— tuberosa L.
186	Palisota mannii C.B. Clarke
187	Tinantia erecta (Jacq.) Schlechtend.
+ 188	Tradescantia crassifolia Cav.
<u>Compositae (Asteraceae)</u>	
189	Achillea ageratifolia (Sibth. & Smith) Boiss.
190	— macrophylla L.
! 191	— millefolium L. (NG)
192	Ammobium alatum R.Br.
193	Antennaria howellii Greene
194	— plantaginea R.Br.
195	Anthemis arvensis L.
196	— austriaca Jacq.
197	— cotula L.
! 198	— tinctoria L. (NG)

Göttingen, Germany, 1970



# Seedlists

Bestell- Nr. / Order Nr.	Name	Herkunft u. Art der Diasporen / origin and type of diaspores	IPEN	Aufsammlungs- daten / collecting data
<b>Asteraceae (cont.)</b>				
30	<i>Centaurea jacea</i> L.	WS	AT-0-WU-0030521	42
31	-- <i>jacea</i> subsp. <i>angustifolia</i> (DC.) Gremli	WS	AT-0-WU-0031475	117
32	-- <i>pseudophrygia</i> C. A. Mey.	WS	AT-0-WU-0027271	125
33	-- <i>scabiosa</i> L.	WS	AT-0-WU-0030447	37
34	-- <i>scabiosa</i> L.	WS	AT-0-WU-0031441	95
35	<i>Cirsium cianum</i> (L.) All.	WS	AT-0-WU-0031471	121
36	-- <i>erichonum</i> (L.) Scop.	WKS	AT-0-WU-0007975	125
37	-- <i>eristhales</i> (Jacq.) Scop.	WS	AT-0-WU-0031498	114
38	-- <i>pannonicum</i> (L. f.) Link	WS	AT-0-WU-0030486	37
39	-- <i>spinossissimum</i> (L.) Scop.	WS	AT-0-WU-0030577	53
40	<i>Crepis pyrenaica</i> (L.) Greuter	WS	AT-0-WU-0030527	45
41	-- <i>pyrenaica</i> (L.) Greuter	WS	AT-0-WU-0031494	114
42	<i>Cyanus segetum</i> Hill	WS	AT-0-WU-0031457	99
43	<i>Echinops ritro</i> subsp. <i>ruthenicus</i> (M. Bieb.) Nyman	WS	AT-0-WU-0030642	70
44	-- <i>ritro</i> subsp. <i>ruthenicus</i> (M. Bieb.) Nyman	WKS	AT-0-WU-0031014	125
45	<i>Erigeron acris</i> L.	WS	AT-0-WU-0030619	65
46	-- <i>glabellus</i> Nutt.	KS	XX-0-WU-0022433	125
47	<i>Eupatorium cannabinum</i> L.	KS	XX-0-WU-0025821	125
48	<i>Filago anensis</i> L.	WS	AT-0-WU-0031534	125
49	-- <i>vulgaris</i> Lam.	WS	AT-0-WU-0031367	111
50	<i>Galatella cana</i> (Waldst. & Kit.) Nees	WS	AT-0-WU-0030647	75
51	-- <i>linosyris</i> (L.) Rchb. f.	WS	AT-0-WU-0030644	72
52	<i>Gnaphalium sylvaticum</i> L.	WS	AT-0-WU-0029392	16
53	<i>Helenium aromaticum</i> (Hook.) L. H. Bailey	KS	XX-0-WU-0031245	125
54	<i>Hieracium prenanthoides</i> Vill.	KS	XX-0-WU-0030679	125
55	-- <i>umbellatum</i> L.	KS	XX-0-WU-0002204	77
56	-- <i>umbellatum</i> L.	WKS	AT-0-WU-0008873	125
57	-- <i>velutinum</i> Hegetschw.	WKS	CH-0-WU-0009851	125
58	-- <i>villosum</i> Jacq.	WKS	FR-0-WU-0008195	125
59	<i>Homogyne sylvestris</i> Cass.	WS	AT-0-WU-0029414	18
60	<i>Inula ensifolia</i> L.	WS	AT-0-WU-0030639	70
61	-- <i>oculus-christi</i> L.	WS	AT-0-WU-0030435	36
62	<i>Jurinea mollis</i> (L.) Rchb.	WS	AT-0-WU-0031346	91
63	<i>Klasea quinquefolia</i> (Willd.) Greuter & Wagenitz	KS	XX-0-WU-00319038	125
64	<i>Leontodon incanus</i> (L.) Schrank	WS	AT-0-WU-0027771	1
65	-- <i>pyrenaicus</i> subsp. <i>helveticus</i> (Mérat) Finch & P. D. Sell	WKS	FR-0-WU-0009648	125
66	<i>Leucanthemopsis alpina</i> (L.) Heywood	WS	AT-0-WU-0030566	53
67	<i>Liatris cylindracea</i> Michx.	WKS	NZ-0-MO-2015-0204	125
68	<i>Onopordum acanthium</i> L.	WS	AT-0-WU-0031459	96
69	<i>Prenanthes purpurea</i> L.	WKS	CH-0-WU-0007600	125
70	-- <i>purpurea</i> L.	WS	AT-0-WU-0031502	114
71	<i>Pseudognaphalium luteoalbum</i> (L.) Hilliard & B. L. Burt	KS	XX-0-WU-0031535	125

## PARS I

### SEMINA PLANTARUM IN LOCO NATALI LECTA

#### MAGNOLIOPHYTA MAGNOLIOPSIDA

##### ADOXACEAE

1. *Viburnum opulus* L. – Vitosha Mts, near Zlatni mostove locality, 19. 09. 2020.
2. *Sambucus racemosa* L. – Rila Mt, near Borovetz resort, 13. 08. 2020.

##### ANACARDIACEAE

3. *Pistacia terebinthus* L. – Managed reserve Izgoryaloto Gijune, near Krichim town, 12. 09. 2020.

##### ASTERACEAE

4. *Inula ensifolia* L. - Nature Park Zlatni pyasatsi, near Varna town, 28. 08. 2020.
5. *Jacobaea paludosa* (L.) G. Gaertn., B. Mey. & Scherb. (syn. *Senecio paludosus* L.) – Rodopi Mt, Tsigov chark locality, 16. 10. 2019.
6. *Tanacetum corymbosum* (L.) Sch. Bip. – Nature Park Zlatni pyasatsi, near Varna town, 28. 08. 2020.

##### BRASSICACEAE

7. *Alliaria petiolata* (M. Bieb.) Cavara & Grande – near Sotrya village, Stara Planina Mt., 05. 09. 2020.
8. *Lunaria annua* L. – Managed reserve Izgoryaloto Gijune, near Krichim town, 12. 09. 2020.

##### CAMPANULACEAE

9. *Campanula sibirica* L. – Nature Park Zlatni pyasatsi, near Varna town, 28. 08. 2020.

##### CAPRIFOLIACEAE

10. *Scabiosa argentea* L. – Nature Park Zlatni pyasatsi, near Varna town, 28. 08. 2020.

##### CARYOPHYLLACEAE

11. *Silene atropurpurea* (Griseb.) Greuter & Burdet (syn. *Viscaria atropurpurea* Griseb.), – near Dolen village, Western Rhodopes Mt., 31. 08. 2020.
12. *Silene atropurpurea* (Griseb.) Greuter & Burdet (syn. *Viscaria atropurpurea* Griseb.) – Samokov plane, near Samokov town, 13. 08. 2020.

##### FABACEAE

13. *Colutea arborescens* L. – Managed reserve Izgoryaloto Gijune, near Krichim town, 12. 09. 2020.

14. *Hippecrepis emerus* (L.) Lassen (syn. *Coronilla emerus* L.) – Managed reserve Izgoryaloto Gijune, near Krichim town, 12. 09. 2020.

15. *Trifolium angustifolium* L. – Managed reserve Izgoryaloto Gijune, near Krichim town, 12. 09. 2020.

##### LAMIACEAE

16. *Clinopodium vulgare* L. – Managed reserve Izgoryaloto Gijune, near Krichim town, 12. 09. 2020.
17. *Salvia sclarea* L. – River Struma valley, near Simidi Town 30. 08. 2020

##### OLEACEAE

18. *Fraxinus ornus* L. – Managed reserve Izgoryaloto Gijune, near Krichim town, 12. 09. 2020.
19. *Syringa vulgaris* L. – Managed reserve Izgoryaloto Gijune, near Krichim town, 12. 09. 2020.

##### PLANTAGINACEAE

20. *Digitalis lanata* Ehrh. - Managed reserve Izgoryaloto Gijune, near Krichim town, 12. 09. 2020

##### PRIMULACEAE

21. *Primula veris* L. – near Dolen village, Rhodopes Mt., 30. 08. 2020.

##### RHAMNACEAE

22. *Paliurus spina-christi* Mill. - Managed reserve Izgoryaloto Gijune, near Krichim town, 12. 09. 2020

##### SOLANACEAE

23. *Physalis alkekengi* L. - near Dolen village, Rhodopes Mt., 30. 08. 2020.

##### LILIOPSIDA

##### ALLIACEAE

24. *Allium flavum* L. – near Luki village, Central Rhodopes Mt. 30. 08. 2020.

##### IRIDACEAE

25. *Iris sibirica* L. – Vitosha Mt., 31. 08. 2020.

##### POACEAE

26. *Danthonia alpina* Vest (syn. *Danthonia calycina* Roem. & Schult.) – Vitosha Mt., 31. 08. 2020.

##### XANTHORRHOACEAE

27. *Asphodeline lutea* (L.) Rchb. – Nature Reserve Kaliakra, 30. 10. 2020.
28. *Asphodelus albus* Mill. – Vitosha Mt., 31. 08. 2020.

Collectors:

Antoaneta Petrova, PhD  
Diana Venkova

Nr	Familie	Gattung	Species (Subspecies, Varietät)	Autor	IPEN
76	Asteraceae	Liatris	elegans	Michx.	XX-0-LI-7086
77	Asteraceae	Liatris	punctata	Hook.	XX-0-LI-10164
78	Asteraceae	Liatris	spicata	Willd.	XX-0-LI-7094
79	Asteraceae	Liatris	spicata	Willd.	XX-0-LI-10268
80	Asteraceae	Onopordum	algeriense	Pomel	XX-0-LI-6994
81	Asteraceae	Onopordum	bracteatum	Boiss. & Heldr.	XX-0- STGAL-130/1981
82	Asteraceae	Osteospermum	hyoseroides	(DC.) Norl.	XX-0-LI-14339
83	Asteraceae	Pilosella	aurantiaca	(L.) F.W. Schultz & Sch. Bip.	XX-0-LI-11633
84	Asteraceae	Scorzonera	suberosa ssp. carlensis	(Boiss.) D.F. Chamb.	XX-0-LI-13055
85	Asteraceae	Senecio	alpinus	Scop.	XX-0-LI-7426
86	Asteraceae	Silphium	perfoliatum	L.	XX-0-LI-10168
87	Asteraceae	Solidago	virgaurea	L.	XX-0-LI-14222
88	Asteraceae	Stockesia	laevis	Greene	XX-0-LI-7111
89	Asteraceae	Tanacetum	corymbosum	(L.) Sch. Bip.	IT-0-SIENA-A-10
90	Asteraceae	Telekia	speciosa	(Schreb.) Baumg.	XX-0-LI-8899A
91	Berberidaceae	Berberis	koreana	PALIB.	XX-0-LI-7898
92	Bignoniaceae	Incarvillea	delavayi	Bureau et Franch	XX-0-LI-6808
93	Boraginaceae	Echium	vulgare	L.	XX-0-LI-8987
94	Boraginaceae	Onosma	ablorosea	Fisch. et C.A. Mey	XX-0-LI-6693
95	Brassicaceae	Alyssum	altaicum	C.A. Mey.	XX-0-LI-6619
96	Brassicaceae	Alyssum	ovirense	A. Kern.	XX-0-LI-7290
97	Brassicaceae	Alyssum	wulfenianum	Bernh.	XX-0-LI-6461
98	Brassicaceae	Arabis	blepharophylla	Hook. et Arn.	XX-0-LI-6398

Vienna, Austria, 2020

Sofia, Bulgaria, 2020

Linz, Austria, 2020

# Seedlists

## Research Engineering task

- Extract plant species names from seedlists

## Scientific goal of the project

- Create database
- Analysis:
  - Curational policies
  - Effects of climate change



# Prompt

You are an expert botanist with deep knowledge of Latin names in binomial nomenclature of plant species.

You have knowledge of plant family names, subspecies, varieties and forms.

The species name in binomial nomenclature consists of two parts: the genus name starting with a capital letter, followed by the epithet name which normally starts with a small letter.

This is often followed by a subspecies name (starting with "subsp. " or "ssp. "), or variety name (starting with "var."), or form (starting with "f.") and the author name(s). Sometimes this is followed by a synonym (starting with a syn.).

Some species names are also followed by a cultivar name (almost always within single quotation marks or with the abbreviation "CV" if the cultivar has no formal name).

The author name is often simply "L.", for "Linnaeus". For other authors it is often an abbreviation describing the author name.

Thus the species name consists of the genus, epithet, subspecies name (if present), variety name (if present), form name (if present), cultivar name (if present), author name(s) (if present) and synonym (if present).

You are given the text from a page from a seedlist of plants from a botanical garden.

Determine if there are any Latin names of plant species present in this text.

Find each and every plant species entry present in the text.

Do not skip or miss any entry.

If any plant species names are present in this text, find, for each plant species present in the text, the species name in Latin, consisting of the genus, epithet, subspecies name (if present), variety name (if present), form name (if present), cultivar name (if present), author name(s) (if present), and synonym (if present).

Return the result for each plant species in valid JSON object format, with the single key "species".

Provide the complete results as a list of valid JSON objects.

If there is no information about plant species on this page, print an empty list, like this: []

Print only the complete list of valid JSON objects, and no other code, text, or explanation.

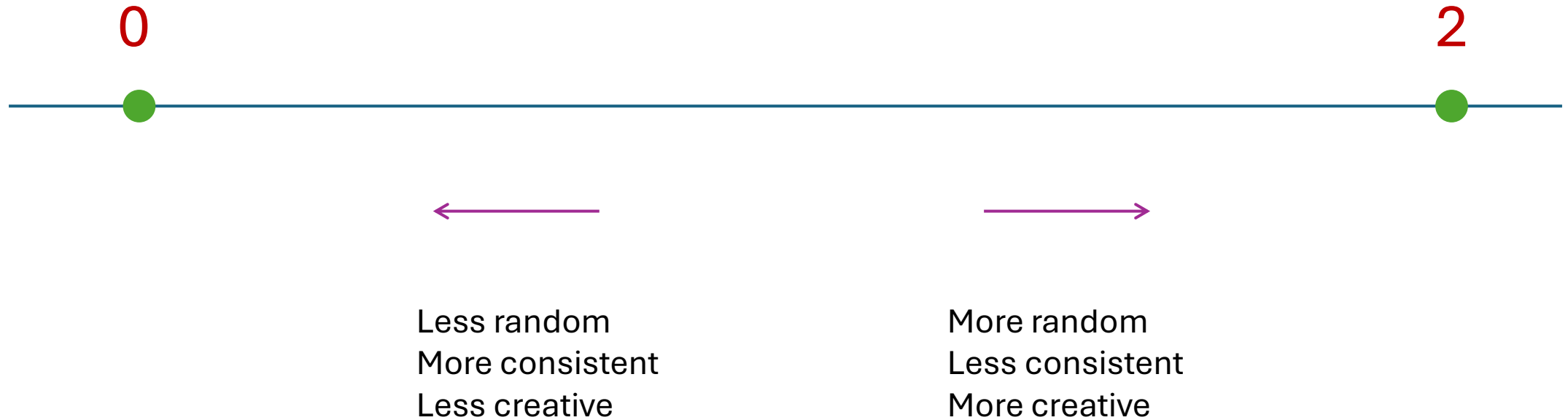
Set  
AI persona

Provide  
background  
information

Describe  
task

Specify  
output format

# Temperature parameter



Temperature = 0 : Maximises both accuracy and consistency

# Seedlists: Results

[illegible]

42 species

*PARS I*  
*SEMINA PLANTARUM*  
*IN LOCO NATALI LECTA*

MAGNOLIOPHYTES  
MAGNOLIOPSIDA

**ADOXACEAE**  
1. *Wibsonia cyathar* L. – V  
moste locality, 19.09.2

**ANACARDIACEAE**  
*A. Pistacia terebinthus* L. –

**ASTERACEAE**  
4. *Aster eriocephalus* L. - Nana  
Varna town, 28.08.2020.

6. *Taxacetum corymbosum*  
Park Ziatai pyasoi, near

8. *Lamarca crassa* L. — Mus. Glouc. near Krishnagar.

9. *Carex ovata* L. -  
pyrasol, near Varna town

**CAPRIEOLIACEAE**

**CARYOPHYLLACEA**  
11. *Silene atropurpurea* (C.

12. *Silene atropurpurea* (C.  
(syn. *Viscaria atropurpurea*)  
plane, near Samokov town

1A. *Colutea arborescens* L.  
Igorjako Gane, near K.

14. *Hypocrepis emerus* (L.) Lassen (syn. *Coronilla emerus* L.) – Managed reserve Izgorjalno Gajne, near Krčin town, 12.09.2020.

15. *Trifolium angustifolium* L. – Managed reserve Izgorjalno Gajne, near Krčin town, 12.09.2020.

**16.** *Clasopodites vulgare* L. – Managed reserve  
Irgayakoto Gijano, near Krichin town, 12.09.2020.

**OLEACEAE**  
**18.** *Fraxinus ornus* L. – Мангуст резерв-Ізгорьскіх  
 Гілля, near Krichim town, 12.09.2020.

**PLANTAGINACEAE**  
**20.** *Digitalis lanata* Ehrh. - Managed source

**PRIMULACEAE**  
21. *Primula veris* L. – near Dolen village, Rhodopes  
Mt., 30. 08. 2021.

22. *Passiflora spiro-chrysi* Mill. - Managed reserve  
Igarka Gorge, near Krichim town, 12.09.2020

**LILIOPSIDA**  
**ALLIACEAE**

**IRIDACEAE**  
25. *Iris sibirica* L. – Vinzha Mt., 31.08.2020.

26. *Danthonia alpina* Vest. (syn. *Danthonia calycina* Roem. & Schult.) – Vinohr. Mt., 31.08.2020.

28. *Asphodelus albus* Mill. – Vaukha Mt., 31.08.2020

Collectors:  
Antoaneta Petrova, PhD  
Diana Mardanova

Nº	Familia	Gattung	Species (Subspecies, Variety)	Author	IPEN
75	Asteraceae	Liatris	elegans	Hook.	XX-01-L-7096
77	Asteraceae	Liatris	punctata	Michx.	XX-01-L-10164
78	Asteraceae	Liatris	spicata	Willd.	XX-01-L-70194
79	Asteraceae	Liatris	spicata	Willd.	XX-01-L-10208
80	Asteraceae	Oncopeltum	algeriense	Pomel	XX-01-L-6994
81	Asteraceae	Oncopeltum	bracteatum	Bolles & Hildebr.	XX-01-STGAL-130-1198
82	Asteraceae	Osteospermum	hyoscyroides	(DC.) Nees	XX-01-L-14339
83	Asteraceae	Pilosella	aurantiaca	(J.F.W.) Schult. (Sch. Bip.)	XX-01-L-11633
84	Asteraceae	Senecio	suberosa ssp. arvensis	D.F. Sauer	XX-01-L-13005
85	Asteraceae	Senecio	alpinus	Schreb.	XX-01-L-74265
86	Asteraceae	Silphium	perfoliatum	L.	XX-01-L-10168
87	Asteraceae	Solidago	virgaurea	L.	XX-01-L-14222
88	Asteraceae	Stoeckia	laevis	Willd.	XX-01-L-17111
89	Asteraceae	Tanacetum	corymbosum	G. Schreb.	(T-0 SIENA-A-11)
90	Asteraceae	Teliskia	speciosa	(Schreb.) Schleg.	XX-01-L-8999A
91	Berberidaceae	Berberis	koranra	(Pav.) Willd.	XX-01-L-7898
92	Digonicaceae	Incarvillea	delavayi	Bureau et Franch.	XX-01-L-6808
93	Gramineaceae	Echium	vilgare	L.	XX-01-L-8987
94	Gramineaceae	Oxonia	albiflora	Flach et A.C. Meyer	XX-01-L-6683
95	Brassicaceae	Alyssum	alticorne	A.C.Mey.	XX-01-L-6819
96	Brassicaceae	Alyssum	ovatum	A.C.Mey.	XX-01-L-7290
97	Brassicaceae	Alyssum	wulfenianum	Bernh.	XX-01-L-6481
98	Brassicaceae	Arabis	blepharophylla	Hook. et Arn.	XX-01-L-6308

23 species

[illegible]

32 species

- All species names were extracted, and were extracted correctly
- This was consistent over 3 runs
- Even OCR errors in the scanned seedlist were corrected!

# Use case: Health Technology Assessment (HTA) documents

**NICE** National Institute for Health and Care Excellence

**HAS**  
HAUTE AUTORITÉ DE SANTÉ



Zorginstituut Nederland

Ivabradine for treating chronic heart failure (TA267)

## 1 Recommendations

- 1.1 Ivabradine is recommended as an option for treating chronic heart failure for people:
- with New York Heart Association (NYHA) class II to IV stable chronic heart failure with systolic dysfunction and
  - who are in sinus rhythm with a heart rate of 75 beats per minute (bpm) or more and
  - who are given ivabradine in combination with standard therapy including beta-blocker therapy, angiotensin-converting enzyme (ACE) inhibitors and aldosterone antagonists, or when beta-blocker therapy is contraindicated or not tolerated and
  - with a left ventricular ejection fraction of 35% or less.
- 1.2 Ivabradine should only be initiated after a stabilisation period of 4 weeks on optimised standard therapy with ACE inhibitors, beta-blockers and aldosterone antagonists.

### 1. CARACTERISTIQUES DU MEDICAMENT

**1.1. Principe actif**  
Ivabradine

**1.2. Indications**

« Traitement de la maladie coronaire :

Traitement symptomatique de l'angor stable chronique chez l'adulte coronarien en rythme sinusal. L'ivabradine est indiquée :

- chez les adultes présentant une intolérance ou une contre-indication aux bêtabloquants,
- ou en association aux bêtabloquants chez des patients insuffisamment contrôlés par une dose optimale de bêtabloquants, et dont la fréquence cardiaque reste supérieure à 60 bpm ».

Traitement de l'insuffisance cardiaque : nouvelle indication

L'ivabradine est indiquée dans le traitement de l'insuffisance cardiaque chronique de classe NYHA II à IV avec dysfonction systolique, chez les patients en rythme sinusal et dont la fréquence cardiaque est supérieure ou égale à 75 bpm, en association au traitement standard comprenant les bêtabloquants, ou en cas de contre-indication ou d'intolérance aux bêtabloquants (cf Pharmacodynamie) ».

**1.3. Posologie**

« Pour les différentes doses thérapeutiques, l'ivabradine est disponible sous forme de comprimés pelliculés dosés à 5 mg et 7,5 mg.

Traitement de l'insuffisance cardiaque chronique :

Le traitement doit être initié uniquement chez les patients atteints d'insuffisance cardiaque stable. Il est recommandé que le médecin ait une expérience de la prise en charge de l'insuffisance cardiaque chronique.

La posologie initiale habituellement recommandée est de 5 mg d'ivabradine deux fois par jour. Après deux semaines de traitement, la dose peut être augmentée à 7,5 mg deux fois par jour si la fréquence cardiaque de repos reste de façon persistante au dessus de 60 bpm, ou diminuée à 2,5 mg deux fois par jour (soit un demi comprimé dosé à 5 mg deux fois par jour) si la fréquence cardiaque de repos est de façon persistante en dessous de 50 bpm ou en cas de symptômes liés à la bradycardie tels que sensations vertigineuses, fatigue ou hypotension. Si la fréquence cardiaque est comprise entre 50 et 60 bpm, la posologie de 5 mg deux fois par jour peut être maintenue.

Si, durant le traitement, la fréquence cardiaque de repos descend de façon persistante en dessous de 50 bpm ou en cas de symptômes liés à la bradycardie, la dose doit être diminuée à la posologie inférieure chez les patients recevant 7,5 mg ou 5 mg deux fois par jour.

### Farmacotherapeutisch rapport ivabradine (Procoralan®) bij de indicatie chronisch hartfalen

De Commissie Geneesmiddelen (CG) heeft een farmacotherapeutisch rapport vastgesteld voor het geneesmiddel ivabradine (Procoralan®) filmomhulde tabletten. Voor de bepaling van de therapeutische waarde is vergeleken met placebo. Hierbij is zij tot onderstaande conclusies gekomen.

- Ivabradine heeft een therapeutische meerwaarde ten opzichte van placebo bij patiënten met chronisch hartfalen NYHA klasse II-IV in sinusritme met systolische disfunctie met rusthartfrequentie  $\geq 75$  spm wanneer toegevoegd aan een individueel, optimaal en richtlijnconform ingestelde medicamenteuze behandeling die in principe een ACE-remmer, diureticum, bètablokker en aldosteronantagonist omvat tenzij de patiënt hiervoor niet is aangewezen, een contra-indicatie heeft of deze niet verdraagt.

#### Ivabradine (Procoralan®), filmomhulde tablet

**Geregistreerde indicatie.** « (...) chronisch hartfalen NYHA-klasse II-IV met systolische disfunctie, bij patiënten in sinusritme en bij wie de hartfrequentie  $\geq 75$  spm is, in combinatie met standaardbehandeling, inclusief behandeling met bètablokkers, of wanneer behandeling met bètablokkers is gecontra-indiceerd of niet wordt getolereerd. »

**Dosering.** 5 mg 2dd. Na een behandelingsduur van twee weken kan de dosis worden verhoogd naar 7,5 mg 2dd wanneer de hartfrequentie in rust voortdurend boven de 60 spm is of worden verlaagd tot 2dd 2,5 mg wanneer de hartfrequentie in rust voortduren onder de 50 slagen per minuut spm is.

**Werkingsmechanisme.** Ivabradine werkt via selectieve en specifieke remming van de cardiale pacemaker  $I_f$ -stroom, welke de spontane diastolische depolarisatie in de sinusknop en de hartfrequentie bepaalt. De cardiale effecten zijn specifiek voor de sinusknop zonder effecten op intra-atriale, atrioventriculaire of intraventriculaire geleidingstijden, evenmin als op myocardiale contractiliteit of ventriculaire repolarisatie.

# HTA documents: Attributes to extract

- Generic name of drug
- Brand name of drug
- Medical indication
- Relative effectiveness
- Cost effectiveness
- ...

# HTA documents

## Research Engineering task

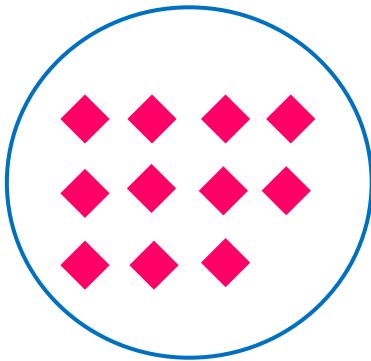
- Extract certain data points of interest from HTA documents

## Scientific goal of the project

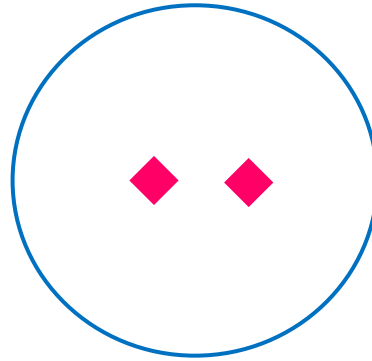
- Create Open Science database that can be used by other researchers, policymakers, etc.

# HTA documents: Results

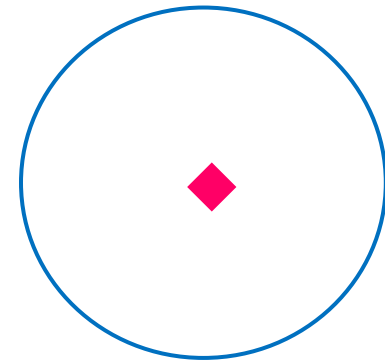
- HTA documents about one drug-indication combination, in **English, Dutch, French**.
- **14 data points** to be extracted.



All languages:  
**11** accurate and consistent  
over all 3 runs



Dutch: **2** semantically accurate  
but lexically different,  
over 3 runs



French: **1** consistent  
over 3 runs, but inaccurate



# Use case : Kickstarter

**KICKSTARTER**

Q Search projects, creators, and categories

Start a project


Art Comics Crafts Dance Design Fashion Film Food Games Journalism Music Photography Publishing Technology Theater Discover

Bring a creative project to life.

ON KICKSTARTER:

264,804 projects funded	\$8,227,855,974 towards creative work	97,266,101 pledges
----------------------------	--	-----------------------

FEATURED PROJECT




**Swedish Machines – New Simon Stålenhag Book**  
Free League  
2 days left • 1028% funded


A new art book from a critically acclaimed illustrator and author, set in a beautiful dystopian world in which two young men set out to explore a mysterious forbidden zone together.

Illustration Stockholm, Sweden


RECOMMENDED FOR YOU




**Valuchi Expedition...**  
Valuchi  
24 days left • 246%...



**Tir na nÓg Reprint...**  
Grand Gamers Guild  
3 days left • 6126%...



**THE DISCOVERIE ...**  
Black Letter Press  
5 days left • 415%...



**Bulletins at The En...**  
Russell Nohetty  
9 days left • 133%...

< 1 2 3 >

# Kickstarter: NAICS codes

← ↻ 🔒 https://www.census.gov/naics/?58967?yearbck=2017

An official website of the United States government

**Census**  
BUREAU OF ECONOMIC ANALYSIS

## North American Industry Classification System

Main History Development Partners Federal Register Notices FAQs Reference Files ▾ NAPCS

### NAICS Search

Enter keyword or 2-6 digit code

2022 NAICS Search Go

Enter keyword or 2-6 digit code

2017 NAICS Search Go

Enter keyword or 2-6 digit code

2012 NAICS Search Go

### 2017 NAICS

The following table provides detailed information on the structure of NAICS.

Sector	Definition
11	Agriculture, Forestry, Fishing and Hunting
21	Mining, Quarrying, and Oil and Gas Extraction
22	Utilities
23	Construction
31-33	Manufacturing
42	Wholesale Trade
44-45	Retail Trade
48-49	Transportation and Warehousing
51	Information
52	Finance and Insurance
53	Real Estate and Rental and Leasing
54	Professional, Scientific, and Technical Services
55	Management of Companies and Enterprises
56	Administrative and Support and Waste Management and Remediation Services
61	Educational Services
62	Health Care and Social Assistance
71	Arts, Entertainment, and Recreation
72	Accommodation and Food Services
81	Other Services (except Public Administration)
92	Public Administration

## Sector 51--Information<sup>T</sup>

### 511 Publishing Industries (except Internet)<sup>T</sup>

#### 5111 Newspaper, Periodical, Book, and Directory Publishers<sup>T</sup>

51111 Newspaper Publishers<sup>T</sup>

511110 Newspaper Publishers

51112 Periodical Publishers<sup>T</sup>

511120 Periodical Publishers

51113 Book Publishers<sup>T</sup>

511130 Book Publishers

51114 Directory and Mailing List Publishers<sup>T</sup>

511140 Directory and Mailing List Publishers

51119 Other Publishers<sup>T</sup>

511191 Greeting Card Publishers

511199 All Other Publishers

#### 5112 Software Publishers<sup>T</sup>

51121 Software Publishers<sup>T</sup>

511210 Software Publishers

Industry code

Number of 4-digit industry codes: 311

# Kickstarter

## Research Engineering task

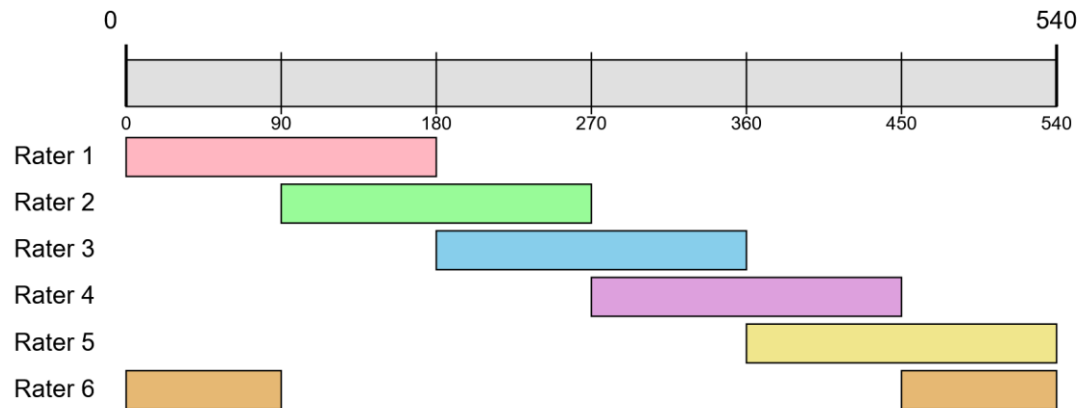
- Assign a NAICS code to every Kickstarter project

## Scientific goal of the project

- Assess whether crowdfunding fosters innovation and economic development

# Kickstarter: Results

- Sample of 540 representative projects, assigned to six human raters in a staggered manner, such that each project is rated by exactly 2 raters



- Highest fraction of industry codes that matched between:
  - Generative AI and a (single) human rater: 53%
  - Two human raters: 60%
- Performance of generative AI broadly comparable to that of a human rater

# Insights derived

# When is generative AI an appropriate tool?

- Amount of data to be processed is large
- No simple, rule-based method is available
- Results are of sufficiently high quality for the research purpose
- Good tool for tasks that are easy for a human to perform, but difficult for a computer
  - Heterogeneous, unstructured data

# Maximising accuracy and consistency

- **Temperature:**
  - Set to 0 to maximise accuracy and consistency
- **Prompt:**
  - Well-posed, unambiguous, clear
  - **Prompt engineering:** Iterate over many different versions of the prompt, with a representative input dataset
  - Prompt should be **general as well as specific:**
    - General wrt **data**
    - Specific wrt **AI's actions**



# Maximising accuracy and consistency

- **User** needs to have:
  - Good understanding of the subject matter (**domain knowledge**)
  - High degree of **familiarity with the data**
- **Data** needs to be:
  - Able to be divided into **chunks**, so that the input and output satisfy context length constraints
- **Generative AI model**:
  - Must have had **similar domain data** to train on



Thank you!  
Questions?  
[m.mitra@uu.nl](mailto:m.mitra@uu.nl)



IEEE International Conference on eScience 2024 paper:  
<https://ieeexplore.ieee.org/document/10678704>



← GitHub repository for data, prompts, code:  
<https://github.com/UtrechtUniversity/generative-ai>



**Utrecht  
University**

Sharing science,  
*shaping tomorrow*

# Prompt engineering

The screenshot displays the ANTHROPIC Workbench interface. At the top, navigation tabs include Dashboard, Workbench (selected), and Settings. On the right, there are links for Docs and Feedback, and a user profile icon labeled MM.

The main workspace is titled "Plant species names 1" with a version indicator "v1". It features a "Prompt" tab and a "Run" button with a keyboard shortcut "Ctrl + ⌘".

The interface is divided into two main sections:

- Prompt Section:**
  - SYSTEM PROMPT:** "You are an expert botanist."
  - EXAMPLES:** A section for providing context or examples.
  - USER:** Contains a list of plant species names: "Chenopodium capitatum (L.) Aschers», 168 - \* ficifolium Sn., ■169 \_ foliosum Asciesse., 170 ~ opulifolium Schaad», 171 - schraderianum Schuit», 172 - urbicum L», 175 Corispermum leptopterum (Asch.) Iljin". Below this list, a red box highlights the instruction: "Extract the plant species names from the text above."
- Response Section:**
  - RESPONSE:** Displays the output of the model, listing the extracted species names: "1. Chenopodium capitatum, 2. Chenopodium ficifolium, 3. Chenopodium foliosum, 4. Chenopodium opulifolium, 5. Chenopodium schraderianum, 6. Chenopodium urbicum, 7. Corispermum leptopterum".

At the bottom of the prompt section, there are options: "Pre-fill response", "Add message pair", "Improve prompt", and a "Beta" badge.

At the bottom of the response section, there is an "Add to Conversation" button and a status bar showing "120" and "93".

# Prompt engineering

The screenshot displays the ANTHROPIC Workbench interface. At the top, navigation links include Dashboard, Workbench (active), Settings, Docs, Feedback, and a user profile icon (MM). The main workspace is titled 'Plant species names 2 v1' and features tabs for 'Prompt' and 'Evaluate'. A 'Run' button with a play icon and 'Ctrl + R' shortcut is visible.

The 'SYSTEM PROMPT' section contains the instruction: 'You are an expert botanist.'

The 'EXAMPLES' section is currently empty.

The 'USER' section contains the following text:

```
Chenopodium capitatum (L.) Aschers»  
168 - * ficifolium Sn.  
■169 _ foliosum Asciesse.  
170 ~ opulifolium Schaad»  
171 - schraderianum Schuit»  
172 - urbicum L»  
175 Corispermum leptopterum (Asch.) Iljin
```

Below the user input, the instruction reads: 'Extract the plant species names from the text above.'

A red rectangular box highlights the instruction: 'Print only the plant species names, and no other additional text or explanation.'

The 'RESPONSE' section shows the output of the model:

```
Chenopodium capitatum  
Chenopodium ficifolium  
Chenopodium foliosum  
Chenopodium opulifolium  
Chenopodium schraderianum  
Chenopodium urbicum  
Corispermum leptopterum
```

At the bottom of the interface, there are buttons for 'Pre-fill response', 'Add message pair', 'Improve prompt', and 'Beta'. The bottom right corner shows 'Add to Conversation', a thumbs up icon, a thumbs down icon, and a count of 68.

# Prompt engineering

The screenshot displays the ANTHROPIC Workbench interface. At the top, navigation links include Dashboard, Workbench (active), Settings, Docs, Feedback, and a user profile icon labeled MM. The main workspace is titled "Plant species names 3" with a version indicator "v1". It features tabs for "Prompt" and "Evaluate", along with icons for undo, redo, and code editing. A prominent orange "Run" button with a play icon and the keyboard shortcut "Ctrl + R" is visible.

The interface is divided into two main panels. The left panel, labeled "SYSTEM PROMPT", contains the following text:

SYSTEM PROMPT  
You are an expert botanist.

Below this is an "EXAMPLES" section, followed by a "USER" section containing a list of plant species names and their authors, along with a request to extract the names:

Chenopodium capitatum (L.) Aschers»  
168 - \* ficifolium Sn.  
■169 \_ foliosum Asciesse.  
170 ~ opulifolium Schaad»  
171 - schraderianum Schuit»  
172 - urbicum L»  
175 Corispermum leptopterum (Asch.) Iljin

Extract the plant species names from the text above.

A specific instruction is highlighted with a red box:

A plant species name consists of the genus, followed by the epithet, followed by the author name(s).

Below this, another instruction states:

Print only the plant species names, and no other additional text or explanation.

The right panel, labeled "RESPONSE", shows the output of the prompt, listing the extracted plant species names in a list format:

Chenopodium capitatum (L.) Aschers.  
Chenopodium ficifolium Sm.  
Chenopodium foliosum Aschers.  
Chenopodium opulifolium Schrad.  
Chenopodium schraderianum Schult.  
Chenopodium urbicum L.  
Corispermum leptopterum (Asch.) Iljin

At the bottom of the interface, there are buttons for "Pre-fill response", "Add message pair", "Improve prompt", and a "Beta" badge. The bottom right corner includes an "Add to Conversation" button and a view count of 101.

# Prompt engineering

Click to go forward, hold to see history

ANTHROPIC Dashboard Workbench Settings Docs Feedback MM

Plant species names 4 v1 Prompt Evaluate Run Ctrl + ⌘

SYSTEM PROMPT You are an expert botanist.

EXAMPLES

USER

Chenopodium capitatum (L.) Aschers»  
168 - \* ficifolium Sn.  
■169 \_ foliosum Asciesse.  
170 ~ opulifolium Schaad»  
171 - schraderianum Schuit»  
172 - urbicum L»  
175 Corispermum leptopterum (Asch.) Iljin

Extract the plant species names from the text above.

A plant species name consists of the genus, followed by the epithet, followed by the author name(s).

Return the result for each plant species in valid JSON object format, with the single key "species". Provide the complete results as a list of valid JSON objects.

Print only this list of JSON objects, and no other additional text or explanation.

RESPONSE

Markdown Text

```
[{"species": "Chenopodium capitatum (L.) Aschers."}, {"species": "Chenopodium ficifolium Sm."}, {"species": "Chenopodium foliosum Aschers."}, {"species": "Chenopodium opulifolium Schrad."}, {"species": "Chenopodium schraderianum Schult."}, {"species": "Chenopodium urbicum L."}, {"species": "Corispermum leptopterum (Asch.) Iljin"}]
```

Pre-fill response Add message pair Improve prompt Beta

195

Add to Conversation 147

# Prompt engineering

The screenshot displays the Anthropic Workbench interface. At the top, there are navigation links for Dashboard, Workbench, and Settings, along with Docs, Feedback, and a user profile icon. The main workspace is titled "Plant species names 5 v1" and has tabs for Prompt, Evaluate, and a Run button. The interface is divided into two main sections: SYSTEM PROMPT and RESPONSE.

**SYSTEM PROMPT:** "You are an expert botanist."

**EXAMPLES:** (Empty)

**USER:**

```
Chenopodium capitatum (L.) Aschers»  
168 - * ficifolium Sn.  
■169 _ foliosum Asciesse.  
170 ~ opulifolium Schaad»  
171 - schraderianum Schuit»  
172 - urbicum L»  
175 Corispermum leptopterum (Asch.) Iljin
```

Extract the plant species names from the text above.

A plant species name consists of the genus, followed by the epithet, followed by the author name(s).

Return the result for each plant species in valid JSON object format, with the single key "species". Provide the complete results as a list of valid JSON objects.

If there is no information about plant species on this page, print an empty list, like this: []

Print only this list of JSON objects, and no other additional text or explanation.

**RESPONSE:**

```
[  
{"species": "Chenopodium capitatum (L.) Aschers."},  
{"species": "Chenopodium ficifolium Sm."},  
{"species": "Chenopodium foliosum Aschers."},  
{"species": "Chenopodium opulifolium Schrad."},  
{"species": "Chenopodium schraderianum Schult."},  
{"species": "Chenopodium urbicum L."},  
{"species": "Corispermum leptopterum (Asch.) Iljin"}  
]
```

At the bottom of the interface, there are links for Pre-fill response, Add message pair, Improve prompt, and a Beta badge. The bottom right corner shows an "Add to Conversation" button and a share icon with a count of 147.



# Prompt engineering

## OCR error correction

ANTHROPIC Dashboard Workbench Settings Docs Feedback MM

Plant species names 5 v1 Prompt Evaluate Run Ctrl + ⌘

SYSTEM PROMPT You are an expert botanist.

Chenopodium capitatum (L.) Aschers»  
168 - \* ficifolium Sn.  
■169 - foliosum Aschiese.  
170 ~ opulifolium Schaad»  
171 - schraderianum Schuit»  
172 - urbicum L»  
175 Corispermum leptopterum (Asch.) Iljin

RESPONSE

```
[{"species": "Chenopodium capitatum (L.) Aschers."}, {"species": "Chenopodium ficifolium Sm."}, {"species": "Chenopodium foliosum Aschers."}, {"species": "Chenopodium opulifolium Schrad."}, {"species": "Chenopodium schraderianum Schult."}, {"species": "Chenopodium urbicum L."}, {"species": "Corispermum leptopterum (Asch.) Iljin"}]
```

167 Chenopodium capitatum (L.) Aschers.  
168 - ficifolium Sm.  
169 - foliosum Aschers.  
170 - opulifolium Schrad.  
171 - schraderianum Schult.  
172 - urbicum L.  
173 Corispermum leptopterum (Asch.) Iljin

"species". Provide the complete results as a list of valid JSON objects.

If there is no information about plant species on this page, print an empty list, like this: []

Print only this list of JSON objects, and no other additional text or explanation.

Pre-fill response Add message pair Improve prompt Beta

217 Add to Conversation 147

# Seedlists: OCR error correction

## Original typewritten text

167 Chenopodium capitatum (L.) Aschers.  
168 - ficifolium Sm.  
169 - foliosum Aschers.  
170 - opulifolium Schrad.  
171 - schraderianum Schult.  
172 - urbicum L.  
173 Corispermum leptopterum (Asch.) Iljin  
174 Kochia scoparia (L.) Schrad.  
  
C i s t a c e a e  
175 Cistus crispus L.  
176 - hirsutus Lam.  
177 - monspeliensis L.  
178 Helianthemum apenninum (L.) Mill.  
179 - var. roseum Grosser  
180 - nummularium (L.) Mill.  
181 Tuberaria guttata (L.) Fourr.  
  
C n e o r a c e a e  
182 Cneorum tricoccon L.  
  
C o m m e l i n a c e a e  
183 Aneilema papuanum Warburg  
184 Commelina communis L.  
185 - tuberosa L.  
186 Palisota mannii C.B. Clarke  
187 Tinantia erecta (Jacq.) Schlechtend.  
+ 188 Tradescantia crassifolia Cav.  
  
C o m p o s i t a e (Asteraceae)  
189 Achillea ageratifolia (Sibth. & Smith)  
Boiss.  
190 - macrophylla L.  
! 191 - millefolium L. (NG)  
192 Ammobium alatum R.Br.  
193 Antennaria howellii Greene  
194 - plantaginea R.Br.  
195 Anthemis arvensis L.  
196 - austriaca Jacq.  
197 - cotula L.  
! 198 - tinctoria L. (NG)

## Extracted OCR

Chenopodium capitatum (L.) Aschers»  
168 - \* ficifolium Sm.  
■ 169 - foliosum Asciesse.  
170 ~ opulifolium Schaad»  
171 - schraderianum Schult»  
172 - urbicum L»  
175 Corispermum leptopterum (Asch.) Iljin  
17^ Kochia scoparia (L») Schrad»  
Cista c e-a e  
175 Cistus crispus L»'  
176 - hirsutus Lam»  
177 - monspeliensis L».,,  
178 Heliantheum apenninum (L») Mill»  
179 - - var» roseum Grossor  
180 - nummularium (L») Mill»  
181 Tuberaria guttata (L») Fourr»  
C n e o r a c e a e  
182 Cneorum tricoccon L»  
C o m m e l i n a c e a \_e  
185 Aneilema papuanum Warburg  
18^ Commelina communis L»  
185 - tuberosa L»  
186 Palisota mannii C»B»ClarkeN  
187 Tinantia erecta (Jacq») Schlechtend»  
+ 188 Tradescantia crassifolia Cav».  
Compositae (Asteraceao)  
189 Achillea ageratifolia (Sibth» & Smith)  
Boiss »  
190 - macrophylla L»  
! 191 - millefolium L» (NG)  
192 Ammobium alatum R»Br.  
195 Antennaria howellii Greene  
19^ - plantaginea R»Br»  
195 Anthemis arvensis L»  
196 - austriaca Jacq»  
197 - cotula L»  
! 198 - tinctoria L» (NG)

## Claude 3 Opus output

0	Chenopodium capitatum (L.) Aschers.
1	Chenopodium ficifolium Sm.
2	Chenopodium foliosum Aschers.
3	Chenopodium opulifolium Schrad.
4	Chenopodium schraderianum Schult.
5	Chenopodium urbicum L.
6	Corispermum leptopterum (Asch.) Iljin
7	Kochia scoparia (L.) Schrad.
8	Cistus crispus L.
9	Cistus hirsutus Lam.
10	Cistus monspeliensis L.
11	Helianthemum apenninum (L.) Mill.
12	Helianthemum apenninum var. roseum Grosser
13	Helianthemum nummularium (L.) Mill.
14	Tuberaria guttata (L.) Fourr.
15	Cneorum tricoccon L.
16	Aneilema papuanum Warburg
17	Commelina communis L.
18	Commelina tuberosa L.
19	Palisota mannii C.B. Clarke
20	Tinantia erecta (Jacq.) Schlechtend.
21	Tradescantia crassifolia Cav.
22	Achillea ageratifolia (Sibth. & Smith) Boiss.
23	Achillea macrophylla L.
24	Achillea millefolium L.
25	Ammobium alatum R.Br.
26	Antennaria howellii Greene
27	Antennaria plantaginea R.Br.
28	Anthemis arvensis L.
29	Anthemis austriaca Jacq.
30	Anthemis cotula L.
31	Anthemis tinctoria L.