

การประมาณค่าการแมปคะแนนหมากรุกข้ามแพลตฟอร์มด้วยการ ถดถอยแบบโมดอล

บันทึกการวิจัย Elo+Chess

29 พฤษภาคม 2569

Abstract

ผู้เล่นมักถามว่าเรตติ้งของ Lichess.org เทียบกับเรตติ้งของ Chess.com เป็นอย่างไร คำถามนี้มีความสำคัญในทางปฏิบัติสำหรับ Elo+Chess เนื่องจากเป็นเกณฑ์มาตรฐาน เส้นโค้งถูกสร้างขึ้นจากตัวอย่างแบบแบ่งชั้นขนาดใหญ่ที่นำมาจากตัวอย่างที่สมบูรณ์ ประวัติเกมรายเดือนของ Lichess.org ในขณะที่ผู้ใช้อาจนำประวัติมาจากทั้งสองอย่าง เว็บไซต์ หมายเหตุนี้อธิบายวิธีการปัจจุบันที่ใช้ในการประมาณค่า การแมปเฉพาะประเภทเกมตั้งแต่เรตติ้ง Lichess ไปจนถึงเรตติ้ง Chess.com ที่ วิธีการใช้ชื่อผู้ใช้ที่ตรงกัน การควบคุมเวลาแบบตรงทั้งหมดหรือใกล้เคียงกัน ตัวกรอง ข้อกำหนดประวัติเกมขั้นต่ำ และขั้นตอนการถดถอยแบบโมดอล ออกแบบมาเพื่อลดอคติจากการจับคู่ชื่อผู้ใช้เดียวกันที่เป็นเท็จ ปัจจุบัน การทำแผนที่การผลิตคือ:

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

การประมาณการนี้เป็นข้อมูลเบื้องต้นและจะได้รับการอัปเดตเมื่อตัวอย่างที่ตรงกันเพิ่มขึ้น โดยเฉพาะอย่างยิ่งสำหรับเกมที่รวดเร็วที่มีความยาวมากกว่า 10 นาที

การทำแผนที่การผลิตปัจจุบัน

สำหรับผู้อ่านที่ต้องการผลลัพธ์ก่อน สมการการผลิตปัจจุบันคือ:

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

เส้นเหล่านี้ติดตั้งด้วยพิกัด Lichess 650–1600 และไม่ควร คาดการณ์เป็นการแปลงสากลในทุกระดับทักษะ

1 วัตถุประสงค์

เป้าหมายคือการจับคู่การให้คะแนนบน Lichess.org กับระดับการให้คะแนนที่เทียบเคียงได้ Chess.com สำหรับประเภทการควบคุมเวลาที่กว้างเหมือนกัน เราประเมินแยกต่างหาก การแมปสำหรับแต่ละหมวดหมู่ที่ใช้โดย Elo+Chess:

- กระสุน 1+0,
- สายฟ้าแลบ 3+0
- 10+0 รวดเร็ว

- เกมที่รวดเร็วนานกว่า 10 นาที

เส้นที่พอดินันถูกใช้เป็นชั้นการแปลงสเกลเท่านั้น มันไม่ได้หมายความถึงสิ่งนั้น กลุ่มการจัดอันดับทั้งสองมีกลไกการจัดอันดับ กลุ่มผู้เล่น หรือเหมือนกัน พฤติกรรมการเลือกเกม

2 แหล่งที่มาของข้อมูล

ด้าน Lichess ของการวิเคราะห์สร้างขึ้นจาก Lichess สาธารณะที่สมบูรณ์ ไฟล์เก็บถาวรเกมระดับมาตรฐาน ตั้งแต่เดือนมกราคม 2025 ถึงเดือนมีนาคม 2026 เรื่องนี้สำคัญ เพราะตัวอย่าง Lichess ไม่ใช่การชุดตามสะดวกหรือตัวอย่างแคบๆ โปรไฟล์ที่มองเห็นได้: สร้างจากไฟล์ประวัติเกมรายเดือนแบบเต็ม เผยแพร่โดย Lichess เพื่อการวิจัยและการใช้งานสาธารณะ

Lichess เผยแพร่การส่งออกฐานข้อมูลเหล่านั้นบน <https://database.lichess.org/> ฐานข้อมูลระบุว่า การส่งออกได้รับการเผยแพร่ภายใต้ “ใบอนุญาต CC0” และอนุญาตอย่างชัดเจนให้ผู้ใช้สามารถ “ใช้เพื่อการวิจัย” และ “publication”.¹ ข้อมูลแบบเปิดนี้ ทำทางเป็นศูนย์กลางในการออกแบบด้านมาตรฐานของ Elo+Chess

ในด้าน Chess.com บัญชีที่มีชื่อผู้ใช้เดียวกันของผู้สมัครจะถูกตรวจสอบเท่านั้น คลังข้อมูลรายเดือนและหมวดหมู่เกมที่จำเป็นในการประมาณค่าการทำแผนที่เรา ให้ความสำคัญเป็นพิเศษกับการจับคู่ประวัติเกมภายในปฏิทินเดียวกัน เดือนเมื่อใดก็ตามที่เป็นไปได้ การจับคู่เดือนตามปฏิทินช่วยลดความเสี่ยงที่ก คะแนนของผู้เล่นเปลี่ยนไปอย่างมากระหว่าง Lichess และ Chess.com การวัด

3 โครงสร้างการแข่งขันของผู้สมัคร

ความท้าทายหลักคือไม่มีแพลตฟอร์มใดที่เปิดเผยข้ามแพลตฟอร์มที่เป็นสากล ตัวระบุผู้เล่น ดังนั้นเราจึงเริ่มต้นด้วยผู้สมัครที่มีชื่อผู้ใช้เดียวกัน: ถ้า ชื่อผู้ใช้ปรากฏบนทั้ง Lichess.org และ Chess.com คุณนั้นจะถือเป็นการแข่งขันของผู้สมัคร นี่เป็นการจงใจส่งเสียงดัง บัญชีชื่อเดียวกันบางบัญชีคือ คนคนเดียวกัน ในขณะที่คนอื่น ๆ เป็นผู้เล่นที่ไม่เกี่ยวข้องกันและบังเอิญใช้สิ่งเดียวกัน จัดการ

เพื่อลดเสียงรบกวนที่หลีกเลี่ยงได้ ไปปีโลนปัจจุบันจะใช้เฉพาะประเภทเกม ตัวกรอง:

1. ผู้ใช้ Lichess จะถูกสุ่มตัวอย่างจากชุดข้อมูลประวัติเกมแบบแบ่งชั้นโดย บัคเก็ตเรตติ้งและประเภทการควบคุมเวลา
2. บัญชีผู้สมัคร Chess.com ได้รับการตรวจสอบตามความเหมาะสม หมวดการควบคุมเวลา
3. คู่ผู้สมัครจะยังคงอยู่เมื่อทั้งสองฝ่ายมีความเกี่ยวข้องเพียงพอเท่านั้น ประวัติเกม
4. สำหรับกระสุน สายฟ้าแลบ และความเร็ว 10 นาที ประวัติขั้นต่ำที่คงไว้คือ 5 เกมต่อด้าน
5. สำหรับการรวดเร็วนานกว่า 10 นาที ประวัติการเก็บรักษาขั้นต่ำคือ 10 เกมต่อด้าน
6. ความพอดีของการถดถอยที่รายงานที่นี้จำกัดอยู่ที่การจัดอันดับ Lichess จาก 650 ถึง 1600 ตรงกับขอบเขตการให้คะแนนที่เกี่ยวข้องกับ Elo+Chess หลักมากที่สุด รายงานเกณฑ์มาตรฐาน

ตัวอย่างปัจจุบันเน้นการจับคู่ชื่อผู้ใช้เดียวกันในเดือนปฏิทินเดียวกัน เมื่อ บัญชีผู้สมัครมีเกมเพียงพอในหมวดการควบคุมเวลาเป้าหมาย เดือนเดียวกับแหล่งข้อมูล Lichess ที่ตรงกับทำให้คะแนนรายเดือนคือที่ต้องการสำหรับการติดตั้ง ข้อมูลการจับคู่ข้ามแพลตฟอร์มที่เก่ากว่าจะถูกเก็บไว้ที่เดิม เพิ่มการสนับสนุนตัวอย่าง แต่เป้าหมายด้านระเบียบวิธีเป็นเป้าหมายเฉพาะประเภทเกม การเปรียบเทียบในเดือนเดียวกัน

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

4 เหตุใดค่าเฉลี่ยหรือค่ามัธยฐานจึงสามารถมีอคติได้

สำหรับที่เก็บข้อมูลการให้คะแนน Lichess แบบคงที่ ชุดตัวเลือกชื่อผู้ใช้เดียวกันจะผสมกัน:

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

โดยที่ x คือเรตติ้ง Lichess, y คือเรตติ้ง Chess.com, f_{true} is the distribution of Chess.com ratings for true การจับคู่ข้ามแพลตฟอร์ม และ f_{false} คือการกระจายที่สร้างขึ้น โดยบัญชี Chess.com ที่ไม่เกี่ยวข้องกับชื่อผู้ใช้เดียวกัน

องค์ประกอบการจับคู่ที่ผิดพลาด ไม่ได้อยู่ตรงกลางที่เทียบเท่า Chess.com ที่ถูกต้อง สำหรับบัคเกิด Lichess มีแนวโน้มที่จะมีลักษณะคล้ายกับประชากร Chess.com ในวงกว้าง สุ่มตัวอย่างตามชื่อผู้ใช้และตัวกรองกิจกรรม เป็นผลให้ค่าเฉลี่ยแบบมีเงื่อนไขคือ ดึงไปสู่การกระจายผู้เล่นที่ใช้งานทั่วโลก:

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

ถ้า $\pi < 1$ โดยทั่วไปแล้วจะไม่เท่ากับผู้เล่นที่ตรงกันที่ต้องการ ความคาดหวัง ปัญหาเดียวกันนี้จะส่งผลต่อค่ามัธยฐานทุกครั้งที่มีการจับคู่ที่ผิดพลาด ส่วนประกอบมีมวลมากพอที่จะเคลื่อนเปอร์เซ็นต์ไทล์ที่ 50

ทิศทางของอคติขึ้นอยู่กับที่ฝากข้อมูล ในที่เก็บข้อมูล Lichess ต่ำ เป็นเท็จ การแข่งขันสามารถดึงค่าเฉลี่ยขึ้นได้หากบัญชี Chess.com ที่ใช้งานแบบสุ่มอยู่ แข็งแกร่งกว่าผู้ใช้ที่ตรงกันอย่างแท้จริง ในถึงสูงก็มีสิ่งปนเปื้อนเหมือนกัน สามารถดึงค่าเฉลี่ยลงได้ เอฟเฟกต์การบีบอัดนี้เป็นสิ่งที่เราต้องการอย่างแท้จริง หลีกเลี่ยงเมื่อสร้างการแปลงขนาด

5 แรงจูงใจแบบกิริยา

วิธีการแบบกิริยาขึ้นอยู่กับสมมติฐานที่อ่อนแอกว่าแบบเฉลี่ย วิธีการ เราไม่ต้องการให้คู่ชื่อผู้ใช้เดียวกันทุกคู่ถูกต้อง แทนเรา สมมติว่าภายในบัคเกิดการจัดอันดับ Lichess ที่แคบเพียงพอ จะมีค่าตรงกันจริง สร้างคลัสเตอร์ท้องถิ่นที่มีความเข้มข้นมากที่สุดในพื้นที่การให้คะแนน Chess.com เท็จ การแข่งขันอาจมีมากมาย แต่ควรกระจายให้กระจายมากกว่านี้เพราะว่า พวกเขาไม่ได้เชื่อมโยงกับที่เก็บข้อมูล Lichess

สำหรับแต่ละบัคเกิด Lichess เราประเมินโหมดของการจัดอันดับ Chess.com การกระจายโดยใช้การประมาณความหนาแน่นของเคอร์เนล จากนั้นจึงทำการรักษาจุดโมดอล เป็น Chess.com ที่น่าเชื่อถือที่สุดสำหรับที่ฝากข้อมูล Lichess นั้น สิ่งนี้ทำให้ ขั้นตอนนี้ไวต่อหางยาวน้อยกว่าและมีความไวต่อหางที่ไม่ตรงกันน้อยกว่า บัญชีชื่อเดียวกันสุ่มตัวอย่างจากประชากร Chess.com ในวงกว้าง

นี่ไม่ใช่การอ้างว่าโหมดนี้สมบูรณ์แบบ มันเป็นความแข็งแกร่งในทางปฏิบัติ ทางเลือกสำหรับปัญหาการแก้ไขเอนทิตีที่มีเสียงดัง วิธีนี้ใช้ได้ผลดีที่สุดเมื่อเป็นจริง การแข่งขันยังคงมีคะแนนใกล้เคียง ระดับข้ามแพลตฟอร์มที่ถูกต้องและเป็นเท็จ การแข่งขันไม่ได้มุ่งไปที่มูลค่าที่ทำให้เข้าใจผิด

6 ขั้นตอนการติดตั้ง

ความพอดีในการผลิตปัจจุบันใช้ขั้นตอนต่อไปนี้แยกกันสำหรับแต่ละรายการ ประเภทเกม:

1. รักษาชื่อผู้ใช้เดียวกันของผู้สมัครให้เป็นไปตามจำนวนเกม เกณฑ์
2. จำกัดความพอดีไว้ที่ $650 \leq R_ZXQPROT3ZXQ \leq 1600$
3. แบ่งพาร์ติชัน Lichess เป็นบัคเกิด 100 จุด

- สำหรับแต่ละที่เก็บข้อมูลที่มีผู้ใช้อย่างน้อย 30 คน ให้ประมาณโมดอล Chess.com การจัดอันดับด้วยการประมาณความหนาแน่นของเคอร์เนลแบบเกาส์เซียนโดยใช้การจัดอันดับแบนด์วิดท์ 75 คะแนน
- แสดงแต่ละที่เก็บข้อมูลด้วยจุดกึ่งกลางและโหนด Chess.com โดยประมาณ
- ใส่เส้นกำลังสองน้อยที่สุดธรรมดาผ่านโมดอลระดับบัคเก็ต คะแนน:

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

เฉพาะจุดโมดอลระดับบัคเก็ตเท่านั้นที่เข้าสู่การถดถอยเชิงเส้นขั้นสุดท้าย รายบุคคล คู่ชื่อผู้ใช้เดียวกันมีอิทธิพลต่อเส้นโค้งผ่านการมีส่วนร่วมใน การประมาณความหนาแน่นภายในถึง

7 การประมาณการปัจจุบัน

ตาราง 1 รายงานค่าคงที่และความลาดชันที่ติดตั้งไว้ในปัจจุบัน ค่าเหล่านี้เป็นค่าที่เชื่อมต่อเข้ากับ Elo+Chess ในปัจจุบันหลังจากวันที่ 29 พฤษภาคม 2026 อัปเดต

Game type	Total users	Baseline	Same-month	In range	α	β
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: โมดอลลิเนียร์ปัจจุบันพอดีมากกว่าพิกัด Lichess 650–1600 “ใน range” นับจำนวนผู้ใช้ที่คงไว้ภายในช่วงเวลา Lichess ที่ติดตั้งไว้ แถวพื้นฐาน มาจากตารางข้ามแพลตฟอร์มการควบคุมที่แน่นอนที่มีอยู่แถวเดือนเดียวกันมา จากตัวอย่างชื่อผู้ใช้เดียวกันรายเดือนเดือนมีนาคม 2569

การประมาณการอย่างรวดเร็วที่นานกว่า 10 นาทีที่มีการสนับสนุนน้อยกว่าอย่างมาก อีกสามหมวดหมู่ จึงมีแนวโน้มมากที่สุดที่จะเปลี่ยนแปลงเพิ่มเติม มีการรวบรวมผู้ใช้ที่ตรงกัน

8 การตีความ

สมการที่ติดตั้งจะถูกตีความ ได้ดีที่สุดว่าเป็นการแปลงในท้องถิ่นสำหรับผู้เริ่มต้น ผู้ใช้ขั้นสูงในช่วงเริ่มต้นไม่ใช่อัตลักษณ์การให้คะแนนสากล ตัวอย่างเช่น การประมาณการอย่างรวดเร็ว 10 นาทีคือ

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

ที่ระดับ Lichess ที่ 1200 ทำให้ค่านี้เทียบเท่ากับ Chess.com ที่ประมาณ 793 ที่ระดับ Lichess ที่ 1500 มันให้ประมาณ 1117

การประมาณการไม่ควรคาดการณ์ไกลออกไปนอกช่วงเวลาที่ติดตั้งไว้ ณ การให้คะแนนที่สูงมาก องค์ประกอบของผู้เล่น-พูล พฤติกรรมบัญชี และแพลตฟอร์ม กลไกการให้คะแนนอาจแตกต่างกันไปในลักษณะที่ไม่ได้ถูกบันทึกด้วยเส้นตรงเส้นเดียว การทำแผนที่ให้พอดีกับช่วงเรตตั้งล่างและกลาง

9 ข้อจำกัดและแผนการอัปเดต

ขั้นตอนปัจจุบันมีเจตนาอนุรักษ์นิยม แต่มีข้อจำกัดหลายประการ ยังคงอยู่:

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

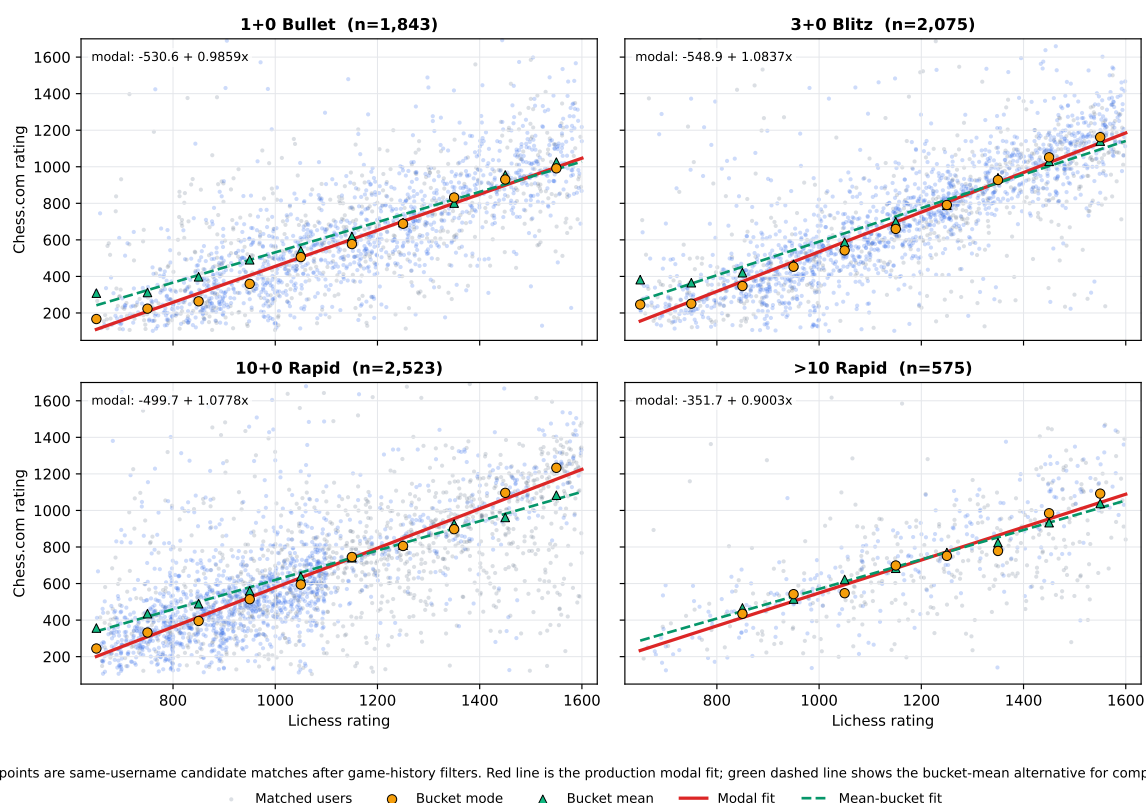


Figure 1: ชื่อผู้ใช้เดียวกันที่ตรงกันและเส้นโค้งการแมปที่พอดี จุดดิบ ยังคงเป็นผู้ใช้ที่เป็นผู้สมัครหลังจากตัวกรองประวัติเกม จุดสีส้มคือ การจัดอันดับ Modal Chess.com ภายในbucketที่ใช้สำหรับการผลิตให้พอดี เส้นสีแดง คือเส้นถดถอยแบบโมดอล สามเหลี่ยมสีเขียวและเส้นประสีเขียวแสดง ทางเลือกถ่วงเฉลี่ยสำหรับการเปรียบเทียบ สิ่งเหล่านี้รวมอยู่ด้วยเพื่อแสดงให้เห็น วิธีดึงข้อมูลสรุปตามค่าเฉลี่ยโดยบัญชีที่มีชื่อเดียวกันที่ไม่ตรงกันและ การสังเกตทาง

- การจับคู่ชื่อผู้ใช้เดียวกันทำให้เกิดเสียงรบกวนและไม่ได้พิสูจน์ตัวตน
- เศษส่วนการจับคู่จริงอาจแตกต่างกันไปตามกลุ่มคะแนนและประเภทเกม
- การจัดอันดับของ Chess.com และ Lichess ไม่ได้วัดจากค่าที่เท่ากันเสมอไป วันสำหรับคู่ผู้สมัครทั้งหมด แม้ว่าตัวอย่างในเดือนเดียวกันจะลดสิ่งนี้ลง ปัญหา
- ตัวอย่างด่วนที่นานกว่า 10 นาทีปัจจุบันมีขนาดเล็กกว่า bullet, blitz และตัวอย่างด่วน 10 นาที
- แบบจำลองเชิงเส้นเป็นการประมาณเชิงปฏิบัติในช่วง 650–1600 ไม่ใช่ a แบบจำลองโครงสร้างของระบบการให้คะแนนอย่างใดอย่างหนึ่ง

การปรับปรุงตามแผนครั้งต่อไปคือการสุ่มตัวอย่างผู้ใช้ที่รวดเร็วขึ้นต่อไป ขึ้นอยู่กับการใช้งาน API และข้อกำหนดแพลตฟอร์มแบบอนุรักษ์นิยม เมื่อได้ตัวอย่างที่ตรงกันแล้ว มีขนาดใหญ่กว่าและมีความสมดุลมากกว่าในบ๊อคเกิดการให้คะแนน ซึ่ง Modal ควรจะพอดี ประมาณการใหม่และค่าคงที่การผลิตได้รับการอัปเดตอีกครั้ง

10 การบำรุงรักษาและการอัปเดต

เอกสารนี้แสดงถึงการแมปการให้คะแนนข้ามแพลตฟอร์ม Elo+Chess ล่าสุด อัปเดต ณ วันที่ 29 พฤษภาคม 2026 การประมาณการจะได้รับการแก้ไขเพิ่มเติม ระบบจะรวบรวมผู้ใช้ที่ตรงกันในเดือนปฏิทินเดียวกัน โดยเฉพาะสำหรับ เกมที่รวดเร็วนานกว่า 10 นาที

รายงานฉบับสาธารณะที่ได้รับการดูแลมีอยู่ที่:

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