PQ-EID: Provenance-Quotient Evidential Interaction Decomposition and Certification in Web Question Answering

Khrystyna Zub, Ivanna Dronyuk and Oksana Mulesa

The material below documents the input records and evaluation protocol used for the 60-question QACC audit, together with the reported settings of the controlled provenance experiments. Question identifiers and evidence-group composition are listed in Table S1; the corresponding source passages are retained in the accompanying CSV file.

S1. QACC sample and evidence units

The evaluation used 60 questions selected from 304 eligible QACC records after a seeded shuffle (20261012). Each question was paired with two annotated answer candidates. Source-domain and answer-claim labels were used as observable grouping proxies. Five units, denoted e0–e4, were retained for each question in the recorded order. Both models received the same questions and evidence. Each five-unit game required 32 coalition scores, including the empty coalition.

Table S1 gives the question identifier and five evidence units. Within each unit, the entry identifies the domain, claim label and zero-based context index of the retained source passage. Several units contain multiple source fragments; the 300 canonical units are represented by 326 source-fragment records in the accompanying CSV. A shared domain is an operational source proxy and does not establish a verified derivation chain.

Table S1. Question identifiers and selected evidence-group metadata.

QACC ID	Question	Evidence groups (unit: domain / claim / context index)
791__122	Doctrine of lapse was one of the case for which rebellion?	e0: en.wikipedia.org / east india / 0 e1: britannica.com / unlabeled / 1 e2: en.wikipedia.org / unlabeled / 2 e3: byjus.com / revolt / 3 e4: pib.gov.in / unlabeled / 4
1271__2301	who plays izabella in transformers the last knight?	e0: transformers.fandom.com / unlabeled / 0 e1: imdb.com / isabela merced / 1; imdb.com / isabela merced / 3 e2: en.wikipedia.org / isabela merced / 7 e3: usatoday.com / isabela moner / 8 e4: hero.fandom.com / isabela moner / 9
1417__2409	the creation of human beings in the kumulipo happens during which wā or period of creation?	e0: quizlet.com / unlabeled / 0; quizlet.com / unlabeled / 2; quizlet.com / unlabeled / 4 e1: higp.hawaii.edu / unlabeled / 1 e2: en.wikipedia.org / unlabeled / 3 e3: hwst270.weebly.com / 16 wā / 5 e4: scholarsbank.uoregon.edu / second wā, / 8
1062__393	When was first teacher's day celebrated in india?	e0: en.wikipedia.org / 5 september / 0 e1: jagranjosh.com / 5 september / 1 e2: en.wikipedia.org / unlabeled / 2 e3: indianexpress.com / 1962 / 3 e4: ndtv.com / 5 september / 5
1200__1773	who was the editor of the journal jugantor published in the time of swadeshi movement?	e0: brainly.in / bhupendranath dutt / 0; brainly.in / bhupendranath dutt / 2 e1: en.wikipedia.org / barindra kumar ghosh / 1 e2: en.wikipedia.org / unlabeled / 3 e3: edurev.in / barindra kumar ghosh / 4 e4: instagram.com / unlabeled / 5
473__80	When was the electronic funds transfer act signed into law?	e0: consumerfinance.gov / unlabeled / 0 e1: investopedia.com / 1978 / 1 e2: en.wikipedia.org / 1978 / 2 e3: ncua.gov / , 2000 / 3 e4: congress.gov / unlabeled / 4
652__304	When does closure of the ductus arteriosus occur?	e0: ncbi.nlm.nih.gov / unlabeled / 0 e1: emedicine.medscape.com / unlabeled / 1 e2: mayoclinic.org / shortly after birth / 2 e3: nationwidechildrens.org / few days after birth, / 3 e4: heart.org / few days after birth, / 7
573__206	What does the c stand for chrysler 300c?	e0: vehiclehistory.com / unlabeled / 0 e1: 300cforums.com / unlabeled / 1 e2: darcarschryslerjeepdodgeofmarlowheights.com / unlabeled / 2 e3: lxforums.com / chrysler / 4 e4: autozine.org / coupe / 5

QACC ID	Question	Evidence groups (unit: domain / claim / context index)
1001__332	Where was tin pan alley in new york city?	e0: en.wikipedia.org / unlabeled / 0 e1: tinpanalley.nyc / 47-55 west 28th street / 1 e2: nyc.gov / west 28th street between sixth / 2 e3: britannica.com / unlabeled / 3 e4: soundamerican.org / west 28th street between sixth / 4
585__220	A drug that might be used in surgery for its amnesic properties is?	e0: ncbi.nlm.nih.gov / diazepam / 0; ncbi.nlm.nih.gov / diazepam / 2 e1: emedicine.medscape.com / unlabeled / 1 e2: adf.org.au / ketamine / 3 e3: ncbi.nlm.nih.gov / ketamine / 4 e4: pubmed.ncbi.nlm.nih.gov / midazolam / 5
1209__2571	what is the system of government in japan?	e0: japan.kantei.go.jp / unlabeled / 0 e1: europarl.europa.eu / monarchy and parliamentary / 1 e2: chathamhouse.org / parliamentary system / 3 e3: japanhpn.org / parliamentary system / 5 e4: commonwealthfund.org / parliamentary system / 8
1029__360	How many episodes are there in a place to call home?	e0: acorn.tv / unlabeled / 0 e1: en.wikipedia.org / thirteen / 1 e2: tvguide.com / unlabeled / 2 e3: kpbs.org / unlabeled / 3 e4: amazon.com / 67 / 4
847__178	Highest run score by player in ipl 2018?	e0: espncricinfo.com / unlabeled / 0 e1: iplt20.com / 735runs / 1 e2: bleacherreport.com / unlabeled / 2 e3: sports.ndtv.com / 684 / 4 e4: timesofindia.indiatimes.com / 684 / 7
116__116	Who presents the people's postcode lottery advert?	e0: postcodelottery.co.uk / jason donovan / 0 e1: postcodelottery.info / unlabeled / 1 e2: youtube.com / unlabeled / 2 e3: cats.org.uk / unlabeled / 3 e4: novamedia.com / judie mccourt, danyl johnson, / 5
874__205	When did the first jungle book movie come out?	e0: en.wikipedia.org / 1967 / 0 e1: en.wikipedia.org / december 25, 1994 / 1 e2: boxofficemojo.com / unlabeled / 2 e3: jonnegroni.com / unlabeled / 3 e4: wdw-magazine.com / 1967 / 4
1065__1715	where is urinary bladder located in human body?	e0: my.clevelandclinic.org / lower part of your abdomen / 0 e1: hopkinsmedicine.org / below the ribs / 1 e2: ncbi.nlm.nih.gov / below the ribs / 3 e3: niddk.nih.gov / lower part of your abdomen / 5 e4: nia.nih.gov / lower part of your abdomen / 6
1459__2715	who was the first elected president of pakistan?	e0: na.gov.pk / unlabeled / 0 e1: pacific.edu / iskander mirza / 1 e2: un.org / unlabeled / 2 e3: en.wikipedia.org / unlabeled / 3 e4: contactpoint.pacific.edu / zardari / 7
810__141	When was the last time newcastle won the fa cup?	e0: en.wikipedia.org / unlabeled / 0; en.wikipedia.org / unlabeled / 2 e1: sportingnews.com / unlabeled / 1 e2: aljazeera.com / unlabeled / 3 e3: theathletic.com / 1955 / 4 e4: goal.com / 1954-55 / 9
1445__3594	where did the world's largest recorded wave occur?	e0: rapturecamps.com / lituya bay, alaska / 0 e1: surfertoday.com / lituya bay, alaska / 1 e2: pbs.org / lituya bay, alaska / 2 e3: usgs.gov / chile / 4 e4: livescience.com / lituya bay, alaska / 5
989__320	When was night of the living dead released?	e0: en.wikipedia.org / october 1, 1968 / 0 e1: britannica.com / october 1, 1968 / 1 e2: imdb.com / october 4, 1968 / 2 e3: rogerebert.com / unlabeled / 3 e4: imdb.com / unlabeled / 4
112__112	When does the fourth season of flash come out?	e0: en.wikipedia.org / unlabeled / 0 e1: arrow.fandom.com / october 10, 2017 / 1 e2: deadline.com / unlabeled / 2 e3: netflix.com / unlabeled / 3 e4: bestbuy.com / 08/28/2018 / 5
939__270	Who sold more albums elvis or michael jackson?	e0: decluttr.com / elvis presley / 0 e1: wpgtalkradio.com / elvis presley / 2 e2: quora.com / not clear / 3 e3: lipstickalley.com / elvis presley / 5 e4: tomsonhighway.com / elvis presley / 7

QACC ID	Question	Evidence groups (unit: domain / claim / context index)
471__77	Who sang the song good morning good morning?	e0: youtube.com / unlabeled / 0; youtube.com / unlabeled / 2; youtube.com / unlabeled / 4; youtube.com / unlabeled / 6 e1: en.wikipedia.org / the beatles / 1 e2: en.wikipedia.org / nacio herb brown / 3 e3: ratedrnb.com / pj morton / 5 e4: goodmorningamerica.com / unlabeled / 7
1409__3069	who sings i wanna rock'n'roll all night?	e0: youtube.com / unlabeled / 0; youtube.com / unlabeled / 4; youtube.com / unlabeled / 6 e1: en.wikipedia.org / american / 1 e2: youtube.com / dominique / 2 e3: songfacts.com / unlabeled / 3 e4: quora.com / unlabeled / 5
1204__1767	consist of the sum of the fixed and variable costs for any given level of production?	e0: home.ubalt.edu / total cost / 3 e1: hafizmrf3.medium.com / total cost / 4 e2: easynotecards.com / total cost / 5 e3: extension.colostate.edu / total cost / 6 e4: investopedia.com / production cost / 8
1134__1887	what is the meaning of auv in cars?	e0: c-magazine.com / asian utility vehicles / 0 e1: autodeal.com.ph / autonomous underwater vehicles / 1 e2: vigattininsurance.com / asian utility vehicles / 3 e3: tsikot.com / asian utility vehicles / 4 e4: en.wikipedia.org / autonomous underwater vehicles / 5; en.wikipedia.org / autonomous underwater vehicles / 8
962__293	Who get the flag at a military funeral?	e0: va.gov / a funeral director / 0; va.gov / a funeral director / 2 e1: cem.va.gov / next-of-kin / 1 e2: carrot-top.com / next-of-kin / 3 e3: defense.gov / next-of-kin / 6 e4: dod.hawaii.gov / next-of-kin / 7
876__207	Where does the last name aranda come from?	e0: forebears.io / santander / 0 e1: houseofnames.com / asturias / 1 e2: ancestry.com / aranda / 2 e3: en.wikipedia.org / unlabeled / 3 e4: namecensus.com / unlabeled / 4
890__221	When did the seed of chucky come out?	e0: en.wikipedia.org / unlabeled / 0 e1: syfy.com / unlabeled / 1 e2: childsplay.fandom.com / november 12 / 2 e3: mtv.com / nov 4, 2004 / 5 e4: nymag.com / nov 4, 2004 / 7
470__76	Who was the viceroy when the simon commission visited india?	e0: toppr.com / lloyd george / 0 e1: en.wikipedia.org / lord irwin / 1 e2: testbook.com / lord irwin / 2 e3: gktoday.in / lord irwin / 3 e4: open.ac.uk / lord irwin / 4
640__290	New mexico was admitted to the union as the . . . state?	e0: guides.loc.gov / 47th / 0 e1: archives.gov / 47th / 1 e2: history.com / 47th / 4 e3: senate.gov / 47th / 6 e4: parks.ca.gov / free, nonslavery state / 8
1377__2394	who played skeletor in the movie masters of the universe?	e0: imdb.com / frank langella / 0; imdb.com / frank langella / 2 e1: en.wikipedia.org / frank langella / 1 e2: reddit.com / jack palance / 3 e3: reddit.com / frank langella / 5 e4: yahoo.com / frank langella / 7
836__167	Who prints currency notes and coins in india?	e0: investopedia.com / the reserve bank of india / 0 e1: en.wikipedia.org / currency note press / 1 e2: home.treasury.gov / unlabeled / 2 e3: drishtiias.com / the reserve bank of india / 3 e4: spmcil.com / currency note press / 6
491__104	The world's tallest jain statue is located in which state of india?	e0: en.wikipedia.org / mangi tungi / 0 e1: gktoday.in / unlabeled / 1 e2: en.wikipedia.org / unlabeled / 2 e3: testbook.com / kevadia / 3 e4: walkthroughindia.com / mangi tungi / 4
895__226	Who wrote the theme song for phineas and ferb?	e0: phineasandferb.fandom.com / unlabeled / 0 e1: disneychannel.fandom.com / jeff "swampy" marsh / 1 e2: reddit.com / bowling for soup / 3 e3: deadline.com / jeff "swampy" marsh / 7 e4: genius.com / jeff "swampy" marsh / 8
979__310	How many total pin codes are there in india?	e0: quora.com / 19100 / 0 e1: bankbazaar.com / unlabeled / 1

QACC ID	Question	Evidence groups (unit: domain / claim / context index)
		e2: dnaindia.com / 19100 / 4 e3: aboutamazon.com / more than 20000 / 5 e4: data.gov.in / 19100 / 7
1554__2224	how many seasons of from dusk till dawn are there?	e0: en.wikipedia.org / 3 / 0; en.wikipedia.org / 3 / 2 e1: imdb.com / unlabeled / 1; imdb.com / unlabeled / 4 e2: screenrant.com / 4 / 3 e3: amazon.com / unlabeled / 5 e4: musingsofamiddleagedgeek.blog / unlabeled / 6
1405__3440	who wrote if she knew what she wants?	e0: en.wikipedia.org / jules shear / 0 e1: namebrandketchup.com / the bangles / 2 e2: simple.wikipedia.org / the bangles / 5 e3: imdb.com / the bangles / 6 e4: secondhandsongs.com / jules shear / 7
453__56	Who sang heard it thru the grapevine first?	e0: en.wikipedia.org / unlabeled / 0 e1: udiscovermusic.com / marvin gaye / 1; udiscovermusic.com / marvin gaye / 3 e2: sites.google.com / marvin gaye / 2 e3: 99wfmk.com / miracles / 7 e4: variety.com / marvin gaye / 9
337__337	Who has scored most runs in t20 international matches?	e0: espncricinfo.com / unlabeled / 0 e1: en.wikipedia.org / virat kohli / 1 e2: amsivri.icar.gov.in / unlabeled / 2 e3: en.wikipedia.org / gayle / 5 e4: mykhel.com / virat kohli / 6
108__108	What laws directed the flow of goods between england and the colonies?	e0: quizlet.com / unlabeled / 0 e1: khanacademy.org / the navigation acts / 1 e2: quizlet.com / the navigation acts / 2 e3: loc.gov / unlabeled / 3; loc.gov / unlabeled / 5 e4: digitalhistory.uh.edu / special laws / 4
1121__1606	which is produced in plants of narora kakrapar tarapur?	e0: world-nuclear.org / atomic energy / 0 e1: brainly.in / unlabeled / 1 e2: researchgate.net / unlabeled / 2 e3: en.wikipedia.org / components for phwrs / 3 e4: world-nuclear-news.org / unlabeled / 4
967__298	Where does most of earths oxygen come from today?	e0: oceanservice.noaa.gov / the ocean / 0 e1: thomasnet.com / phytoplankton / 1 e2: nationalgeographic.org / unlabeled / 2 e3: theconversation.com / the ocean / 3 e4: whoi.edu / the ocean / 5
574__207	Who was the first executive president of guyana?	e0: en.wikipedia.org / unlabeled / 0; en.wikipedia.org / unlabeled / 1 e1: pdba.georgetown.edu / unlabeled / 2 e2: facebook.com / linden forbes sampson burnham / 3 e3: 2009-2017.state.gov / cheddi jagan / 4 e4: bbc.com / linden forbes sampson burnham / 7
227__227	Where does the coin go in a french drop?	e0: youtube.com / made to disappear / 0; youtube.com / made to disappear / 4; youtube.com / made to disappear / 6 e1: masterclass.com / made to disappear / 1; masterclass.com / made to disappear / 8 e2: youtube.com / dropped / 2 e3: quora.com / made to disappear / 3 e4: howcast.com / dropped / 7
968__299	What genre is the book the adventures of huckleberry finn?	e0: study.com / picaresque,bildungsroman,satir / 0 e1: sparknotes.com / picaresque / 1 e2: study.com / adventure / 2 e3: amazon.com / unlabeled / 3 e4: gale.com / unlabeled / 4
598__235	Mention the chemical change that proinsulin undergo to be able to act as mature insulin?	e0: byjus.com / the 'c' peptide is cleaved / 0 e1: toppr.com / unlabeled / 1 e2: shaalaa.com / disulphide bridges / 2 e3: vedantu.com / disulphide bridges / 4 e4: ncbi.nlm.nih.gov / disulphide bridges / 7; ncbi.nlm.nih.gov / disulphide bridges / 9
139__139	Who plays aang in avatar the last airbender movie?	e0: en.wikipedia.org / noah andrew ringer / 0 e1: imdb.com / unlabeled / 1 e2: en.wikipedia.org / noah ringer / 2 e3: whats-on-netflix.com / unlabeled / 3 e4: cbc.ca / noah ringer / 7
505__121	Where does the last name hansen come from?	e0: en.wikipedia.org / denmark / 0 e1: ancestry.com / unlabeled / 1 e2: forebears.io / scandinavian / 2

QACC ID	Question	Evidence groups (unit: domain / claim / context index)
		e3: selectsurnames.com / denmark / 4 e4: quora.com / scandinavian / 6
1025__356	What actor holds the record for sexiest man alive?	e0: people.com / unlabeled / 0 e1: insider.com / chris evans / 1 e2: ca.sports.yahoo.com / chris evans / 2 e3: stltoday.com / chris evans / 4 e4: yardbarker.com / johnny depp / 8
287__287	Who has won the us open golf 2018 title in new york?	e0: en.wikipedia.org / unlabeled / 0 e1: en.wikipedia.org / brooks koepka / 1 e2: usopen.com / unlabeled / 2 e3: nytimes.com / johnson / 3 e4: nytimes.com / brooks koepka / 5
1301__3189	who sings gone gone gone she been gone so long?	e0: youtube.com / chilliwack / 0; youtube.com / chilliwack / 2; youtube.com / chilliwack / 4 e1: en.wikipedia.org / chilliwack / 1 e2: talkbass.com / the recess monkeys / 3 e3: genius.com / unlabeled / 5 e4: reddit.com / unlabeled / 6; reddit.com / unlabeled / 9
271__271	When was alice in chains jar of flies released?	e0: en.wikipedia.org / january 25, 1994 / 0 e1: loudwire.com / january 25, 1994 / 1 e2: revolvermag.com / 1994 / 2 e3: discogs.com / 1994 / 4 e4: alice-in-chains.fandom.com / january 25, 1994 / 5
676__7	What area of san diego is sdsu in?	e0: admissions.sdsu.edu / the heart / 0 e1: en.wikipedia.org / the college area / 1 e2: sdsu.edu / unlabeled / 2 e3: en.wikipedia.org / south bay / 3 e4: calstate.edu / unlabeled / 4
1423__3602	who were the two mathematicians that invented calculus?	e0: en.wikipedia.org / isaac newton, gottfried wilhel / 0 e1: amsi.org.au / isaac newton, gottfried wilhel / 2 e2: mathtutordvd.com / isaac newton, gottfried wilhel / 5 e3: tutorocean.com / jacob bernoulli / 7 e4: study.com / isaac newton, gottfried wilhel / 9
1347__3357	who has won the cma entertainer of the year the most?	e0: en.wikipedia.org / garth brooks / 0 e1: yardbarker.com / garth brooks / 3 e2: theboot.com / garth brooks / 4 e3: wideopencountry.com / garth brooks / 5 e4: theboot.com / brad paisley / 7
1295__2996	where is the largest salt mine in the us?	e0: en.wikipedia.org / livingston county, new york / 0 e1: compassminerals.com / unlabeled / 1 e2: americanmineservices.com / unlabeled / 2 e3: allensamuelshutch.com / lake huron / 3 e4: draglamsalt.com / unlabeled / 4
1394__2895	who played the princess in kingsman secret service?	e0: imdb.com / hanna carolina alström / 0 e1: en.wikipedia.org / hanna carolina alström / 1 e2: kingsman.fandom.com / unlabeled / 2 e3: m.imdb.com / hanna carolina alström / 6 e4: 0b.medium.com / colin firth / 8
1357__2669	who wrote the theme to last of the mohicans?	e0: en.wikipedia.org / trevor jones / 0 e1: amazon.com / trevor jones / 1 e2: cnn.com / trevor jones / 2 e3: maintitles.net / randy edelman / 3 e4: moviemusicuk.us / trevor jones / 4
988__319	How old is was the oldest person to ever live?	e0: cnbc.com / 122 years / 0 e1: en.wikipedia.org / 122 years / 1 e2: sciencefocus.com / 122 years / 2 e3: a-z-animals.com / 116 years / 3 e4: statista.com / 122 years / 6

S2. Prompt composition

Evidence units were inserted in fixed order for each coalition. Each retained evidence text was truncated to its first 1,100 characters and displayed with a consecutive source number. For the empty coalition, the evidence field contained the literal string “(no retrieved evidence)”. The system instruction was “You are a careful question-answering evaluator.” The user message had the following form:

Question: {question}
Evidence:
{evidence}
Choose the better-supported answer using the provided evidence. Output exactly one letter, A or B, with no explanation.
A: {candidate_A}
B: {candidate_B}
Answer:

The system and user messages were passed to the model-specific tokenizer chat template with add_generation_prompt=True. Each coalition was evaluated twice, with the two answer candidates exchanged between positions A and B while the question and evidence remained fixed.

S3. Fixed-choice model scores

The models were Qwen/Qwen2.5-0.5B-Instruct and HuggingFaceTB/SmolLM2-1.7B-Instruct. They were evaluated as causal language models on CPU in float32 inference mode. No free-form answer generation was used in calculating coalition values. Instead, the scorer extracted next-token logits for the label continuations " A" and " B" and checked that each label mapped to one token.

For each answer-position arrangement, the scorer calculated the logit difference $z(A) - z(B)$. The two positions were combined as a correct-answer-minus-alternative margin by reversing the sign of the swapped-position comparison and averaging the resulting margins. The empty-coalition margin was subtracted from every coalition margin. Möbius dividends were then calculated by exact subset inversion of the 32-value game.

The primary audit recorded 70 model-forward evaluations per question and model, giving 8,400 evaluations in total for the two models. A separate order-sensitivity follow-up tested three non-original permutations of the full evidence set, with two candidate positions for each permutation (720 additional forward evaluations overall). The resulting measures are fixed-choice candidate preferences, not open-generation exact-match accuracy.

S4. Canonicalization-error experiment

Five canonical units were represented by three raw passages each (15 passages overall). The canonicalization operators were MERGE, SPLIT, ROOT, SEM and ALL; nominal error rates were 0, 0.05, 0.10, 0.20 and 0.30. Each operator–rate combination used 1,000 Monte Carlo repetitions. The NumPy generator seed was 20261002 (the base seed 20260925 plus 77).

The nonzero singleton coefficients were 0.8, 0.6, −0.4, 0.2 and 0.1. Pairwise coefficients were 0.9 for (0,1), −0.5 for (1,2) and 0.4 for (3,4). Third-order coefficients were 1.2 for (0,1,3) and −0.7 for (1,2,4). MERGE combined distinct units, SPLIT separated copies, and ROOT reassigned provenance roots. SEM combined splitting and merging at half the nominal rate for each operation; ALL applied root reassignment, splitting and merging at one third of the nominal rate each. The fixed-partition auxiliary sign audit used a probability floor of 0.10 and a perturbation bound of 0.005.

S5. Dirichlet structural-uncertainty experiment

Two third-order target interactions had prescribed ground-truth coefficients of +1.40 and −1.10. The candidate partitions included the true provenance partition, incorrect-merge configurations and a split configuration. Partition weights were sampled from Dirichlet distributions corresponding to low, medium and high structural uncertainty. These regimes varied the concentration of weight across candidate partitions. Each uncertainty regime and target interaction was evaluated over 5,000 independent Monte Carlo runs. The reported measures were MAP-partition coverage and sign accuracy, dual-certificate coverage, accuracy among certified signs, false-sign frequency and abstention rate.

These sampled weights represent controlled structural ambiguity rather than calibrated probabilities for real-world source origins.

S6. Data source and accompanying records

The QACC benchmark and its documentation are available through Amazon Science: https://www.amazon.science/code-and-datasets/open-domain-question-answering-with-conflicting-contexts. The accompanying file PQ_EID_QACC_60_Evidence_Groups.csv provides question IDs, selected source fragments, group IDs, observed domain and claim labels, original context indices, source URLs and the two candidate answers.