

# मोडल रिग्रेशन के साथ क्रॉस-प्लेटफॉर्म शतरंज रेटिंग मैपिंग का अनुमान लगाना

Elo+Chess अनुसंधान नोट्स

29 मई 2026

## Abstract

खिलाड़ी अक्सर पूछते हैं कि Lichess.org पर रेटिंग की तुलना Chess.com पर रेटिंग से कैसे की जाती है। यह प्रश्न Elo+Chess के लिए व्यावहारिक रूप से महत्वपूर्ण है क्योंकि यह इसका बेंचमार्क है वक्र पूर्ण से लिए गए एक बहुत बड़े स्तरीकृत नमूने से बनाए जाते हैं Lichess.org मासिक गेम इतिहास, जबकि उपयोगकर्ता इनमें से किसी एक से इतिहास ला सकते हैं साइट. यह नोट अनुमान लगाने के लिए उपयोग की जाने वाली वर्तमान पद्धति का वर्णन करता है Lichess रेटिंग से Chess.com रेटिंग तक गेम-प्रकार-विशिष्ट मैपिंग। द विधि समान-उपयोगकर्ता नाम उम्मीदवार मिलान, सटीक या लगभग-सटीक समय-नियंत्रण का उपयोग करती है फ़िल्टर, न्यूनतम गेम-इतिहास आवश्यकताएँ, और एक मोडल रिग्रेशन प्रक्रिया झूठे समान-उपयोगकर्ता नाम मिलान से पूर्वाग्रह को कम करने के लिए डिज़ाइन किया गया। वर्तमान उत्पादन मानचित्रण है:

$$\hat{R}_{\text{Chess.com}} = \alpha + \beta R_{\text{Lichess}}$$

अनुमान प्रारंभिक हैं और मिलान किए गए नमूने के बढ़ने पर इसे अपडेट किया जाएगा। विशेष रूप से 10 मिनट से अधिक लंबे तीव्र गेम के लिए।

## वर्तमान उत्पादन मानचित्रण

उन पाठकों के लिए जो पहले परिणाम चाहते हैं, वर्तमान उत्पादन समीकरण हैं:

$$\begin{aligned}\hat{R}_{\text{Chess.com,bullet}} &= -530.60 + 0.9859R_{\text{Lichess}}, \\ \hat{R}_{\text{Chess.com,blitz}} &= -548.88 + 1.0837R_{\text{Lichess}}, \\ \hat{R}_{\text{Chess.com,10minrapid}} &= -499.68 + 1.0778R_{\text{Lichess}}, \\ \hat{R}_{\text{Chess.com,>10rapid}} &= -351.72 + 0.9003R_{\text{Lichess}}.\end{aligned}$$

ये लाइनें Lichess रेटिंग 650-1600 से ऊपर फिट की गई हैं और नहीं होनी चाहिए सभी कौशल स्तरों पर सार्वभौमिक रूपांतरण के रूप में एक्सट्रपलेशन किया गया।

## 1 उद्देश्य

लक्ष्य Lichess.org पर एक रेटिंग को तुलनीय रेटिंग पैमाने पर मैप करना है समान व्यापक समय-नियंत्रण श्रेणी के लिए Chess.com। हम एक अलग अनुमान लगाते हैं Elo+Chess द्वारा उपयोग की जाने वाली प्रत्येक श्रेणी के लिए मैपिंग:

- 1+0 गोली,
- 3+0 ब्लिट्ज़,
- 10+0 तीव्र,

- 10 मिनट से अधिक लंबे तीव्र खेल।

फिट की गई लाइन का उपयोग केवल स्केल रूपांतरण परत के रूप में किया जाता है। इसका यह मतलब नहीं है दो रेटिंग पूलों में समान रेटिंग मैकेनिक्स, प्लेयर पूल या हैं खेल-चयन व्यवहार.

## 2 डेटा उद्गम

विश्लेषण का Lichess पक्ष संपूर्ण सार्वजनिक Lichess से बनाया गया है जनवरी 2025 से मार्च 2026 तक मानक रेटेड गेम संग्रह। यह मायने रखता है क्योंकि Lichess नमूना कोई सुविधा स्कैप या संकीर्ण नमूना नहीं है दृश्यमान प्रोफाइल: इसका निर्माण पूर्ण मासिक गेम-इतिहास फ़ाइलों से किया गया है अनुसंधान और सार्वजनिक उपयोग के लिए Lichess द्वारा जारी किया गया।

Lichess इन डेटाबेस निर्यातों को <https://database.lichess.org/> पर प्रकाशित करता है। डेटाबेस पेज बताता है कि निर्यात “CC0 लाइसेंस” के तहत जारी किए जाते हैं और स्पष्ट रूप से उपयोगकर्ताओं को “अनुसंधान के लिए उनका उपयोग करने” की अनुमति देता है “प्रकाशन”<sup>1</sup> यह ओपन-डेटा आसन Elo+Chess के बेंचमार्क पक्ष के डिजाइन के केंद्र में है।

Chess.com पक्ष पर, उम्मीदवार के समान-उपयोगकर्ता नाम वाले खातों की ही जाँच की जाती है मैपिंग का अनुमान लगाने के लिए आवश्यक मासिक अभिलेखागार और खेल श्रेणियाँ। हम एक ही कैलेंडर के भीतर खेल इतिहास के मिलान पर विशेष जोर दें जब भी संभव हो महीना. कैलेंडर-माह मिलान से यह जोखिम कम हो जाता है कि a Lichess और Chess.com के बीच खिलाड़ी की रेटिंग में काफी बदलाव आया माप.

## 3 उम्मीदवार मिलान निर्माण

मुख्य चुनौती यह है कि कोई भी प्लेटफ़ॉर्म सार्वभौमिक क्रॉस-प्लेटफ़ॉर्म को उजागर नहीं करता है खिलाड़ी पहचानकर्ता. इसलिए हम समान-उपयोगकर्ता नाम वाले उम्मीदवारों से शुरुआत करते हैं: यदि ए उपयोगकर्ता नाम Lichess.org और Chess.com दोनों पर दिखाई देता है, उस जोड़ी को एक माना जाता है उम्मीदवार मिलान. यह जानबूझकर शोर है. कुछ समान नाम वाले खाते हैं वही व्यक्ति, जबकि अन्य असंबंधित खिलाड़ी हैं जो उसी का उपयोग करते हैं संभाल.

टाले जा सकने वाले शोर को कम करने के लिए, वर्तमान पाइपलाइन गेम-प्रकार-विशिष्ट को लागू करती है फ़िल्टर:

1. Lichess उपयोगकर्ताओं को स्तरीकृत गेम-इतिहास डेटासेट से नमूना लिया जाता है रेटिंग बकेट और समय-नियंत्रण प्रकार।
2. Chess.com उम्मीदवार के खातों की जांच की जाती है समय-नियंत्रण श्रेणी.
3. उम्मीदवार जोड़े तभी बनाए रखे जाते हैं जब दोनों पक्षों के पास पर्याप्त प्रासंगिकता हो खेल का इतिहास.
4. बुलेट, ब्लिट्ज और 10 मिनट की रैपिड के लिए, न्यूनतम संरक्षित इतिहास है प्रति पक्ष 5 खेल।
5. 10 मिनट से अधिक की तीव्र गति के लिए, न्यूनतम बनाए रखा गया इतिहास 10 है प्रति पक्ष खेल.
6. यहां रिपोर्ट किया गया प्रतिगमन फिट Lichess रेटिंग तक सीमित है 650 से 1600, मुख्य Elo+Chess के लिए सबसे अधिक प्रासंगिक रेटिंग क्षेत्र से मेल खाता है बेंचमार्क रिपोर्ट.

वर्तमान नमूना समान-कैलेंडर-माह समान-उपयोगकर्ता नाम मिलान पर जोर देता है। जब ए उम्मीदवार के खाते में लक्ष्य समय-नियंत्रण श्रेणी में पर्याप्त गेम हैं Lichess स्रोत डेटा के समान महीने, जो मासिक रेटिंग से मेल खाता है फिटिंग के लिए पसंदीदा. पुराने क्रॉस-प्लेटफ़ॉर्म मिलान डेटा को वहीं रखा जाता है जहां वे हैं नमूना समर्थन बढ़ाएँ, लेकिन पद्धतिगत लक्ष्य खेल-प्रकार-विशिष्ट है एक ही महीने की तुलना।

<sup>1</sup><https://database.lichess.org/>, accessed May 29, 2026. The quoted phrases are from the ZXQPROT3ZXQ database page.

## 4 माध्य या माध्यिका पक्षपाती क्यों हो सकती है?

एक निश्चित Lichess रेटिंग बकेट के लिए, समान-उपयोगकर्ता नाम उम्मीदवार सेट एक मिश्रण है:

$$f(y | x) = \pi f_{\text{true}}(y | x) + (1 - \pi) f_{\text{false}}(y),$$

जहां  $x$  Lichess रेटिंग है,  $y$  Chess.com रेटिंग है,  $f_{\text{true}}$  is the distribution of Chess.com ratings for true क्रॉस-प्लेटफॉर्म मिलान, और  $f_{\text{false}}$  उत्पन्न वितरण है समान उपयोगकर्ता नाम वाले असंबद्ध Chess.com खातों द्वारा।

गलत-मिलान घटक सही Chess.com समकक्ष पर केंद्रित नहीं है Lichess बकेट के लिए. यह व्यापक Chess.com जनसंख्या से मिलता जुलता है उपयोगकर्ता नाम और गतिविधि फ़िल्टर द्वारा नमूना। परिणामस्वरूप, सशर्त माध्य है वैश्विक सक्रिय-खिलाड़ी वितरण की ओर खींचा गया:

$$E[y | x] = \pi E_{\text{true}}[y | x] + (1 - \pi) E_{\text{false}}[y].$$

यदि  $\pi < 1$ , तो यह आम तौर पर बांछित मिलान-खिलाड़ी के बराबर नहीं है उम्मीद. जब भी गलत मिलान होता है तो यही समस्या माध्यिका को प्रभावित करती है घटक 50वें प्रतिशतक को स्थानांतरित करने के लिए पर्याप्त द्रव्यमान का योगदान देता है।

पूर्वाग्रह की दिशा बाल्टी पर निर्भर करती है। निम्न Lichess बाल्टियों में, असत्य यदि यादृच्छिक सक्रिय Chess.com खाते हैं तो मिलान माध्य को ऊपर की ओर खींच सकता है वास्तविक मिलान वाले उपयोगकर्ताओं से अधिक मजबूत। ऊंची बाल्टियों में, वही संदूषण माध्य को नीचे की ओर खींच सकता है। यह संपीड़न प्रभाव बिल्कुल वही है जो हम चाहते हैं स्केल रूपांतरण बनाते समय इससे बचें।

## 5 मॉडल प्रेरणा

मोडल दृष्टिकोण माध्य-आधारित की तुलना में कमजोर धारणा पर आधारित है दृष्टिकोण. हमें यह आवश्यक नहीं है कि प्रत्येक समान-उपयोगकर्ता नाम युग्म सही हो। इसके बजाय, हम मान लें कि पर्याप्त रूप से संकीर्ण Lichess रेटिंग बकेट के भीतर, वास्तविक मिलान होता है Chess.com रेटिंग क्षेत्र में सबसे अधिक केंद्रित स्थानीय क्लस्टर बनाएं। मिथ्या मिलान असंख्य हो सकते हैं, लेकिन उन्हें अधिक व्यापक रूप से वितरित किया जाना चाहिए क्योंकि वे Lichess रेटिंग बकेट से बंधे नहीं हैं।

प्रत्येक Lichess बकेट के लिए, हम Chess.com रेटिंग के मोड का अनुमान लगाते हैं कर्नेल घनत्व अनुमान का उपयोग करके वितरण। फिर मोडल बिंदु का उपचार किया जाता है उस Lichess बकेट के लिए सबसे प्रशंसनीय Chess.com समकक्ष के रूप में। यह बनाता है यह प्रक्रिया लंबी पूँछों के प्रति कम संवेदनशील और बेमेल के प्रति कम संवेदनशील होती है व्यापक Chess.com जनसंख्या से समान नाम वाले खाते का नमूना लिया गया।

यह कोई दावा नहीं है कि मोड उत्तम है। यह एक व्यावहारिक मजबूती है शोरगुल वाली इकाई-रिज़ॉल्यूशन समस्या के लिए विकल्प। सही होने पर विधि सर्वोत्तम कार्य करती है सही क्रॉस-प्लेटफॉर्म रेटिंग और गलत के निकट मिलान बहुलता में रहते हैं मैच स्वयं भ्रामक मूल्य पर केंद्रित नहीं हैं।

## 6 फिटिंग प्रक्रिया

वर्तमान उत्पादन फिट प्रत्येक के लिए अलग से निम्नलिखित प्रक्रिया का उपयोग करता है खेल का प्रकार:

1. गेम-गणना को संतुष्ट करते हुए उम्मीदवार के समान-उपयोगकर्ता नाम जोड़े को बनाए रखें दहलीज
2. फिट को  $650 \leq R_{\text{ZXQPROT3ZXQ}} \leq 1600$  तक सीमित रखें।
3. Lichess रेटिंग को 100-पाइंट बकेट में विभाजित करें।
4. कम से कम 30 उपयोगकर्ताओं वाली प्रत्येक बकेट के लिए, Chess.com मोडल का अनुमान लगाएं बैंडविड्थ 75 रेटिंग का उपयोग करके गॉसियन-कर्नेल घनत्व अनुमान के साथ रेटिंग अंक.

5. प्रत्येक बकेट को उसके मध्यबिंदु और अनुमानित Chess.com मोड द्वारा निरूपित करें।
6. बकेट-लेवल मोडल के माध्यम से एक सामान्य न्यूनतम-वर्ग रेखा फिट करें अंक:

$$\hat{R}_{\text{Chess.com}} = \alpha + \beta R_{\text{Lichess}}.$$

केवल बकेट-स्तरीय मोडल बिंदु ही अंतिम रैखिक प्रतिगमन में प्रवेश करते हैं। व्यक्तिगत समान-उपयोगकर्ता नाम जोड़े अपने योगदान के माध्यम से वक्र को प्रभावित करते हैं बाल्टी के भीतर घनत्व का अनुमान।

## 7 वर्तमान अनुमान

तालिका 1 वर्तमान फिट स्थिरांक और ढलान की रिपोर्ट करती है। ये वे मान हैं जो वर्तमान में 29 मई, 2026 के बाद Elo+Chess में वायर्ड हैं अद्यतन.

Game type	Total users	Baseline	Same-month	In range	$\alpha$	$\beta$
1+0 bullet	2,999	1,474	1,525	1,843	-530.60	0.9859
3+0 blitz	3,101	1,306	1,795	2,075	-548.88	1.0837
10+0 rapid	3,084	1,509	1,575	2,523	-499.68	1.0778
> 10 rapid	750	445	305	575	-351.72	0.9003

Table 1: वर्तमान मोडल लीनियर Lichess रेटिंग 650–1600 से अधिक फिट बैठता है। “में रेंज” फिट किए गए Lichess अंतराल के अंदर बनाए गए उपयोगकर्ताओं की गणना करता है। आधारभूत पंक्तियाँ मौजूदा सटीक-नियंत्रण क्रॉस-प्लेटफॉर्म तालिका से आएं; समान-माह पंक्तियाँ आती हैं मार्च 2026 के मासिक समान-उपयोगकर्ता नाम नमूने से।

10 मिनट से अधिक लंबे त्वरित अनुमान को वास्तव में इसकी तुलना में कम समर्थन प्राप्त है अन्य तीन श्रेणियां। इसलिए इसके अतिरिक्त के रूप में बदलने की सबसे अधिक संभावना है मिलान किए गए उपयोगकर्ता एकत्र किए जाते हैं.

## 8 व्याख्या

शुरुआती लोगों के लिए फिट किए गए समीकरणों को स्थानीय रूपांतरण के रूप में सबसे अच्छी तरह से समझा जाता है प्रारंभिक-उन्नत उपयोगकर्ता, सार्वभौमिक रेटिंग पहचान के रूप में नहीं। उदाहरण के लिए, 10 मिनट का त्वरित अनुमान है

$$\hat{R}_{\text{Chess.com}} = -499.68 + 1.0778 R_{\text{Lichess}}.$$

1200 की Lichess रेटिंग पर, यह लगभग 793 के बराबर Chess.com देता है। 1500 की Lichess रेटिंग पर, यह लगभग 1117 देता है।

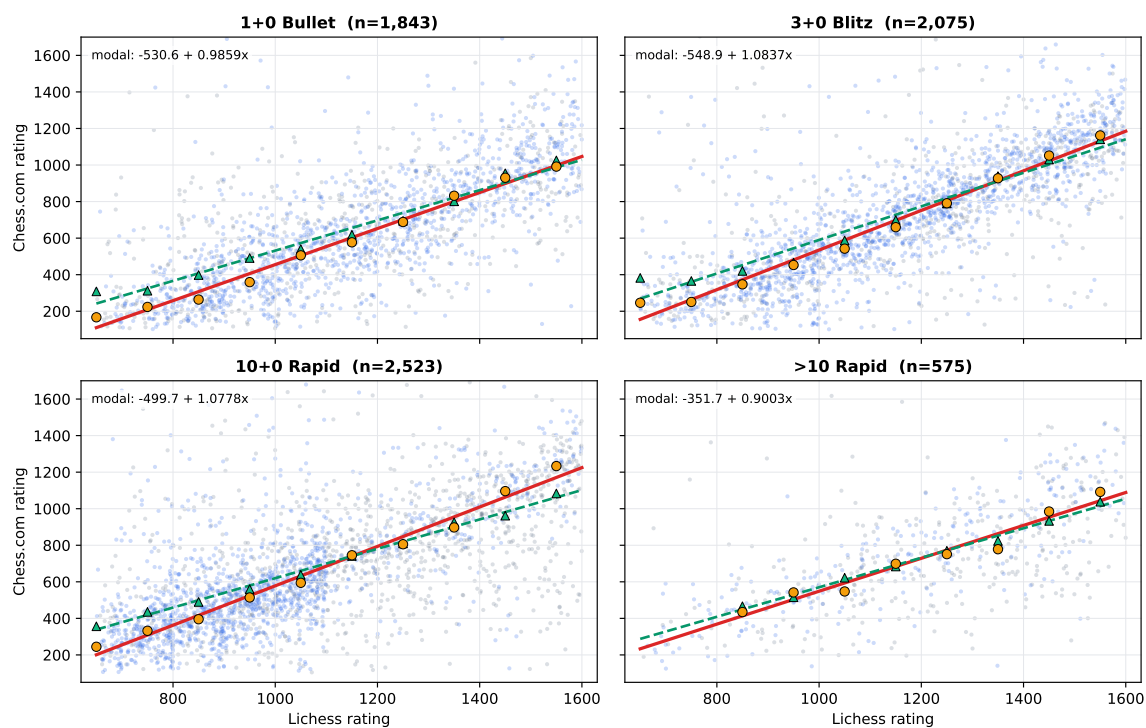
अनुमानों को फिट किए गए अंतराल से बहुत बाहर नहीं लगाया जाना चाहिए। पर बहुत उच्च रेटिंग, प्लेयर-पूल संरचना, खाता व्यवहार और प्लेटफॉर्म रेटिंग यांत्रिकी उन तरीकों से भिन्न हो सकती है जो एक एकल रैखिक द्वारा कैप्चर नहीं किए जाते हैं मैपिंग को निचले और मध्य रेटिंग रेंज में फिट किया गया है।

## 9 सीमाएँ और अद्यतन योजना

वर्तमान प्रक्रिया जानबूझकर रूढ़िवादी है, लेकिन कई सीमाएँ हैं शेष:

- समान-उपयोगकर्ता नाम मिलान शोरपूर्ण है और पहचान साबित नहीं करता है।
- टू-मैच अंश रेटिंग बकेट और गेम प्रकार के अनुसार भिन्न हो सकता है।

### Cross-platform same-username rating matches and modal mapping fits



Raw points are same-username candidate matches after game-history filters. Red line is the production modal fit; green dashed line shows the bucket-mean alternative for comparison.  
 Matched users    Bucket mode    Bucket mean    Modal fit    Mean-bucket fit

Figure 1: समान-उपयोगकर्ता नाम वाला उम्मीदवार मेल खाता है और मैपिंग वक्र फिट किए गए हैं। कच्चे अंक गेम-इतिहास फ़िल्टर के बाद उम्मीदवार उपयोगकर्ताओं को बनाए रखा जाता है। नारंगी बिंदु हैं उत्पादन फिट के लिए इन-बकेट मोडल Chess.com रेटिंग का उपयोग किया जाता है। लाल रेखा मोडल रिग्रेशन लाइन है। हरे त्रिकोण और धराशायी हरी रेखा दिखाते हैं तुलना के लिए बकेट-मीन विकल्प; इन्हें चित्रित करने के लिए शामिल किया गया है बेमेल समान नाम वाले खातों द्वारा माध्य-आधारित सारांश कैसे निकाले जा सकते हैं और पूँछ अवलोकन.

- Chess.com और Lichess रेटिंग हमेशा एक जैसी नहीं मापी जातीं सभी उम्मीदवार जोड़ियों के लिए दिन, हालांकि उसी महीने का नमूना इसे कम कर देता है मुद्दा.
- 10 मिनट से अधिक लंबा तीव्र नमूना वर्तमान में इससे छोटा है बुलेट, ब्लिट्ज़ और 10 मिनट के तेज़ नमूने।
- रैखिक मॉडल 650-1600 से अधिक का व्यावहारिक सन्निकटन है, ए नहीं किसी भी रेटिंग प्रणाली का संरचनात्मक मॉडल।

अगला नियोजित सुधार लंबे समय तक तीव्र उपयोगकर्ताओं का नमूनाकरण जारी रखना है, रूढ़िवादी API उपयोग और प्लेटफॉर्म शर्तों के अधीन। एक बार सैंपल का मिलान हो गया रेटिंग बकेट में बड़ा और अधिक संतुलित है, मोडल फिट होना चाहिए पुनः अनुमान लगाया गया और उत्पादन स्थिरांक फिर से अद्यतन किया गया।

## 10 रखरखाव और अद्यतन

यह दस्तावेज़ नवीनतम Elo+Chess क्रॉस-प्लेटफॉर्म रेटिंग मैपिंग को दर्शाता है 29 मई, 2026 तक अद्यतन। अनुमानों को अतिरिक्त के रूप में संशोधित किया जाएगा समान-कैलेंडर-माह से मेल खाने वाले उपयोगकर्ताओं को विशेष रूप से एकत्र किया जाता है 10 मिनट से अधिक लंबे तीव्र खेल।

इस रिपोर्ट का अनुरक्षित सार्वजनिक संस्करण यहां उपलब्ध है:

<https://www.elopluschess.com/static/docs/cross-platform-elo-mapping-modal-method.pdf>