

# Digital Gastronomy

## Ingredients Embedding

Yuval Friedmann

Ariel Bezaleli-Mizrahi

GitHub repository: [github.com/yuvfried/rec2vec](https://github.com/yuvfried/rec2vec)

### מבוא

גסטרונומיה/בישול דיגיטלי (DG) היא תחום מחקר חדש שחוקר אינטגרציה של טכנולוגיות ממוחשבות במטבח (1) בתור מערכות לתכנון מתכונים ומנות (2) וככלי בישול חדשים. באופן מסורתי תכנון מתכונים מתבסס על היכרות מעמיקה של טבחים ושפים עם ריאקציות המתקבלות בהפעלת טכניקות בישול (טכ"ב) על חומרי גלם (חומ"ג). ההיכרות הזו מסוננת לתוך הוראות ההכנה במתכונים ובעשורים האחרונים מהפכת הבישול המולקולרי מסירה את המסתורין האופף ריאקציות אלו באמצעות שימוש בידע אשר נצבר בתצפיות מדעיות בתהליכי בישול. בעבודה זו נציג גישה אחרת לניתוח קשרים בין טכ"ב וחומ"ג באמצעות ניתוח מבני של מתכונים ושימוש בטכניקת (WE) word embedding. אנו משערים שגישה זו תאפשר לנו לזקק באופן חלקי את האינטואיציה אותה מסנתזים השפים אל המתכונים. בפרק הבא נביא רקע לביסוס ההנחות בהן אנו משתמשים לניתוח המתכונים. בהמשך נציג עבודות קשורות ממספר תחומים ולאחר מכן נביא גרסה מפורטת של ההשערה שהצגנו כאן והשיטה לבחינתה. לאחר שנציג תוצאות אמפיריות, נדון בהן ונעריך את תקפות ההשערה. נסכם במסקנות העולות מהתוצאות ונפרט עבודה עתידית הנדרשת בתחום.

### רקע

בשנים האחרונות פרויקטים רבים בוחנים ומדגימים את פוטנציאל התרומה של כלים חישוביים וטכנולוגיות ייצור דיגיטליות לעולם הקולינרי. בין הפרויקטים הללו טכנולוגיות להדפסת אוכל ב-2D, 3D ו-4D (Wang et al. 2017; Mizrahi et al. 2016) שפים רובוטיים (Samsung 2019; Gibson 2015) ומחקר תאורטי על סמנטיקה ומערכות יחסים פרוצדורליות במתכונים (Kiddon et al. 2015; Huang et al. 2017). כניסת כלים אלו אל המטבח הופכת פרסונליזציה של מתכונים ומנות קלה מאי פעם. לדוגמא, צוות Habit (Habit 2019) מתבסס על מחקרים בתחום התזונה הגנטית (nutrigenomics) לבניית תזונה בהתאמה אישית בהתבסס על מידע אישי כגון בדיקות DNA והתייחסות למצב רפואי (סוכרת, רגישות ללקטוז, צליאק וכו'). לצד העדפות בריאותיות קיימות מוטיבציות נוספות להתאמה אישית של מתכונים כמו התאמות תרבותיות, דתיות (כשרות, חלאל, בודהיזם), אקולוגיות, מוסריות או מטעמי גיוון קולינרי.

מתכונים מורכבים מרשימת מרכיבים וסדרת הוראות מפורטת על מנת להגיע לתוצאה סופית שהיא המנה. לאחר סקירה של כ-40 מתכונים זיהינו חזרתיות במבנה ההוראות, דפוס זה ניתן לחלק לשלושה חלקים, שלצורך העבודה נכנה חלקי-בשל:

1. **כלי בישול** (U) – כלי הבישול בו נעשה שימוש במהלך ביצוע פעולת הבישול (לדוגמא: מחבת, קרש חיתוך, סכין וכד'), לעיתים יופיע עם תיאור נלווה.

2. **פעולת בישול (C)** – פעולה לביצוע על תת-קבוצה של חומ"ג (לדוגמא: חיתוך, קיצוץ, אפייה, טיגון וכד') לעיתים בליווי פרמטרים כמו טמפרטורה או תיאור כמו השחמה או הזהבה (בטיגון למשל).

3. **חומ"ג (I)** – תת-קבוצה של חומ"ג לעיתים בליווי כמויות או שמות תואר (לדוגמא: עגבנייה בשלה, שלושה שיני שום וכד')

הוראות הבישול שדגמנו מאתר [allrecipes.com](http://allrecipes.com) ומהקורפוס הציגו כל סידור אפשרי של חלקי-בשל, אך עם הטיה לסדר שלעיל. נגדיר את  $d(C, I)$ , פונקציית המרחק בין **פעולת בישול לחומ"ג** בהוראת בישול נתונה, להיות גודל החלון המינימלי מסביב ל**פעולת בישול** בו מופיעה קבוצת **חומ"ג**. ב-72.5% מההוראות שדגמנו התקיים  $d(C, I) \leq 3$  (ראה מספר דוגמאות ב-Table 1).

הוראת בישול	חלקי-בשל לפי סדר הופעה			$d(C, I)$
transfer to the wire rack in the <b>oven</b> while you <b>fry</b> the remaining <b>chicken</b>	oven	fry	chicken	3
set the <b>colander</b> over a large <b>saucepan</b> of <b>boiling water</b> and <b>steam</b> the <b>rice</b> covered with a kitchen towel and the lid for 15 minutes or until it is fluffy and dry	colander saucepan	boiling steam	water rice	1 2
<b>melt butter</b> in heavy medium <b>saucepan</b> over medium low heat	melt	butter	saucepan	1
<b>melt butter</b> in heavy small <b>skillet</b> over low heat	melt	butter	skillet	1
working in batches <b>fry bean cakes</b> until firm and crisp adding more oil as needed about 6 minutes per side	fry	bean cakes		1
blanch <b>bean sprouts</b> in a <b>pot</b> of <b>boiling</b> water (for) 15 seconds and drain well	bean sprouts	pot	boiling	5
spread <b>hazelnuts</b> in a shallow <b>bake pan</b> and <b>toast</b> in <b>oven</b> until they begin to turn pale golden 7 to 10 minutes	hazelnuts	bake pan oven	toast	5
bring <b>cream</b> and <b>milk</b> to <b>simmer</b> in medium <b>saucepan</b>	cream milk	simmer	saucepan	2 and 4

Table 1: פירוק הוראות בישול לחלקי-בישל וגודל החלון המינימלי הנדרש מסביב לפעולת הבישול בו מופיעה קבוצת החומ"ג עליה היא מופעלת

## עבודות קשורות

בישול היא אחת המלאכות העתיקות ביותר כאשר תיעוד ראשון למתכון כתוב מתוארך למאה ה-15 לפנה"ס. אופן העברת הידע בעולם הבישול בצורת מתכונים לאורך תקופה של למעלה מ-3000 שנה ביחד עם התעוררות תרבותית סביב בישול כיום (תכניות ראליטי, אתרי בישול וכד') מציגים כמויות גדולות של מידע קולינרי. כמויות המידע הללו מעוררת תיאבון בקרב קהילת הדאטה, Jaan Altosaar (MIT) מציג

מיפוי של מטבחי העולם ומערכת המלצה למתכונים על בסיס סל מרכיבים זמין או דמיון בין מתכונים באמצעות הפעלת word2vec על רשימות מרכיבים (בלבד) (Ahn et al. 2017). (Jaan Altosaar) מצא דפוסים שונים בבחירת שילוב בין תרכובת טעמים במטבחים שונים באמצעות פיתוח "רשת טעמים דו-צדדית" (FlavorNetwork) (למשל, במטבח המערבי קיימת נטייה לזווג בין חומ"גים שחולקים מספר גדול של תרכובות טעמים, ולעומת זאת במטבח המטבח מזרח אסיאתי נוטה להימנע משילוב של חומ"גים שחולקים אותן תרכובות טעם (Ahn et al. 2011)).

## היפותזה

WE היא שיטה למידול של יחסים סמנטיים בין מילים, המתייחסת להקשרן של מילים ולהופעתן עם מילים אחרות בטקסט נתון. התווית הווקטורים המתארים מילים מניבה דיאגרמה בה קרבה מרחבית מתארת קרבה סמנטית ולהפך. מניתוח מבנה הוראות הבישול במתכונים שהצגנו ברקע אנו צופים ש-WE של קורפוס הוראות הבישול ממתכונים אמיתיים יניב מיפוי בו חומ"ג קרובים מרחבית יהיו דומים קולינרית. הדמיון הקולינרי אותו אנו צופים לזהות הוא בבחינת פעולות בישול שניתן לבצע על חומ"ג ובתוך כך מיפוי אספקט מסוים של חליפיות בין חומ"גים. אבחנה זו יכולה להוות בסיס למערכת המאפשרת להתאים מתכונים להעדפות קולינריות, בריאותיות, תרבותיות ואף אקולוגיות של מתכונים.

## שיטה

### דאטה: הוראות הבישול

לצורך העבודה השתמשנו ב-597 מתכונים בשפה האנגלית מאתר המתכונים [Epicurious](#). את המתכונים הורדנו מפלטפורמת הדאטה [Kaggle](#). לכל מתכון ישנן 1-24 הוראות בישול, כאשר סך כל כמות ההוראות שהדאטה מכיל היא 69458.

### פרה-פרוססינג וניקוי הדאטה

לצורך הזנת הדאטה לאלגוריתם של, היה עלינו לבחור איזו יחידה בדאטה שלנו תיחשב 'מסמך'. היינו מעוניינים למנוע מצב בו ה-context-window אותו פותח האלגוריתם סביב המילה הנתונה, יתחשב גם במילים שלא נמצאות באותו המשפט עם מילה זו, שכן דבר זה היה עלול לגרום לבלבול בין פעולת הבישול לרכיב עליה היא מתבצעת (למשל, האלגוריתם היה עלול להבין מהטקסט "טגן את העגבניות עם הבצל. חמם תנור לטמפ' של..." שפעולת 'חמם' מבוצעת על הבצל). לפיכך, החלטנו להגדיר 'מסמך' כ'משפט'. בסך הכל נוצרו מהדאטה 256578 משפטים שמהווים חלק מהוראות בישול.

החלטנו להסיר את הניקוד, כיון שלעיתים הניקוד במשפט יוצר אבחנה בין אזכור של אותה מילה. למשל, "קצוץ את הבצל..." ו-"חתוך את הבצל" (קצוץ אותו באופן דק) ("ייצור שני מופעים שונים – קצוץ ו-"קצוץ"). החלטה זו לא משנה הרבה את מהלך השפה הטבעית, כיון שמראש חילקנו את המסמכים ליחידות של משפטים, ובאופן גס ניתן לומר שלמשפטים בפני עצמם ישנה בדרך-כלל הבנה מרכזית אחת גם ללא ניקוד. לעומת זאת, לא הורדנו מילות קישור (אין הכוונה כאן למושג התחבירי, אלא לכינוי למילים שאותן נהוג לכוונות stopwords), כיוון שרצינו לשמור על שטף השפה טבעי כמה שאפשר ומשום שאלגוריתם השיבוץ מטפל במילים אלו באופן מרומז (subsampling). למרות זאת, שני שינויים אכן הוכנסו על ידינו באופן מכוון בתוך הטקסט: איחודם של צירופי מילים שלדעתנו ייצגו רכיב מזון אחד, ואיחודן של הטיות לשון למילים שמייצגות פעולות בישול, כפי שנפרט להלן.

## רשימת חומ"ג (ingredient list)

כיון שברצוננו היה לצייר את הרכיבים שמופיעים בקורפוס על מרחב דו-ממדי, היה עלינו ליצור רשימת רכיבים שאותם נשלף מתוך הקורפוס.<sup>1</sup> השתמשנו ב-[Recipe Ingredients Dataset](#) הנמצא גם הוא באתר Kaggle, כדי ליצור את רשימת הרכיבים. כיוון שהקורפוס שלנו קטן יחסית ואיתור phrases ע"י אלגוריתם השיבוץ עשוי להיות מורכב, השתמשנו ברשימה זו כדי לאחד בין מילים שונות המייצגות את אותו הרכיב על-ידי מקף (למשל "זיתים שחורים" הפכו ל"זיתים-שחורים").

## איחוד הטיות לשון (stemming) של טכ"ב

השערנו הייתה שקיים יחס בטקסט בין הרכיבים כיון שנעשות עליהם פעולות ביטול דומות. על מנת שנוכל להציג בוויזואליזציה הסופית את פעולות הביטול שמבוצעות על כל רכיב, החלטנו לאחד הטיות לשון של פעולות ביטול. כלומר, רצינו שכל פעולת ביטול בפלט של האלגוריתם תתורגם לאינסטנס אחד. לצורך כך, שלפנו את פעולות הביטול המוזכרות בערך ויקיפדיה List of cooking techniques (Wikibooks 2017). אלגוריתם Porter Stemming (Porter 1979) הוא תהליך להסרת סיומות (commoner morphological and inflexional) והשימוש העיקרי שלו הוא נורמליזציה של מושגים – במקרה שלנו פעולות הביטול בקורפוס וברשימה הנ"ל. ה-stemming נעשה בעזרת מימוש האלגוריתם הזמין בספריית NLTK.

## **שיבוץ המילים באלגוריתם Word2Vec**

WE ניתן לביצוע במספר שיטות, בהן word2vec, TF-IDF ו-GloVe. בעוד שתי האחרונות מניבות embedding איכותי, GloVe עובדת טוב יותר עם קורפוסים גדולים מאוד ומשום שהקורפוס שלנו קטן יחסית word2vec הייתה אופציה עדיפה. שיקול נוסף לבחירה ב-word2vec הוא גישת negative sampling שהופכת אותו מהיר יותר ממימושים אחרים (McCormick 2017).

אלגוריתם Word2vec מניב ייצוג ווקטורי למילים. הווקטורים ממודלים כך שווקטורים המייצגים מילים הנחשבות 'קרובות' זו לזו יהיו קרובים זה לזה מבחינה מרחבית (למשל cosine similarity). 'קרבה' בין מילים מחושבת לפי ההקשר שלה בתוך הטקסט, כך שמילים שמתפקדות בטקסט באופן דומה, אמורות לקבל ווקטורים קרובים יחסית לזוג מילים שאינן מתפקדות באופן דומה. לפיכך, האלגוריתם מושפע ממספר היפר-פרמטרים הנקבעים מראש – ביניהם context window וכן ממד המרחב הווקטורי של המילים. בעבודה שלנו בחנו מספר קומבינציות של context-window באורכים 2, 3 ו-5 וכן ממדים מגודל 50 ו-100 (Goldberg and Levy 2014).

## **וויזואליזציה: הורדת ממד ל- $R^2$**

על מנת לנתח את המודל בחרנו לבנות וויזואליזציה באמצעות הורדת ממד הווקטורים שהתקבלו מממד 100 או 50 לממד 2 וכך להציגם בגרף פיזור דו-ממדי. גם כאן, קיימות מספר אלטרנטיבות כמו PCA אבל הנפוצה ביותר בעבודות עם word2vec היא t-Distributed Stochastic Neighbor Embedding (t-SNE) (Wattenberg, Viégas, and Johnson 2017; Heuer 2016) ולכן בחרנו בה.

---

<sup>1</sup> הדאטה המקורי שהוזכר לעיל הכיל אף הוא עמודות רכיבים שהוזכרו בעמודת ההוראות. אך הרכיבים הוזכרו שם בתוספת קדימויות ש'הרעישו' את המילים הללו. למשל, "חצי כוס קמח", "כוס קמח" ו-"קמח". גם אם היה ברצוננו להשתמש ברשימה זו לצורך זיהוי הרכיב בפני עצמו ("קמח"), והפרדתו מהתוספות המיותרות, היה עלינו בכל אופן להשתמש ברשימת רכיבים חיצונית.

t-SNE יוצר מרחב הסתברות המייצג הסתברויות משותפות עבור כל זוג ווקטורים בממד המקורי, ומנסה 'לחקות' מרחב הסתברות זה בממד הקטן יותר לפי הדרישה (במקרה שלנו דו-ממדי). היפר-פרמטר חשוב של אלגוריתם זה הוא ה-perplexity, והוא אחראי במידה רבה על האופן בו יוצג הממד המוקטן ביחס לממד המקורי. ערכים גבוהים ייקחו בחשבון את המבנה הכללי של הדאטה, ואילו ערכים קטנים יתייחסו יותר לנקודה ברמה הלוקלית – כל נקודה בפני עצמה – ופחות במבנה הכללי של הדאטה. ברשת מוזכר כי העמימות בעניין בחירת ההיפר פרמטרים באלגוריתם זה עדיין גדולה, וכי הבחירה נשענת בעיקר על כללי אצבע. שני כללי אצבע מרכזיים שנפגשו בהם היו  $perp = N^{(1/2)}$  וכן  $perp = N * 0.05$ . במקרה שלנו הערכים הם 120 ו-49 בהתאמה. לעיל הזכרנו שהרכבנו ארבע קומבינציות של word2vec, ולכן לאחר הפעלת t-SNE קיבלנו לבסוף 8 תוצאות שונות המייצגות קומבינציות שונות של ההיפר פרמטרים של האלגוריתמים.

## תוצאות ודיון

בחלק זה נציג את תוצאות המודלים שאימנו, לכל תמונה קיים קישור לגרף אינטראקטיבי (HTML) שנמצא ב-repository בפרויקט ה-GitHub המצורף. בגרפים ב Figure 1 ניתן לראות וויזואליזציה של המודלים ברורים ביותר מבחינת אשכולות (clustering) עם צביעה לפי טכ"ב הקרובה ביותר לכל חומ"ג. ניתן לראות חלוקה ברורה למספר אשכולות עם 4-5 טכניקות בישול עיקריות. טכניקות בישול בשלב ההכנה (mise en place) לדוגמה סוגים שונים של קציצה (julienning, mincing), נפוצות מאוד שכן רוב חומ"ג עוברים את התהליך. לאור תובנה זו החלטנו לבחור חתך תצוגה נוסף שאינו מוטה לפעולות כלליות. בגרף שב Figure 2 בחרנו בששת טכניקות הבישול הנפוצות ביותר בקורפוס ללא פעולות הכנה:

**baking, boiling, simmering, roasting, grilling and frying**

ומשום שחלק מהוראות המתכונים הראו גם  $d(C, I) \leq 2$  החלטנו להוסיף מודל שאומן עם חלון בגודל 2 אותו ניתן לראות לצד מודלים עם חלונות וממד מרחב ווקטורי משתנה לצורך השוואה. בחתך הצביעה הזה אמנם ניתן לראות חלוקה פחות גסה אך החומ"גים המופיעים יחד נראים מתאימים יותר, למשל כל החומ"ג שעוברים הקרמה (creaming) אכן קרובים וחליפיים.

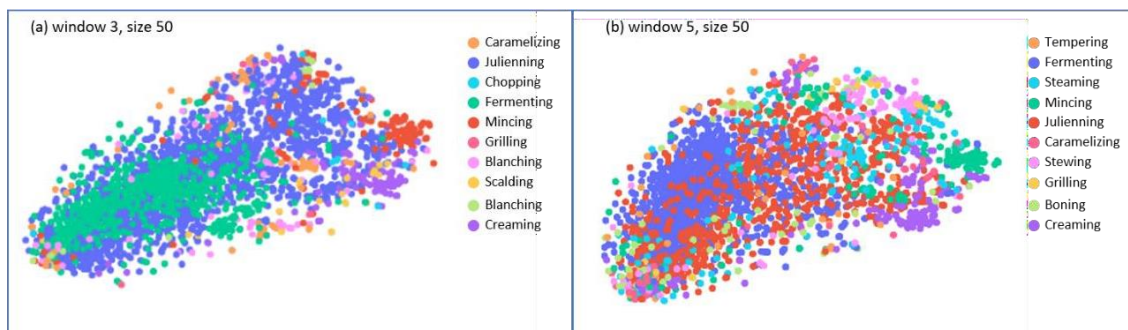


Figure 1: וויזואליזציה למודלים לאחר הורדת ממד באמצעות t-SNE, הצביעה היא לפי טכ"ב הקרובה ביותר לכל חומ"ג. (a) מודל שאומן עם חלון בגודל 3 ואורך ומרחב ווקטורי ממד 50, (b) מודל שאומן עם חלון בגודל 5 ומרחב ווקטורי ממד 50

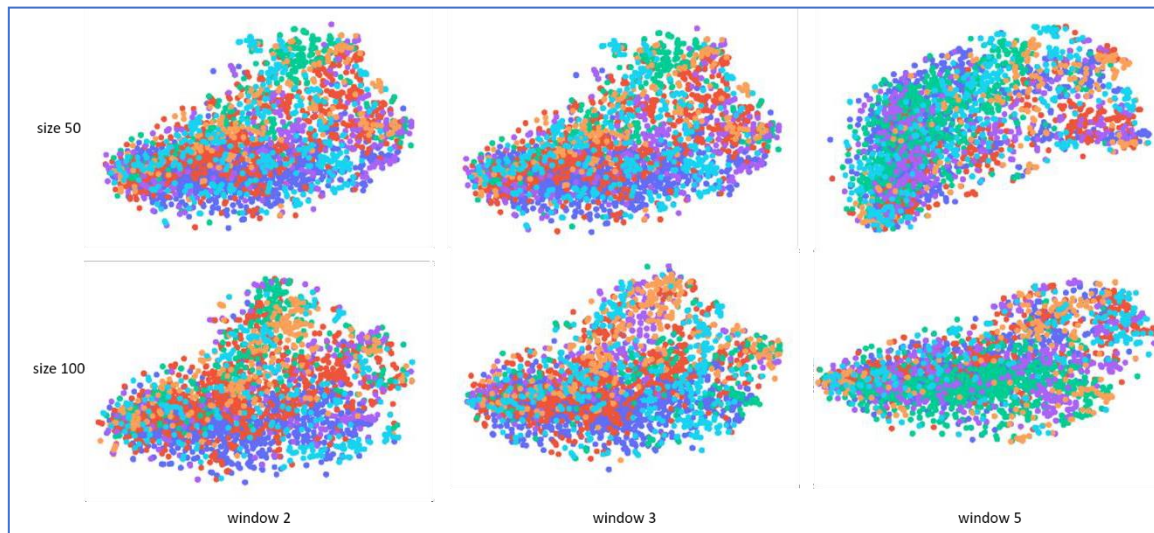


Figure 2: מודלים שאומנו לפיה פרמטרים שבגרף - בוויזואליזציה הזו טל"ב שנבחרו להצגה הן הנפוצות ביותר בקורפוס

## עבודה עתידית

אחד החסרונות בשיטת t-SNE הוא חוסר האינדיקטיביות של הויזואליזציה בין קלסטרים שונים. עבודות נוספות עשויות להתמקד בשיטת הויזואליזציה של הוקטורים שנוצרו מה-WE, כדי להסיק ממנה על היחסים בין פעולות הבישול עצמן (למשל, עד כמה קרובה פעולת הבישול לפעולת האידוי).

כאמור, בעבודה זו השתמשנו בהוראות בישול, מתוך הנחה כי טקסט זה יעיד על התאמתם רכיבים למתכון מבחינת התאמת פעולות הבישול שמתבצעות עליהם. ייתכן ועבודות שייעשו עם דאטא שונה יהיו מסוגלות להעיד גם על התאמת הרכיבים מבחינות אחרות כגון ערך תזונתי, מרקם, הגבלה חוץ-תזונתית (טבעונות/צמחונות, כשרות/חלאל).

- Ahn, Yong Yeol, Sebastian E Ahnert, James P Bagrow, and Albert-László Barabási. 2011. "Flavor Network and the Principles of Food Pairing." *Scientific Reports* 1. <https://doi.org/10.1038/srep00196>.
- Gibson, Megan. 2015. "Meet The Robot Chef That Can Prepare You Dinner | Time." TIME Tech Innovation. 2015. <https://time.com/3819525/robot-chef-moley-robotics/>.
- Goldberg, Yoav, and Omer Levy. 2014. "Word2vec Explained: Deriving Mikolov et Al.'s Negative-Sampling Word-Embedding Method." <https://code.google.com/p/word2vec/>.
- Habit. 2019. "Habit Essential." Habit. 2019. <https://habit.com/how-it-works/>.
- Heuer, Hendrik. 2016. "Text Comparison Using Word Vector Representations and Dimensionality Reduction." *PROC. OF THE 8th EUR. CONF. ON PYTHON IN SCIENCE*, 13. <http://arxiv.org/abs/1607.00534>.
- Huang, De An, Joseph J Lim, Li Fei-Fei, and Juan Carlos Niebles. 2017. "Unsupervised Visual-Linguistic Reference Resolution in Instructional Videos." In *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-Janua:1032–41. <https://doi.org/10.1109/CVPR.2017.116>.
- Jaan Altosaar. 2017. "Food2vec - Augmented Cooking with Machine Intelligence." 2017. <https://jaan.io/food2vec-augmented-cooking-machine-intelligence/>.
- Kiddon, Chloé, Ganesa Thandavam Ponnuraj, Luke Zettlemoyer, and Yejin Choi. 2015. "Mise En Place: Unsupervised Interpretation of Instructional Recipes." In *Conference Proceedings - EMNLP 2015: Conference on Empirical Methods in Natural Language Processing*, 982–92. Stroudsburg, PA, USA: Association for Computational Linguistics. <https://doi.org/10.18653/v1/D15-1114>.
- McCormick, Chris. 2017. "Word2Vec: Negative Sampling." <http://mccormickml.com/2017/01/11/word2vec-tutorial-part-2-negative-sampling/>.
- Mizrahi, Moran, Amos Golan, Ariel Bezaleli Mizrahi, Rotem Gruber, Alexander Zoonder Lachnise, and Amit Zoran. 2016. "Digital Gastronomy." In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology - UIST '16*, 541–52. New York, New York, USA: ACM Press. <https://doi.org/10.1145/2984511.2984528>.
- Porter, Martin. 1979. "Porter Stemming Algorithm." <https://tartarus.org/martin/PorterStemmer/>.
- Samsung. 2019. "Samsung Bot Chef." SAMSUNG. 2019. <https://news.samsung.com/global/get-a-glimpse-of-the-next-generation-innovations-on-display-at-samsungs-technology-showcase>.
- Wang, Wen, Lining Yao, Teng Zhang, Chin-Yi Cheng, Daniel Levine, and Hiroshi Ishii. 2017. "Transformative Appetite." In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, 6123–32. New York, New York, USA: ACM Press. <https://doi.org/10.1145/3025453.3026019>.
- Wattenberg, Martin, Fernanda Viégas, and Ian Johnson. 2017. "How to Use T-SNE Effectively."

*Distill* 1 (10). <https://doi.org/10.23915/distill.00002>.

Wikibooks. 2017. "Cookbook: Cooking Techniques." 2017.  
[https://en.wikibooks.org/wiki/Cookbook:Cooking\\_Techniques](https://en.wikibooks.org/wiki/Cookbook:Cooking_Techniques).