arXiv:2602.03092v1 [cs.CE] 3 Feb 2026

生成式人工智能创造美味、可持续且营养丰富的汉堡

Vahidullah Tac[1*]、Christopher Gardner[2] 和Ellen Kuhl[1]

[1]斯坦福大学机械工程系，美国斯坦福。[2]斯坦福大学医学院预防研究中心，美国斯坦福。

*通讯作者。电子邮箱：vtac@stanford.edu；共同作者：cgardner@stanford.edu；ekuhl@stanford.edu；

摘要

食物选择影响人类健康与地球健康；然而，要设计出美味、营养且可持续的食物仍颇具挑战。在此，我们展示生成式人工智能如何直接从大规模人工生成的食谱数据中学习人类味觉结构，并在结构化的设计空间中创造新颖食物。以汉堡为模型，生成式人工智能无需明确监督便重新发现了经典巨无霸，并生成了以美味度、可持续性或营养为优化目标的新型汉堡。在一项于餐厅环境中开展、共有101名参与者参加的盲法感官评价中，与巨无霸相比，其美味汉堡在总体喜好度、风味和质地方面得分相同或更高；其蘑菇汉堡的环境影响评分低了一个数量级以上；而其豆类汉堡的营养评分则接近翻倍。综上，这些结果确立了生成式人工智能作为一种定量框架，可用于学习人类口味，并以有原则的方式驾驭食品设计中的复杂权衡。

关键词： 适口性、可持续性、营养、生成式人工智能、机器学习

食物选择是人类所作最具深远影响的决策之一，对个人健康和地球健康都具有广泛影响 [1]。全球粮食系统对气候变化作出了重大贡献 [2], , 推动土地利用变化并造成生物多样性损失 [3], , 消耗淡水资源 [4], , 并污染陆地和水生生态系统 [5]。要让地球维持在人类安全运行空间之内，就必须迅速转向更加可持续的粮食系统 [6]。

同一系统也对人类健康产生主导影响 [7]。超过十亿人摄入的膳食无法满足基本营养需求 [8], , 还有更多人存在微量营养素缺乏 [9]。不良饮食模式是全球慢性非传染性疾病 [10],负担的重要成因，其中包括2型糖尿病 [11] 、心血管疾病 [12] 和某些癌症 [13]。因此，饮食将环境可持续性与人类健康通过共同的消费模式联系起来 [14]。

消费者接受度，而非可获得性，仍然是采用可持续且营养丰富食物的核心瓶颈 [15]。即使环境和营养益处显而易见，口味 [16], 、质地 [17], 和文化熟悉度 [18]方面的不足仍持续限制其接受程度。要设计出既满足环境和营养目标、又符合感官预期的食物，需要定量理解人类味觉[19]。然而，传统食品开发严重依赖手工制作专长和渐进式反复试错，这限制了对大型设计空间的系统性探索，并减缓了创新进程 [20]。

为弥补这一空白，我们利用生成式人工智能，将人类食物偏好建模为高维概率分布，并以汉堡作为模型系统。我们并非

通过规则编码风味或质地，而是让模型直接从观察到的食材组合和用量中学习统计规律。这种概率建模方式能够发现新颖食谱，并根据环境影响和营养质量等外部标准进行优化，同时确保设计与消费者息息相关。要实现这些能力，模型必须内化人类口味偏好，而不能仅停留在表面相关性上。

在此，我们检验如下假设：生成式人工智能能够学习人类味觉，并创造出口味与经典巨无霸®相同或更佳的汉堡。

巨无霸®位列全球食用最广泛的汉堡之列 [21], 它在100多个国家销售 [22],其价格被用作衡量各货币购买力平价的非正式指数 [23]。这种全球熟悉度使巨无霸®成为评估生成式模型能否捕捉人类集体口味偏好的严格基准。为检验这一前提，我们采用两项发现基准来评估该模型：对巨无霸的再发现®仅凭统计结构，以及创造出仍保持较高消费者接受度的新型汉堡。

除验证之外，我们还将环境可持续性与营养质量作为额外标准，用于从大量生成食谱的集成中进行筛选。我们利用土地利用、温室气体排放、富营养化潜势以及按水资源稀缺性加权的用水量 [24], 来量化环境影响，并采用既有的特征分析框架（包括健康饮食指数[25]）评估营养质量。通过广泛抽样并筛选同时满足适口性、可持续性与营养的食谱，我们表明，生成式人工智能能够识别出在不放弃文化熟悉度的前提下大幅降低环境影响并提升营养质量的汉堡。对100多名参与者开展的感官验证证实，该模型正确捕捉了人类味觉偏好，并创造出在感官吸引力上匹配或超越经典巨无霸®的汉堡食谱。

结果

生成式人工智能成功创作出汉堡食谱。
我们首先通过检验生成式人工智能模型能否复现2,216 份由人类设计的汉堡食谱的关键统计特性，同时生成新颖的食材组合，来验证该模型（图1）。该模型的架构将用于食材选择的多项式扩散模型与用于食材定量的基于评分的生成模型相结合，二者共同生成完整的汉堡食谱，每份食谱由 146种配料及其用量定义（图 1a）。将生成样本与训练数据进行比较后发现，多种边缘统计量和高阶统计量均高度吻合：模型尤其能够复现食材用量的分布（图 1b）以及食材流行度，即食材在各食谱中出现的概率（图 1c），这表明它既学会了食材出现的频率，也学会了食材的用量。模型还捕捉到了高阶结构，包括真实食谱中常见的食材配对之间浓烈的正相关和负相关（图 1d），并准确匹配了食谱长度的分布，食谱长度以每个汉堡所含食材的总数衡量（图1e）。在确认统计保真度后，我们生成了一百万份汉堡食谱，并绘制其适口性、环境和营养评分，以揭示生成的设计空间结构（图 1f）。适口性得分较高的食谱聚集在与受欢迎的传统食材组合相关的区域，这些食谱的营养评分较低，环境评分居中；而适口性得分较低的食谱则分布在与罕见或非传统组合相关的区域，这与人类的烹饪偏好一致。总之，这些结果表明，生成模型学会了真实人类设计的汉堡食谱的内在分布，并能系统地探索适口性、营养与环境影响之间的权衡。

生成式人工智能重新发现经典汉堡并创造新颖设计。
接下来，我们评估生成式人工智能模型能否既重新发现经典汉堡食谱，又生成新颖且诱人的替代方案（图 2）。我们使用实质差异分数来量化生成样本与参考配方之间的相似度，其中实质差异分数 $= 0$ 表示食材和用量相匹配，而实质差异分数 > 0 则衡量新颖性不断增加的程度（图2a）。从随机样本中，模型成功重新发现了经典巨无霸®，其食材和重量均准确无误，尽管巨无霸®从未出现在初始训练数据中（图2a）。在十次独立随机化中，重新发现巨无霸®平均需要 730 万个样本，这表明在学习到的分布下，食谱的精确复现是低概率事件（图2e）。除重新发现经典汉堡外，模型还生成了新颖程度各异的汉堡，两个具有代表性的食谱分别是美味汉堡1（实质差异分数 $= 3$）和美味汉堡2（实质差异分数 $= 6$）。它们的食材构成逐渐呈现出更显著的差异，同时保留了熟悉的汉堡结构（图。

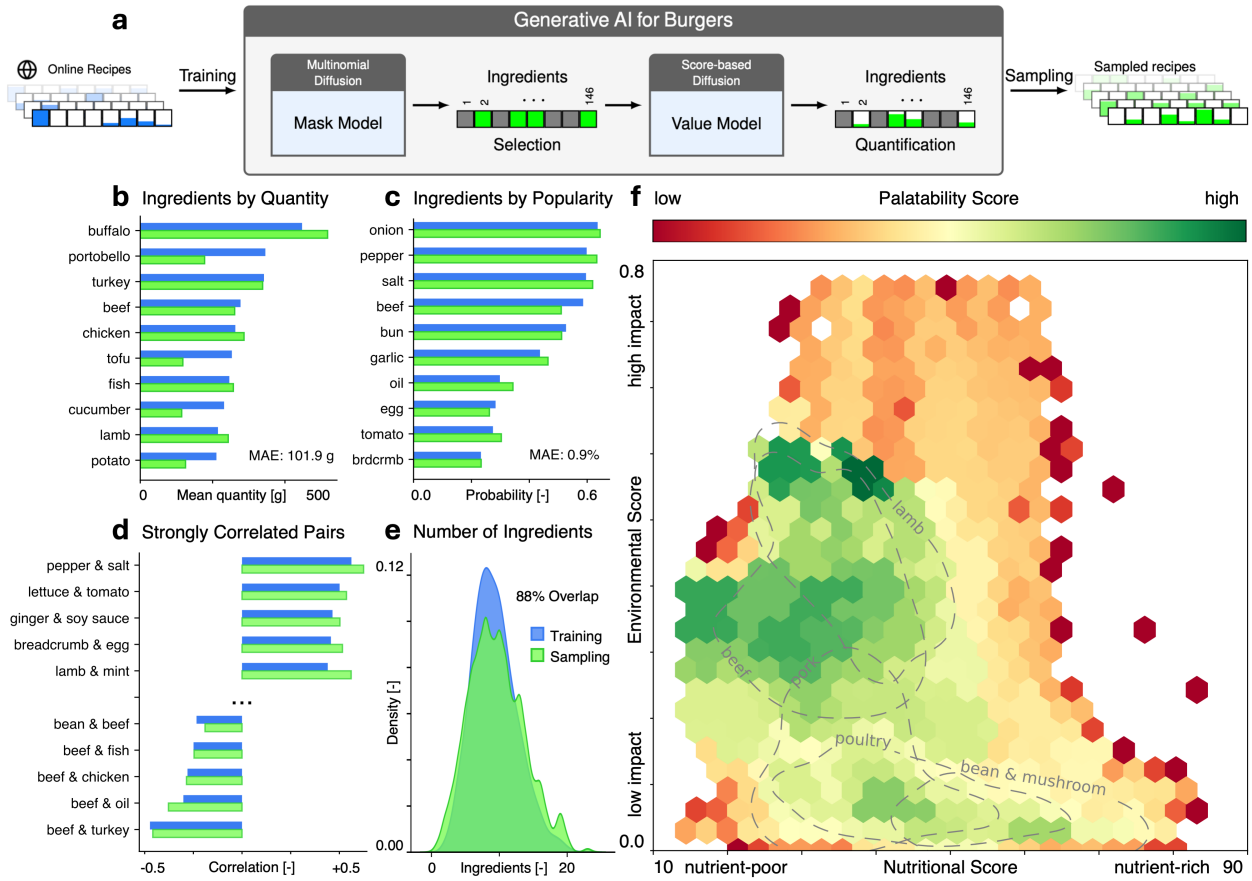

图1 汉堡的生成式人工智能。 高层模型架构包括用于食材选择的多项式扩散模型和用于食材定量的基于分数的扩散模型（a）；按数量（b）和受欢迎程度（c）比较训练数据与采样数据中的食材；强相关配对（d）；以及食材总数（e）；100万个样本中适口性得分、环境得分和营养评分的分布，其中高适口性（绿色）表示受欢迎的食材组合，低适口性（红色）表示罕见的食材组合（f）。

2c、d）。感官评价显示，这两款美味汉堡的消费者评分与经典巨无霸®相当，有时甚至更高（图2f、g）：美味汉堡1 的风味评分显著高于巨无霸®（5.8 ± 1.3 对比 5.4 ± 1.5）；美味汉堡2 的总体喜好度（5.7 ± 1.2 对比 5.3 ± 1.5）和风味评分（5.8 ± 1.3 对比 5.4 ± 1.5）也显著更高，而两款汉堡的质地评分与巨无霸®相比均无显著差异（$n = 101$，$p < 0.05$）。与巨无霸®相比，参与者更常将美味汉堡1描述为肉味浓郁（67% 对比 42%）、湿润（60% 对比 32%）和油腻（40% 对比 12%）；更常将美味汉堡2 描述为肉味浓郁（66% 对比 42%）和带有烟熏味（47% 对比 4%）（$n = 101$；配对比较，$p < 0.05$）。总之，这些示例验证了生成式人工智能模型的能力：它能够内化经典汉堡食谱，并生成新颖、美味且可口的设计。

生成式人工智能创造出可持续汉堡。我们接下来评估生成式人工智能模型能否识别并生成环境影响较低、同时保持消费者接受度的汉堡食谱（图 3）。我们使用环境影响评分来量化可持续性，该评分汇总了土地利用、富营养化潜势、按水资源稀缺性加权的用水量和温室气体排放所产生的成分级贡献（图 3a）。对训练数据的分析显示，不同主要蛋白质来源的食谱之间，环境影响评分存在显著差异；羊肉和牛肉配方的环境影响系统性高于禽肉和蘑菇配方（图 3b）。从模型中抽取一百万份食谱，可以识别出可持续汉堡候选方案，其中两个具有代表性的例子是可持续汉堡1和可持续汉堡2；二者在成分组成和主要蛋白质来源方面存在差异（图 3c,d）。可持续汉堡1采用蘑菇基配方，环境影响评分为0.06，比

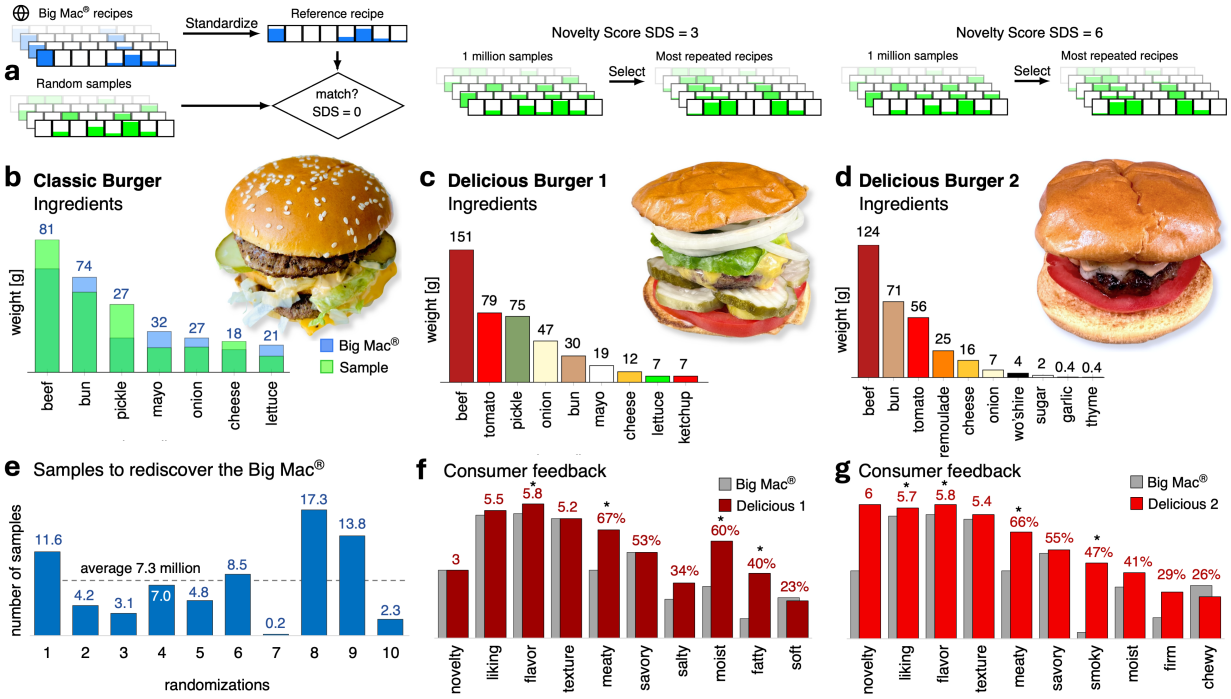

图2 重现经典汉堡，发现美味汉堡。生成式人工智能从随机生成的人工智能样本中重现经典汉堡并发现新颖汉堡：实质差异分数为零（SDS $=0$）表示经典食谱与样本相匹配，实质差异分数大于零（SDS >0）则表示样本的新颖度分数（a）；重现的巨无霸®食谱，作为代表性经典汉堡（b）；发现的美味汉堡1食谱（SDS >3）和美味汉堡2食谱（SDS >6）（c、d）；在$n=10$次随机化中重现巨无霸®所需的样本数量（e）；消费者对美味汉堡1和美味汉堡2的反馈，并与巨无霸®进行比较（f、g）；$*$表示具有统计显著性（$p<0.05$）。

巨无霸低一个数量级®（巨无霸评分为0.93）；而可持续汉堡2采用蘑菇与牛肉混合配方，评分为1.02，与巨无霸相当®（图3e）。消费者反馈表明，可持续汉堡1在总体喜好度、风味和质地方面的评分略低于巨无霸®，而可持续汉堡2在这些类别中的表现与巨无霸®相当。感官评估表明，两种可持续汉堡在消费者评分方面均以系统性的方式不同于经典巨无霸®（图 3f、g）：可持续汉堡1在总体喜好度（4.8 ± 1.8 对比 5.3 ± 1.5）、风味（5.0 ± 1.9 对比 5.4 ± 1.5）和质地（4.5 ± 1.9 对比 5.2 ± 1.5）方面的评分均显著低于巨无霸®；而可持续汉堡2在这些属性上的评分与巨无霸®均无显著差异（$n=101$；$p<0.05$）。与巨无霸®相比，参与者更常将可持续汉堡1描述为带有泥土味（63% 对比 2%）、浓烈（37% 对比 14%）、湿润（53% 对比 32%）和柔软（50% 对比 25%）；而将可持续汉堡2描述为带有烟熏味（55% 对比 4%）、湿润（49% 对比 32%）和油腻（30% 对比 12%）（$n=101$；配对比较，所有报告属性均为 $p<0.05$）。这些结果共同验证了生成式人工智能模型能够成功权衡可持续性与适口性，并发现环境影响显著降低且不影响口味的汉堡。

生成式人工智能创造出营养且个性化的汉堡。接下来，我们评估生成式人工智能模型能否识别出针对营养质量进行优化的汉堡食谱，并根据个人膳食需求对其进行调整（图4）。我们使用健康饮食指数来量化营养状况；该指数汇总了有益食物类别、脂肪酸组成和需限制的《营养素》的贡献（图4a）。对训练数据的分析显示，使用不同主要蛋白质来源的食谱，其健康饮食指数存在显著差异；以豆类和蘑菇为基础的食谱，其营养评分系统性地高于以牛肉和羊肉为基础的食谱（图4b）。对一百万份食谱进行采样后，我们得以识别出具有代表性的营养汉堡，其营养评分较高；与其他人工智能生成的汉堡相比，它位于营养—环境设计空间中更有利的区域（图4c、d）。该营养汉堡采用豆类配方，

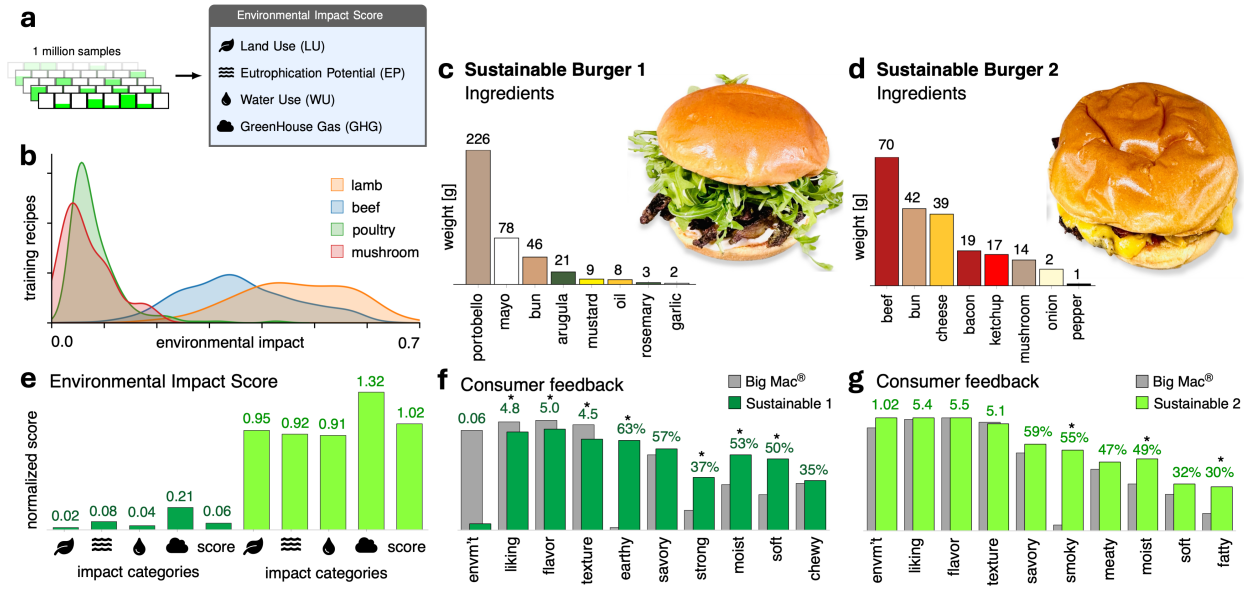

图 3 生成可持续汉堡。生成式人工智能根据随机生成的人工智能样本生成可持续汉堡，其中环境影响评分表示食谱中所有食材对土地利用、富营养化潜势、按水资源稀缺性加权的用水量和温室气体排放的综合影响 (a)；所有含羊肉、牛肉、禽肉和蘑菇的训练食谱的环境影响评分分布 (b)；可持续汉堡1和可持续汉堡2的生成食谱 (c,d)；可持续汉堡1和可持续汉堡2的环境影响评分 (e)；与巨无霸相比，消费者对可持续汉堡1和可持续汉堡2的反馈Ⓡ(f,g)； $*$ 表示具有统计显著性（$p < 0.05$）。

其健康饮食指数为63.12，几乎是巨无霸Ⓡ（33.71）的两倍，同时将其环境影响评分降低了六倍（图4c）。与巨无霸Ⓡ相比，该营养汉堡在健康饮食指数的多个组成部分上与膳食指南的契合度更高，包括增加蔬菜、全谷物和植物蛋白的贡献，同时减少精制谷物、钠和饱和脂肪（图4e）。消费者反馈显示，与巨无霸Ⓡ相比，营养汉堡的愉悦度评分明显降低（图4f）：营养汉堡在总体喜好度（3.8 ± 1.7 对 5.3 ± 1.5）、风味（4.0 ± 1.8 对 5.4 ± 1.5）和口感（3.7 ± 1.8 对 5.2 ± 1.5）方面的评分均显著低于巨无霸Ⓡ（$n = 101$； $p < 0.05$）。与巨无霸Ⓡ相比，参与者更常将营养汉堡描述为带有泥土味（55% 对 2%）、清淡无味（43% 对 21%）、干燥（51% 对 18%）、柔软（50% 对 25%）和颗粒感（42% 对 7%），而较少描述为鲜美（27% 对 53%）（$n = 101$；配对比较，所有报告属性均为 $p < 0.05$）。除了群体层面优化之外，人工智能模型还可以生成根据个人营养需求定制的个性化汉堡食谱。我们通过为一名活动量很高的15岁男性和一名活动量中等的70岁女性制作个性化食谱，展示了这一功能；两者在成分组成方面存在差异并根据年龄和活动水平调整用量，以满足相应的膳食需求（图 4g）。最后，对全部六种汉堡进行直接比较，揭示了食材数量、新颖性、环境影响和营养之间的系统性权衡，并将人工智能生成的汉堡置于这一多目标设计空间中的不同区域（图 4h）。在这一设计空间中，美味汉堡1和美味汉堡2 的总体喜好度、风味和质地评分最高，而可持续汉堡1 和营养汉堡在营养和环境方面表现最佳（图 4i）。总之，这些结果证实，生成式人工智能模型能够在维持感官接受度的同时，在人群和个体层面优化营养质量。

讨论

学习食品设计的结构。这项研究的一个核心结果是，生成式人工智能可以直接从大规模、由人类生成的食谱数据中学习食品设计的潜在结构 [26]。该模型并非只是再现表层统计特征，而是捕捉了定义汉堡这一烹饪类别的高阶规律，包括食材共现、典型食谱长度和特征性的数量分布。

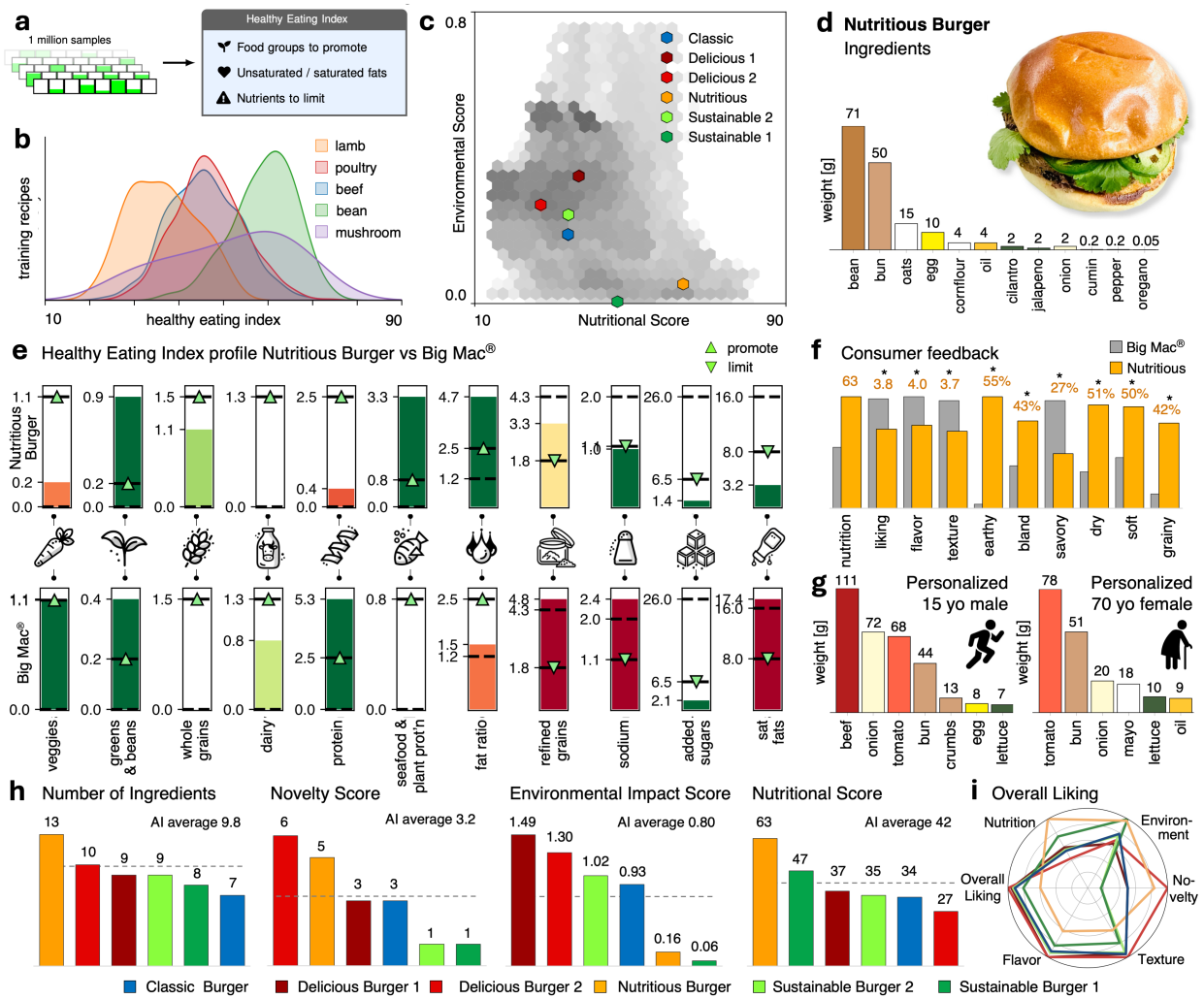

图 4 生成营养汉堡和个性化汉堡。生成式人工智能根据随机生成的人工智能样本生成营养汉堡，其中，健康饮食指数定义了食谱中所有食材对需促进的食物类别、不饱和脂肪和饱和脂肪，以及需限制的营养素的综合影响（a）；含羊肉、禽肉、牛肉、豆类和蘑菇的所有训练食谱的健康饮食指数分布（b）；六种汉堡的营养与环境评分（c）；生成的营养汉堡食谱（d）；营养汉堡与巨无霸的健康饮食指数概况对比®（e）；营养汉堡与巨无霸的消费者反馈对比®（f）；为15岁、身高180厘米、体重80千克的活跃男性，以及70岁、身高170厘米、体重70千克的中等活跃女性生成的个性化食谱（g）；按食材数量、新颖性、环境影响和营养评分直接比较六种汉堡（h）；总体比较（i）；$*$表示具有统计显著性（$p < 0.05$）。

这一视角呼应了数据驱动烹饪科学领域更广泛的发现，例如风味网络方法能够揭示全球烹饪体系数万份食谱中食材共现的基本模式 [27]。模型将食谱的统计语法内化的能力，反映了人类口味的高维非线性本质；人类口味源于食材之间复杂的相互作用，而非简单的加性规则 [28]。重要的是，在没有明确监督的情况下，再发现诸如 BigMac®这样的经典基准，表明具有文化主导地位的食物位于高概率区域的所学习到的设计空间，并依据广泛认可的烹饪参照点验证模型。这些发现共同支持这样一种观点：食谱承载着数十年来积累的集体人类口味知识，而生成模型能够提取这些知识，形成一个可解释且可导航的设计流形。这重新界定了食品配方设计：它不再是依赖工匠技艺、反复试错的实践，而成为一门数据驱动的设计科学；同时，食谱也成为连接烹饪传统与人工智能的一种自然且以人为本的界面。

探索新颖性，而不牺牲适口性。除了学习现有的烹饪结构，生成模型还能够在保留感官吸引力的同时，对新颖的汉堡设计进行可控探索。通过用实质差异分数明确量化新颖性，我们表明，偏离经典食谱并不必然导致消费者接受度下降。相反，该模型能够识别设计空间中兼具新颖性与适口性的区域：它发现了一些汉堡，其食材组成和比例存在显著差异，但在喜好度、风味和质地方面的评分仍与广泛食用的参照汉堡相当。这种表现反映了人类味觉的高维性与非线性特征，人类味觉源于食材之间的复杂相互作用，仅凭直觉或简单的经验法则无法可靠地加以探索[19]。这也与系统性分析的结论一致：新颖食物的消费者接受度既取决于口味和熟悉度等内在感官属性，也取决于食物新恐惧和食物体验等心理因素 [16]。经典汉堡的再发现，以及成功的新颖变体，表明文化熟悉度在烹饪设计空间中构成了广泛可接受区域的基石，使创新得以实现，同时不会疏远消费者的期待 [29]。在此背景下，生成式人工智能为探索熟悉度与新颖性之间的非线性权衡提供了系统化框架，并将烹饪创新重新界定为对结构化、数据驱动的设计流形进行导航，而非反复试错式实验。
</output>

将环境可持续性作为系统级设计目标。本研究的一项关键贡献在于证明，生成式人工智能能够在立足于与消费者相关的设计的同时，大幅降低汉堡的环境影响。畜牧业生产是温室气体排放、土地利用和生物多样性丧失的主要驱动因素 [30]；然而，消费者对低环境影响替代品的采用仍受到口味、质地和熟悉度的限制 [31]。我们的结果表明，可持续性并非源于孤立的食材替换，而是来自对整份食谱的协同调整。由此产生的设计通过对食材进行系统级重新平衡，大幅降低了环境影响。这一发现与更广泛的证据相符：饮食可持续性有赖于综合性的食物系统重设计，而非替换单一食材 [32]。值得注意的是，纯蘑菇汉堡的环境影响得分比传统参照汉堡低一个数量级以上，而蘑菇混合汉堡则保留了相近的环境影响和感官表现。这凸显了环境绩效取决于食材之间的相互作用，而非任何单一成分；也说明生成模型非常适合处理这类非线性的多目标权衡。通过将生命周期考量直接融入生成过程，该模型能够找出既可持续、又仍然让人熟悉为汉堡的设计，并解决替代蛋白的采用中的一个核心瓶颈——接受度，而非供应不足——这一挑战在有关消费者对植物基和混合型肉类替代品反应的研究中屡被强调 [33]。总之，这些发现弥合了环境指标与面向消费者的设计之间的鸿沟，并使生成式人工智能成为实现食品可持续性系统级优化的实用工具。

将营养与个性化作为明确的设计标准。 本研究的另一项贡献是证明，营养质量可以作为明确且可量化的设计目标，而非食品配方设计的次要结果。通过依据健康饮食指数 [25], 优化生成的食谱，该模型识别出营养评分远超广泛消费基准的汉堡，其营养评分几乎是传统参考汉堡的两倍，同时将环境影响降低了六倍。这一结果凸显，营养与可持续性的改善并非必然相互排斥，而是可以通过对成分组成和用量进行协调的系统级调整来实现 [32]。重要的是，人工智能生成的营养汉堡在达到这一表现的同时，并未增加配方的复杂程度，因而直接回应了人们对超加工食品以及可能损害消费者信任的冗长配料表日益增长的担忧。超越群体层面优化，人工智能模型还展示了将膳食参考摄入量转化为面向消费者的具体食品设计的能力：根据年龄、性别和活动水平个性化定制食谱，这也建立在此前依据葡萄糖水平实现营养个性化的举措之上[34]。这些发现共同表明，生成式人工智能能够切实连接营养科学与日常饮食，并在统一的设计框架内实现对食物的客观比较、优化和个性化。

局限性与范围。 本研究存在若干局限性，界定了所报告结果的适用范围。首先，生成模型从现有的人工设计食谱中学习，因此也继承了源数据中的文化、地域和时间偏见，

其次，食谱表示仅包含食材身份和用量，并未明确考虑影响质地和风味的加工步骤、烹饪方法或物理转化。第三，环境评分和营养评分依赖汇总数据库和全球平均值，因此无法反映特定生产实践或供应链所带来的差异。第四，感官验证仅针对有限数量的生成汉堡和参与者；要确定其在不同人群和情境中的普遍适用性，还需开展更广泛的研究。尽管存在这些限制，该模型仍为多目标食品设计提供了灵活且可扩展的框架。

更广泛的影响与未来展望。这项研究展示了生成式人工智能如何推动食品配方设计从手工反复试错转向定量数据驱动的设计科学。人工智能模型直接从大规模食谱数据中学习，进而成为汉堡设计的领域专用世界模型，捕捉塑造出可行且受欢迎食物的统计规律与权衡。这使得我们能够系统地探索美味度、营养和可持续性，同时立足于文化熟悉度；在食品创新中，真正的核心挑战是如何促成采用，而非追求新颖性。更广泛而言，食品为生成式人工智能提供了一个以人为本的独特领域，在这里，模型能够将优化目标与感官反馈、健康及环境限制相协调。若进一步将加工、烹饪和供应链动态纳入此类方法，最终有望实现端到端食品设计，并使生成式人工智能成为协作工具，推动人类创造力、工程学与地球健康交汇融合。

结论

本研究表明，生成式人工智能能够将食品配方设计重新构想为一种以人类偏好和可测量约束为基础、由数据驱动的定量设计流程。通过直接从大规模食谱数据中学习，该模型掌握了汉堡设计的结构，并能系统地探索口味、营养和可持续性之间的关系。消费者反馈证实，生成模型设计的产品符合人类口味偏好，同时相较于传统基准，在营养质量和环境影响方面均有显著改善。除汉堡之外，这一方法还指向人工智能驱动的食品设计新范式，将烹饪创意与人类健康和地球健康融为一体。

方法

模型架构、训练与验证

食品食谱是离散—连续混合对象，需要 *ingredient selection* 和 *ingredient quantification*。在这里，我们仅用食材及其对应重量来表示每种汉堡食谱。因此，食谱 $\mathbf{x}_0 = \{\mathbf{m}, \mathbf{w}\}$，由二元食材掩码 $\mathbf{m} \in \{0,1\}^K$ 和食材重量 $\mathbf{w} \in \mathbb{R}^K$ 组成；掩码指示每种食材是否存在。我们采用一种基于扩散的两阶段框架，将食材选择与食材定量分离：具体而言，我们结合 *multinomial diffusion model* [35] 来生成食材掩码，并结合 *score-based generative model* [36] 在给定掩码的条件下生成食材重量（图1a）。

基于扩散的食谱生成。扩散模型定义了一个前向随机过程 $q(\mathbf{x}_t|\mathbf{x}_{t-1})$，逐步向数据样本 $\mathbf{x}_0$ 中添加噪声，以生成潜变量 $\mathbf{x}_t$，使原始数据在 $t = 1,\ldots,T$ 个时间步 [37]中日益难以辨识。可学习的部分是反向过程 $p(\mathbf{x}_{t-1}|\mathbf{x}_t)$，它逐步去除噪声，并支持从无结构先验中生成数据。我们通过最大化数据似然的变分下界来训练扩散模型，这对应于证据下界，

$$\log P(\mathbf{x}_0) \geq \mathbb{E}_{\mathbf{x}_1,\cdots,\mathbf{x}_T \sim q}\left[\log p(\mathbf{x}_T) + \sum_{t=1}^{T} \log \frac{p(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q(\mathbf{x}_t|\mathbf{x}_{t-1})}\right].$$

这里，$P(\mathbf{x}_0)$表示生成模型下观测到的汉堡食谱的边缘似然，其通过对所有潜在扩散变量进行积分获得，其中$\mathbf{x}_0$是由人设计的汉堡食谱，$\mathbf{x}_t$是其逐步加噪后的表示，q 是固定的前向过程，p 是学习得到的反向去噪过程，而 $\mathbb{E}$ 是期望算子，表示对从前向扩散过程中抽取的噪声实现取平均。

通过多项式扩散进行食材选择。我们使用多项式扩散过程对食材选择进行建模，其中将食材是否存在视为分类变量。我们将前向过程定义为

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{C}(\mathbf{x}_t \,|\, (1-\beta_t)\mathbf{x}_{t-1} + \beta_t/K),$$

其中 $\mathcal{C}$ 表示参数为竖线后所列内容的分类分布 $|$，β_t 控制时间步 t的噪声水平，K 是类别数。在我们的应用中，食材选择为*binary* $K = 2$，也就是说，一种食材要么存在，要么不存在。上述方程可简化为参数为 $(1-\beta_t)\mathbf{x}_{t-1} + \beta_t(1-\mathbf{x}_{t-1})$ 的伯努利分布，它以概率 β_t 将食材纳入状态从存在切换为不存在，或反之；并以概率 $(1-\beta_t)$ 保持不变。随着 t 增大，食材掩码会逐渐随机化，而训练得到的反向过程则会重建统计上合理的食材组合，其中自然体现了汉堡食谱中常共同出现的食材之间的依赖关系。

基于分数的扩散实现食材定量。在给定食材掩码的条件下，我们使用通过随机微分方程构建的基于分数的生成模型来生成食材用量。正向 *noising process* 为

$$\mathrm{d}\mathbf{x}_t = f(\mathbf{x}_t, t)\,\mathrm{d}t + g(t)\,\mathrm{d}B_t,$$

而反向时间 *denoising process* 为

$$\mathrm{d}\mathbf{x}_t = \left[g^2(t)\nabla_{\mathbf{x}} \log p_t(\mathbf{x}) - f(\mathbf{x}_t, t)\right]\mathrm{d}t + g(t)\,\mathrm{d}\tilde{B}_t,$$

其中 B_t 是一个 K维布朗运动，$\tilde{B}_t$是其时间反演，而 f 和 g 分别定义漂移系数和扩散系数 [36, 38]。这里，我们不直接学习概率密度 $p_t(\mathbf{x})$，而是使用神经网络近似分数函数 $\nabla_{\mathbf{x}} \log p_t(\mathbf{x})$，从而能够高效采样食材重量，使其符合人工设计的汉堡食谱中观察到的分布。

数据集与训练。我们使用经过整理的汉堡数据集训练模型。该数据集源自 Food.com 上一个包含五十多万份人工设计食谱的开源集合 [39, 40]。我们筛选出所有汉堡食谱，从自由文本中提取食材、用量和单位，并对数据进行标准化，转换为结构化表示。最终的数据集包含 2,216 份汉堡食谱，由 146种配料组成。补充文件1提供了预处理步骤的更多详情。

模型验证与统计保真度。训练后的模型能够准确再现训练数据的一阶和高阶统计特性：该 *ingredient selection model*模型对随机汉堡食谱中每种食材出现的边际概率估计，最大绝对误差低于 1%（图 1b）。该 *ingredient quantification model* 模型预测此前未见食谱中食材用量的平均绝对误差为 101.9 克（图1c），尽管这一问题具有外推性且高度随机。除了边缘统计量，模型还捕捉到了高阶结构，包括 *pairwise ingredient correlations* 1（图d）和*recipe number of ingredients per* 1（图 b、e）。训练过程中并未明确约束这些特性，而是从学得的生成过程中自然涌现。补充文件1提供了更多验证结果和基准。

实质差异分数，用于量化相似性。量化食谱之间的接近程度并将相似的食谱分组，可用于多种场景，例如量化人工智能生成的食谱的新颖性。为此，我们定义了两个食谱 r_1 与r_2之间的半离散*substantial difference score* ，

$$\mathsf{SDS}(r_1, r_2) = \sum_{i=1}^{n_{\mathrm{ing}}} d_i(r_1, r_2),$$

作为数据库中所有 $n_{\mathrm{ing}} = 146$食材的二元距离 d_i 之和，其中

$$d_i(r_1, r_2) = \begin{cases} 1 & \text{if } r_{1i} + r_{2i} \neq 0 \text{ and } r_{1i} \cdot r_{2i} = 0 \\ 1 & \text{if } \max(r_{1i}, r_{2i}) / \min(r_{1i}, r_{2i}) \geq 2 \\ 0 & \text{otherwise.} \end{cases}$$

在 *rediscovery*期间，我们使用为零的实质差异分数SDS $= 0$，来量化人工智能生成的食谱与人工设计的食谱之间的匹配。在*discovery*期间，我们使用大于零的值SDS > 0，来量化人工智能生成的食谱相较于训练集中人工设计的食谱的新颖性。

流行度分数，用于量化适口性。适口性是指使食物符合人类味觉喜好的特质，例如风味、香气和质地。尽管不同个体的人类味觉存在显著差异，我们仍可通过测量群体中的 *popularity score* 来量化食品的整体适口性。食物制备模式也体现了这一点，例如，某些食材组合比其他组合更受欢迎（图1d）。在此，我们提出以受欢迎程度作为适口性的代理指标。我们的人工智能模型学习人类味觉的概率分布，并为受欢迎的食谱、模式和组合赋予更高的概率。在食谱层面，重复出现得越频繁，实际上就意味着该食谱越适口，并与更高的流行度分数相关联。

汉堡的生成式人工智能

经典汉堡。作为概念验证，我们使用生成模型来 *rediscover* 巨无霸®，它是全球消费最广泛的汉堡之一 [21]，，在100多个国家供应 [22]。这种全球普及反映了它在不同人群中具有很高的适口性，也使巨无霸®成为一项严格的基准，用于评估模型是否捕捉到了广泛共享的偏好。由于官方配方属于专有信息，我们将四种独立的开源复刻配方综合起来，近似得到统一的参考配方 [41–44]（图1b）。随后，我们在模型随机生成的样本中搜索这一参考配方（图1e）。我们将再发现定义为与参考配方相比实质差异分数为零的样本，SDS $= 0$。训练数据集中不包含参考巨无霸® 配方。

美味汉堡。接下来，我们使用人工智能来 *discover* 制作美味汉堡，并预先设定*novelty score*。具体而言，我们采用实质差异分数来量化人工智能生成样本的新颖性，并将其与训练集中的人工设计食谱进行比较。对于美味汉堡1，我们抽取100万个样本，筛选出所有满足以下条件的样本：SDS ≥ 3，并从筛选结果中选出重复次数最多的样本，作为适口性最高、流行度分数最高的食谱（图2c，f）。对于美味汉堡2，我们执行相同步骤，但此时采用SDS ≥ 6（图2d，g）。补充文件4提供了有关其加工方式的详细信息。

可持续汉堡。我们使用生命周期评价数据来评估环境可持续性。这些数据依据全球生产者调查 [45]，估算农产品在生产和分销链中的总体环境影响。我们从跨越 $n = 570$ 项研究的协调环境数据库中获取成分层面数据 [30]。由于

该数据库未包含蘑菇，因此我们补充了来自美国农业部的土地利用数据 [46] ，以及欧洲蘑菇生产的淡水富营养化潜势、按水资源稀缺性加权的用水量和温室气体排放数据 [47]。我们通过一个单一指标来量化可持续性 *environmental impact score* 该指标对各成分的土地利用、水生富营养化潜势、按水资源稀缺性加权的用水量和温室气体排放进行归一化后取平均，并按各成分的用量加权 [24]。对于可持续汉堡 1，即环境影响评分为 0.06 的纯蘑菇汉堡，我们抽取 100 万个样本，按环境影响评分排序，并选出总体上出现次数最多的食谱（图 3c，f）。对于可持续汉堡 2，即环境影响评分为 1.02 的牛肉蘑菇混合汉堡，我们执行相同的步骤，但这次选出同时含有牛肉和蘑菇且出现次数最多的食谱（图3d，g）。补充文件 1 报告了归一化因子和类别划分。

营养汉堡。我们使用成熟的营养画像模型量化营养质量。这些模型将食物和营养素组成与膳食指南进行比较，包括健康饮食指数 [25], 营养评分 [48], 和健康星级评级 [49]。此处我们采用*healthy eating index* 美国农业部制定的模型，评估食物与《美国人膳食指南》的一致程度；该模型侧重于食物组充足性，而非单一营养素 [50, 51]。我们从美国农业部食物模式等价数据库获取食物组当量 [52] ，并从美国农业部FoodData Central数据库获取营养素组成数据 [53]，再汇总各成分对食物组和营养素的贡献，计算健康饮食指数，并将数值标准化为每份 500 千卡。对于营养汉堡，这是一款以豆类为基础、健康饮食指数为 63.12 的配方，我们抽取 100 万个样本，按健康饮食指数排序（图 4e），并选取前 5% 中出现频率最高的食谱（图 4d、f）。补充文件 1 列出了替代营养评分对应的食谱，补充文件 4 提供了加工方式的详细信息。

个性化汉堡。 我们采用个性化营养素概况模型，根据年龄、性别、身体组成和身体活动水平调整食谱，以应对个体之间营养需求的差异 [54]。我们计算一个 *personalized nutrition score* 基于个人特征（包括年龄、性别、体重、身高和身体活动水平）的0–100分数。我们根据膳食参考摄入量和可接受的常量营养素分布范围 [55, 56],以及世界卫生组织关于钠、游离糖和饱和脂肪最高摄入量的指南 [57–59]，得出营养素特定目标范围。我们将这些目标汇总为每种汉堡的单一个性化营养评分。利用这一框架，我们为两种具有代表性的人口统计特征生成个性化汉堡食谱：一名15岁、身高180厘米、体重80千克且活动量大的男性，以及一名70岁、身高170厘米、体重70千克且活动量适中的女性（图 4g）。补充文件 1报告了其他个性化食谱。

汉堡验证

汉堡制备。 我们的人工智能生成的食谱仅列出食材和用量，不包括制作实际汉堡所需的加工或烹饪步骤。我们邀请一名行政主厨评估食材清单，并确定每份食谱适用的加工、烹饪和组装流程。我们将最终的制备方案交给一个独立的厨师团队，由他们制作我们的五种人工智能生成的汉堡——美味汉堡1、美味汉堡2、可持续汉堡1、可持续汉堡2和营养汉堡——并购买原版巨无霸®进行比较，以开展感官调查。补充文件 4包括所有食谱以及厨师的操作说明。

感官调查。 我们按照斯坦福大学机构审查委员会的指南，在加利福尼亚州旧金山一家营业中的餐厅，对来自一般人群的 $n = 101$ 名自愿参与者进行盲法感官评价。每位参与者使用7点李克特量表，对全部六种汉堡的 *overall liking*、*flavor*和 *texture* [60]进行评价，并回答有关 *12 flavor-* 和 *15 texture*相关属性的多选题。所有调查答复均经过完全匿名处理。补充文件 2提供了完整的问卷。

人口统计特征。 在 $n = 101$ 名参与者中，47.5%为男性，47.5%为女性，3%为非二元性别，2%不愿透露；22%为18-25岁，26%为26-35岁，19%为36-45岁，18%为46-55岁，16%超过55岁；65%为杂食者，35%为弹性素食者；最高教育程度为高中学历的占4%、大学学历的占24%、学士学位的占50%、硕士学位的占11%、博士学位或更高学历的占8%，职业学校学历的占3%；4%每天吃汉堡，20%每周吃2-3次，31%每周吃一次，27%每月吃2-3次，16%每1-2个月吃一次，3%每年吃4-5次。

统计功效与样本量。 我们选择了参与者人数 $n = 101$，以兼顾可行性与统计功效。对于李克特量表评分，该样本量足以通过双侧 Welch t检验检测小到中等的效应量。对于二元感官属性，在显著性水平为 $p < 0.05$ 时，该样本量可提供 $> 80\%$ 的统计功效，通过配对比较检测汉堡之间超过 20% 的差异。我们未进行事前功效计算；不过，样本量 $n = 101$ 与食品消费者调查中常用的样本量相当或更大。

统计分析。我们以7点李克特量表报告总体喜好度、风味和质地的感官评分，并以均值 ± 标准差表示。我们使用双侧 Welch t检验，将人工智能生成的汉堡与 Big Mac 进行比较®。我们以百分比值报告二元风味和质地属性，并进行配对比较，使用双侧二项检验评估统计显著性。由于所有检验均为预先计划且由假设驱动，因此我们不对多重比较进行校正。我们将统计显著性报告为 $p < 0.05$（图 2f,g，3f,g，4f）。

补充信息。本手稿附有以下补充文件：补充文件 1 介绍了数据准备、模型架构、训练和推理的详情，说明如何生成美味、可持续、营养丰富且个性化的汉堡，并提供了其他汉堡食谱。补充文件 2 概述了人工智能生成的汉堡食谱。补充文件 3 提供了行政主厨制定的加工方式和烹饪说明。补充文件 4 提供了调查参与者的人口统计特征、感官反馈调查问卷以及其他调查结果。

致谢。我们感谢行政主厨 Justin Schneider 凭借其烹饪专长编写食谱制作说明，感谢 Food System Innovations 的 NECTAR 团队 Caroline Cotto 与我们进行了富有启发性的讨论，并感谢 Palate Insights 的 Alice Wistar 和 Alex Weissman 开展消费者调查。我们感谢斯坦福 Marlowe 计算平台提供高性能计算资源。本研究得到 Schmidt Science Fellowship（与罗德信托基金合作）对 Vahidullah Tac 的资助，以及斯坦福大学多尔可持续发展学院加速器、斯坦福 Bio-X Snack 资助计划、贝索斯地球基金、美国国家科学基金会 CMMI 项目资助 2320933 和欧洲研究委员会高级资助 101141626（资助 Ellen Kuhl）的支持。

参考文献

[1] 社论。膳食、健康与环境。《自然·食品》**5**, 717（2024年）。[2] Vermeulen, S. J., Campbell, B. M. & Ingram, J. S. 气候变化与粮食系统。《环境与资源年度评论》**37**, 195–222（2012年）。[3] Foley, J. A.等。 土地利用的全球影响。《科学》**309**, 570–574（2005年）。[4] Wada, Y. 等。 全球地下水资源枯竭。《地球物理研究快报》**37**, L20402（2010年）。[5]Xu, X. 等。 动物性食品产生的全球温室气体排放量是植物性食品的两倍。《自然·食品》**2**, 724–732（2021年）。[6]Springmann, M. 等。 使粮食系统维持在环境限度之内的方案。*Nature* **562**, 519–525（2018年）。[7] Herforth, A. W., Bai, Y., Venkat, A. & Masters, W. A. 健康膳食篮是有效的全球标准，凸显了人们无法获得健康且可持续的膳食这一问题。《自然·食品》**6**, 622–631（2025年）。[8] Canton, H. 载于联合国粮食及农业组织（*FAO*）第23版，297–305页（Europa Publications，伦敦，2021年）。

[9] He, P. 等。膳食的健康—环境效率在 1990–2011 年间呈现非线性趋势。《自然·食品》5, 116–124 (2024年)。[10]Popkin, B. M. 低收入国家的营养转型：一场正在显现的危机。《营养学评论》52, 285–298 (2009年)。[11] Aune, D.、Ursin, G. 与 Veierød, M. B. 肉类消费与 2 型糖尿病风险：队列研究的系统综述与荟萃分析。Diabetologia 52, 2277–2287 (2009年)。[12] Lopez Barrera, E. 与 Hertel, T. 应对营养不良双重负担的方案也能带来健康与环境效益。《自然·食品》4, 616–624 (2023年)。[13] Huang, T. 等。素食者的心血管疾病死亡率与癌症发病率：一项荟萃分析与系统综述。Annals of Nutrition and Metabolism 60, 233–240 (2012)。[14] Lapp´e, F. M.《小星球的饮食》。(Ballantine Books, 纽约, 1971)。[15] Friedrich, B. 变革有 12,000 年历史的技术。《自然·食品》3, 807–808 (2022年)。[16] Laureati, M.、Delgado, A. M.、De Boni, A. 与 Sinesio, F. 着眼于欧盟更有韧性且可持续的粮食系统，决定消费者接受度与新型食品的采用的因素：一项系统文献综述。Foods 13, 1534 (2024年)。[17] St. Pierre, S. R. 与 Kuhl, E. 模拟力学：肉类与肉类替代品的比较。Foods 13, 3495 (2024年)。[18]Mellor, C. 等。消费者对“藻类”作为蛋白质替代品的认知与接受度：一项基于英国的定性研究。Foods 11, 1703 (2022年)。[19] Spence, C. 多感官风味感知。《细胞》161, 24–35 (2015年)。[20] Kuhl, E. 食品人工智能：加速并普及发现与创新。食品科学。《npj 食品科学》9, 82 (2025)。[21] Spencer, E. H.、Frank, E. 与 McIntosh, N. F. 麦当劳未来售出的 1000 亿个汉堡可能产生的影响。《美国预防医学杂志》28, 379–381 (2005)。[22] 麦当劳在全球交出强劲表现；11 月可比销售额上升 7.4%。美通社 (2011)。[23] 巨无霸指数：巨无霸货币。[24] Clark, M. 等。估算 57,000 种食品的环境影响。《美国国家科学院院刊》119, e2120584119 (2022年)。

[25]Krebs-Smith, S. M. 等。健康饮食指数的更新：HEI-2015。《营养与饮食学会杂志》118, 1591–1602 (2018年)。[26]Datta, B. 等。面向可持续食品未来的人工智能：从数据到晚餐。arXiv doi:10.48550/arXiv.2509.21556 (2025)。[27] Ahn, Y.-Y.、Ahnert, S. E.、Bagrow, J. 与 Barabasi, A.-L. 风味网络与食物搭配原理。《科学报告》1, 196 (2011)。[28]Barabasi, A.-L.、Menichetti, G. 与 Loscalzo, J. 我们饮食中尚未绘制的化学复杂性。《自然·食品》1, 33–37 (2020年)。[29]Torrico, D. D.、Fuentes, S.、Gonzalez Viejo, C.、Ashman, H. 与 Dunshea, F. R. 食品产品熟悉度对消费者感官可接受性及无创生理反应的跨文化影响。《国际食品研究》115, 439–450 (2019年)。[30]Poore, J. 与 Nemecek, T. 通过生产者与消费者减少食品的环境影响。《科学》360, 987–992 (2018年)。[31] Godfray, H. D. J. 等。肉类消费、健康与环境。《科学》361, eaam5323 (2018年)。[32] Willett, W. 等。人类世中的食物：来自可持续粮食系统的健康饮食之 eat–《柳叶刀》委员会。《柳叶刀》393, 447–492 (2019年)。[33] Bryant, C. 与 Barnett, J. 培养肉的消费者接受度：一项系统综述。《肉类科学》143, 8–17 (2020年)。[34]Zeevi, D. 等。通过预测血糖反应实现个性化营养。《细胞》163, 1079–1094 (2015年)。[35] Hoogeboom, E.、Nielsen, D.、Jaini, P.、Forr´e, P. 与 Welling, M. Ranzato, M.、Beygelzimer, A.、Dauphin, Y.、Liang, P. 与 Vaughan, J. W.（编）Argmax 流与多项扩散：学习类别分布。（编 Ranzato, M.、Beygelzimer, A.、Dauphin, Y.、Liang, P. 与 Vaughan, J. W.）《神经信息处理系统进展》，第 34 卷，12454–12465 (2021)。[36] Song, Y. 等。通过随机微分方程进行基于分数的生成建模。arXiv doi:10.48550/arXiv.2011.13456 (2021)。[37] Ho, J.、Jain, A. 与 Abbeel, P. Larochelle, H.、Ranzato, M.、Hadsell, R.、Balcan, M. 与 Lin, H.（编）去噪扩散概率模型。（编 Larochelle, H.、Ranzato, M.、Hadsell, R., Balcan, M. & Lin, H.) 神经信息处理系统进展，第33卷，6840–6851（2020年）。[38]Ta。c, V., Rausch, M. K., Bilionis, I., Sahli Costabal, F. & Tepole, A. B. 利用物理信息概率扩散场构建生成式超弹性模型。计算机工程**41**，51–69（2024年）。[39] Alvin。Food.com - 食谱与评论（2020年）。[40] Wei, Alexander。Food.com 食谱及其食材和标签（2023年）。[41] Taylor Stinson。最佳自制巨无霸食谱（2024年）。[42] Theodora Kaloudis。巨无霸（2024年）。[43] Krystle Smith。巨无霸食谱（2024年）。[44] Ron Dimpflmaier。自制巨无霸。[45]Hellweg, S. & Mil`a I Canals, L. 生命周期评价的新兴方法、挑战与机遇。科学**344**，1109–1113（2014年）。[46]蘑菇。技术报告，ISSN：1949-1530，国家农业统计局（NASS）（2016年）。[47] Goglio, P. 等对欧洲各地双孢蘑菇（((J.E.Lange) Imbach)）生产系统的环境评估。《欧洲农学杂志》**155**，127108（2024年）。[48] Julia, C., Etil´e, F. & Hercberg, S. 法国包装正面 Nutri-Score 标签：一项循证政策。《柳叶刀·公共卫生》**3**，e164（2018年）。[49] Barrett, E. M.等修订健康星级营养素分析算法，以纳入超加工因素。《营养与饮食学》**82**，53–63（2025年）。[50] 2015-2020年《美国人膳食指南》。技术报告，第8版，美国卫生与公众服务部（USDHHS）和美国农业部（USDA）（2015年）。[51] Herforth, A. 等全球基于食物的膳食指南综述。《营养进展》**10**，590–605（2019年）。[52] Bowman, S. A., Clemens, J. C., Friday, J. E. & Moshfegh, A. J. *2017—2018*年食物模式等价数据库：方法与用户指南（美国农业部，马里兰州贝尔茨维尔，2020年）。[53] 美国农业部（USDA）。*FoodData Central*：*Foundation Foods*（农业研究局，2024年）。

[54] Mainardi, F., Drewnowski, A. & Green, H. 食物模式的个性化营养素分析：将 Nestl´e 的营养算法应用于美国国家健康与营养调查（NHANES）的膳食摄入量。《营养素》**11**，379（2019年）。[55]钠和钾膳食参考摄入量审查委员会、食品与营养委员会、健康与医学部及美国国家科学院、工程院和医学院。钠和钾的膳食参考摄入量（美国国家科学院出版社，华盛顿特区，2019年）。[56]膳食参考摄入量框架审查常务委员会、食品与营养委员会、健康与医学部及美国国家科学院、工程院和医学院。重新思考*21*世纪可接受宏量营养素分布范围：信函报告（美国国家科学院出版社，华盛顿特区，2024年）。[57] 世界卫生组织。世卫组织全球减少钠摄入量报告第1版（世界卫生组织，日内瓦，2023年）。[58] 世界卫生组织。指南：成人和儿童糖摄入量（世界卫生组织，日内瓦，2015年）。[59] 世界卫生组织。成人和儿童饱和脂肪酸与反式脂肪酸摄入量：世卫组织指南（世界卫生组织，日内瓦，2023年）。[60] St. Pierre, S.等植物肉与动物肉的机械和感官特征。《*npj* 食品科学》**8**，94（2024年）。

补充文件 1
生成式人工智能创造美味、可持续且营养的汉堡

Vahidullah Tac[1*]、Christopher Gardner[2] 和Ellen Kuhl[1]

[1]机械工程系，斯坦福大学，美国斯坦福。[2]预防研究中心，斯坦福大学医学院，美国斯坦福。

*通讯作者。电子邮箱：vtac@stanford.edu；贡献作者：cgardner@stanford.edu；ekuhl@stanford.edu；

数据准备

本研究使用的数据集由以下网站的食谱开源数据库提炼而成：Food.com [1, 2]。该网站的原始数据库 [1] 缺少食材的计量单位，因此我们将其与一个补充数据库 [2]合并。合并后的数据库包含 522,517 道食谱，由 57,178 位作者提供，涵盖各类食品。表1展示了合并数据库中的一份典型食谱。

识别汉堡食谱

识别汉堡食谱有几种方法，但每种方法都有其局限。一种方法是直接选择名称中提到“汉堡”一词的食谱。然而，这种方法无法正确处理某些条目，例如汉堡汤：它虽然含有“汉堡”一词，却是一种汤；又如迷你牛肉汉堡，它属于汉堡，却不含“汉堡”一词。另一种方法是使用大语言模型。我们结合这两种方法，以尽可能增加训练数据点并减少假阳性。对于每一道食谱，我们都会向大语言模型 Mistral 7B [3], 提供提示，根据食谱名称判断它是否为汉堡食谱。如果大语言模型的搜索结果为肯定，并且名称中含有“汉堡”一词，我们就将其判定为汉堡食谱。

解析汉堡食材

如表1所示，食材及其用量和计量单位均以纯文本形式给出。有些食材描述并不规范，例如：带倍数的用量，如两罐 0.5 夸脱装的罐头；没有计量单位的可数物品，如两个鸡蛋；缺少用量，如一些盐；使用范围而非确切数值，如一到两个番茄；以及带有修饰词的食材，如新鲜牛至。这使得解析食谱颇具挑战。为了规范在线食谱中各种各样的食材描述，我们使用大型语言模型将原始食材文本解析为结构化三元组，

$$\{\texttt{ingredient name}, \texttt{quantity}, \texttt{unit}\}.$$

该模型将用量归一化为单一数值，在可能的情况下将计量转换为质量或体积，对报告的范围取平均值，并移除诸如大小或加工方式等定性描述词。我们

表 1 代表性汉堡食谱。 来自合并后的 food.com数据库的一份典型食谱包含食谱编号和名称、作者编号和姓名、烹饪时间、加工方式和总时间、发布日期、描述、图片、食谱类别、关键词、食材、食谱说明和评分。在 food.com数据库的 522,517 份食谱中，我们筛选出 3,882 份汉堡食谱，并据此生成训练数据。

食谱编号	2734
Name	Ken's Hamburgers
作者编号	1575
作者姓名	lindaWWJD
烹饪时间	10 分钟
准备时间	10 分钟
总时间	20 分钟
发布日期	1999-08-25 05:21:00
描述	制作并分享这份来自 Food.com 的 Ken's Hamburgers 汉堡食谱。
图片	https://img.sndimg.com/food/image/upload/w 555,h 416,c fit, fl-progressive,q 95/v1/img/食谱/27/34/picgsfaLB.jpg, https://img.sndimg.com/food/image/upload/w 555,h 416,c fit,fl progressive,q 95/v1/img/recipes/27/34/picaPkIS9.jpg
食谱类别	Meat
关键词	极低碳水、高蛋白、适合儿童、富含……、炙烤/烧烤，¡ 30 分钟，烤箱，简单
食材（生）	• 1/3 磅瘦牛肉末（每个汉堡） • 1 1/2 茶匙烧烤酱 • 1/4 茶匙脱水洋葱，切碎
食谱说明	• 每个汉堡使用 1/3 磅瘦牛肉末。 • 每个汉堡加入 1 1/2 茶匙烧烤酱和 1/4 茶匙脱水洋葱碎。 • 放入炙烤箱烹制，每隔约 4 分钟翻一次面。 • 用烧烤架烹制也很美味。 • 搭配更多烧烤酱汁，而不是番茄酱。
综合评分	4.0
评论数	5
卡路里	273.2
脂肪含量	15.2
⋮	⋮

随后人工审核并修正解析结果，以解决任何剩余的不一致之处。由此得到的整理后的数据集包含3,882份食谱，其中有3,683种独特食材，以171种不同的计量单位表示。

统一计量单位

进一步处理计量单位，以找出共性并规范化数据。部分加工步骤包括：统一计量单位的拼写，例如 g 与 gr，或茶匙与 tsp；排除使用独特计量单位的食谱；排除使用非标准计量单位（例如罐或包）的食谱；以及解决解析步骤中出现的格式问题。最终，食谱数量进一步精简至 3,725 个，涉及 3,535 种独特食材，并使用 30 种独特计量单位。

统一食材名称

在此阶段，所有食材名称都经过人工处理，以消除大语言模型产物。示例包括：简化文本，例如，将 X 转换为 Y 或 X → Y，而不是仅保留 Y；删除大语言模型生成的额外说明；删除特殊字符，例如 *、"、。；将复数改为单数，例如将 leaves 改为 leaf，或将 tomatoes 改为 tomato；统一同一种食材的多种拼写，例如 chile 和 chili。随后进一步精简食材名称，删除不必要的描述：删除

营销术语和品牌名称，例如将 Heinz 番茄酱改为番茄酱；将所有专有调味料混合物和稀有调味料混合物归入调味料这一总类；合并所有牛肉制品，例如将碎牛肉、牛排、牛肩肉等统称为牛肉；合并鸡肉和猪肉制品；合并名称不同但相同的食材，例如阿多波调味料和阿多波酱，或洋蓟和洋蓟芯；将相似产品合并为一类，例如泡打粉和小苏打；将所有面包分为面包、餐包和汉堡面包三类；并纠正拼写错误。经过这些处理，独特食材减少至仅 709 种。接下来，对食谱列表进行第一次频率筛选，删除所有含有稀有食材的食谱（此类食材的使用次数少于 10 次），因为如此低的使用频率会妨碍对这些食材的学习。然后进一步标准化独特食材列表，包括：将蔬菜和水果归并到各自的总类中，例如将芸豆和斑豆归入豆类；合并所有醋类；合并所有液态油；合并所有鱼类；将所有面包和汉堡面包归并为汉堡面包；并从食材中删除水。最后，对食谱列表进行第二次频率筛选，删除所有含有稀有食材的食谱（此类食材的使用次数少于 15 次）。完成此步骤后，最终整理后的数据集包含 2218 道食谱和 146 种独特食材。

将计量单位转换为克

此步骤的目标是将所有计量单位替换为克。为此，计量单位将经过以下加工步骤：如果没有提供数量，则删除计量单位；将所有体积单位（如加仑、升、汤匙、茶匙）转换为毫升；将所有重量单位（如千克、磅、盎司）转换为克；并手动纠正大语言模型的错误归因，例如，将“一块牛肉”更正为“一块牛肉高汤块”。接下来，使用食物密度数据库中的密度值，将所有食材体积从毫升转换为食材重量（克） [4–6]。最后，将所有剩余的计量单位转换为克 [7]。具体做法是，搜索数据库中的食材名称，并选择最接近的匹配项，为食材名称配对唯一的 USDA FDC 编号。此步骤借助了`usda_fdc` Python 包。搜索结果按以下优先级排序：完全匹配；描述词最少的匹配项，例如，搜索肉汤时，优先选择“高汤块”而非“牛肉高汤块”；来自Foundation Foods 类别的匹配项；以及来自`SR Legacy` `foods` 类别的匹配项。表2 汇总了最终的配料表及其 FDC 编号。

食材用量处理

前文所述的使用代码或大语言模型进行自动化处理的步骤，可能会产生错误条目，例如用量出现非数字条目，如 1/2 而不是 0.5。这些错误会通过人工处理进行更正。随后，根据卡路里对食谱进行标准化：利用 FDC 数据库中每种食材的热量密度，使每份食谱中的食材总计为 500 千卡。在最后这个阶段，数据库中的部分食材仍可能缺少用量。我们用所有包含该食材的食谱中的平均用量替换这些缺失值。综合以上步骤，最终得到汉堡数据库 ，其中每份食谱均表示为

$$\{\text{食材名称；重量（克）}\}.$$

并存储在一个 2,216 x 146 的数组中，其中列对应食谱，行对应食材用量。表 2 汇总了全部 146 种食材及其对应的美国农业部食品数据中心代码。

图 1 展示了数据的一些统计信息，包括按频次和重量计算的最常用的前 10 种食材。图 2 展示了图 1a 中按频次排名前 10 的最常用食材的典型相对重量分布。

建模

在本研究中，食谱指的是用于制作汉堡的食材选择及其用量。总共有146种独特食材（表2），每份食谱都是从该列表中选取少数食材组合而成。典型食谱约含十种食材，很少超过二十种（主文图1）。对此类数据进行建模颇具挑战，因为它由离散数据和连续数据组合而成。我们在此提出一种分两个阶段依次进行的方法：首先，识别将用于制作

表2 本研究中使用的全部146种食材清单。 汉堡配料从最初的3,683种中筛选而来：先合并配料类别，将其缩减至709种，再经过第一次频率筛选；随后再次合并配料类别并进行第二次频率筛选，最终得到按字母顺序排列的146种食材及其美国农业部食品数据中心代码 [7]。

食材名称	FDC ID	食材名称	FDC ID	食材名称	FDC 编号
多香果	171315	咖喱粉	170924	意大利面	168927
杏仁	2707486	dill	172233	花生	2515376
苹果	2709215	面团	172791	胡椒	170931
芝麻菜	2709791	沙拉酱	173592	青酱	171582
牛油果	2709223	egg	171287	腌黄瓜	169379
培根	168277	feta	173420	甜椒	168559
烧烤酱	174523	fish	2706284	菠萝	169124
罗勒	172232	面粉	169761	pork	2705863
bean	173744	大蒜	169230	波多贝罗蘑菇	169255
beef	2514743	姜	169231	土豆	170027
beer	168749	山羊奶酪	2705716	普罗卧干酪	170850
beet	2685576	豪达奶酪	171241	藜麦	168874
蓝纹奶酪	172175	葱	170005	牧场沙拉酱	173592
肉汤	171562	绿叶菜	2709792	调味酱	168561
面包屑	1930877	鳄梨酱	2709307	rice	168879
肉汤	172883	海鲜酱	172886	迷迭香	173473
红糖	2710260	蜂蜜	169640	sage	170935
水牛肉	175299	辣根	173472	莎莎酱	174524
碎小麦	2710820	辣酱	171186	salt	173468
bun	2707784	意大利综合香料	1887966	酱汁	174524
黄油	173430	墨西哥辣椒	168576	香肠	174584
卷心菜	169975	番茄酱	168556	芝麻	170150
卡真调味料	1887349	lamb	2705907	火葱	170499
胡萝卜	2709660	柠檬	2709168	酸奶油	2346387
卡宴辣椒	170932	扁豆	172420	酱油	174278
芹菜	2709778	生菜	2709789	菠菜	2709614
切达奶酪	2705709	lime	168155	牛排酱	171825
其他奶酪	2705764	液体烟熏剂	2680455	牛排调味料	2157193
鸡肉	2706091	芒果	2709242	糖	169655
鸡肉调味料	2060346	马郁兰	170928	葵花籽	170562
鹰嘴豆	173756	蛋黄酱	168112	塔可调味料	172243
辣椒	170497	milk	2705385	芝麻酱	2707587
辣椒碎	579086	mint	173474	龙蒿	170937
chip	2709434	蒙特雷杰克奶酪	2705720	照烧酱	171167
烟熏墨西哥辣椒	2009291	马苏里拉奶酪	170845	百里香	173470
细香葱	2709781	蘑菇	169251	tofu	172476
香菜	169997	芥末	171043	番茄	170457
肉桂	171320	面条	2708354	玉米饼	175036
亮姜籽	170922	肉豆蔻	171326	火鸡肉	171505
corn	2709783	oat	2708489	醋	172241
玉米粉	2710835	oil	748278	核桃	170187
茅屋奶酪	172182	橄榄	2710089	芥末	168583
薄脆饼干	2708184	洋葱	170000	小麦胚芽	168892
蔓越莓	2709279	橙子	169918	wine	173185
奶油	2705597	生蚝	171328	伍斯特酱	171610
奶油奶酪	173418	其他调味料	171331	酵母	167717
黄瓜	2709784	红椒粉	171329	酸奶	171284
孜然	170923	帕尔马干酪	171247	西葫芦	168565
咖喱	2710178	欧芹	170416		

汉堡的食谱中所含食材，使用掩码模型；其次，确定所选食材的用量，使用值模型。第一步是离散建模过程，第二步则处理连续数据。我们使用多项式扩散对食材掩码进行建模，并使用基于分数的生成模型对食材重量进行建模。

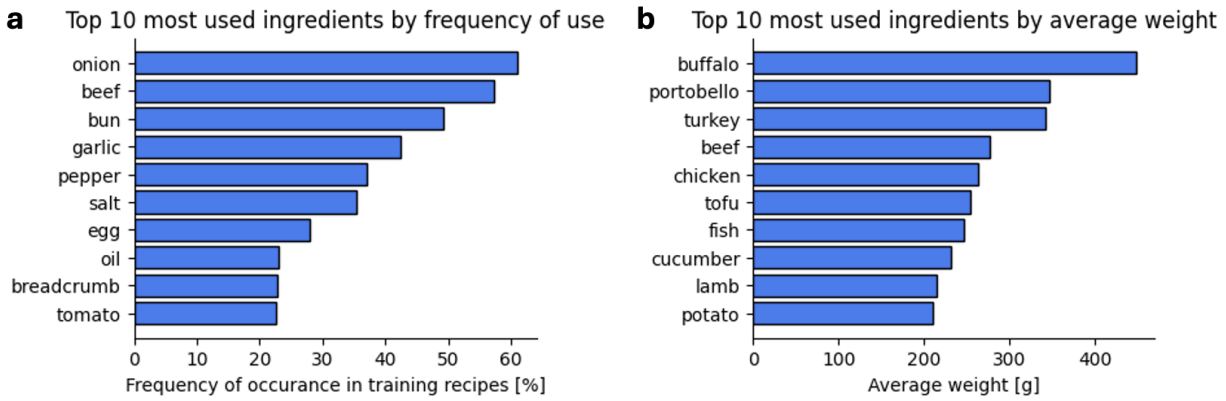

图1 使用频率最高的10种食材。在训练数据的2,216份食谱中，总计146种食材里使用频率最高的食材，按使用频率排序，其中洋葱、牛肉和汉堡面包出现在半数以上的食谱中（a）；按平均重量排序，水牛肉、波多贝罗蘑菇和火鸡肉的用量均超过300克（b）。

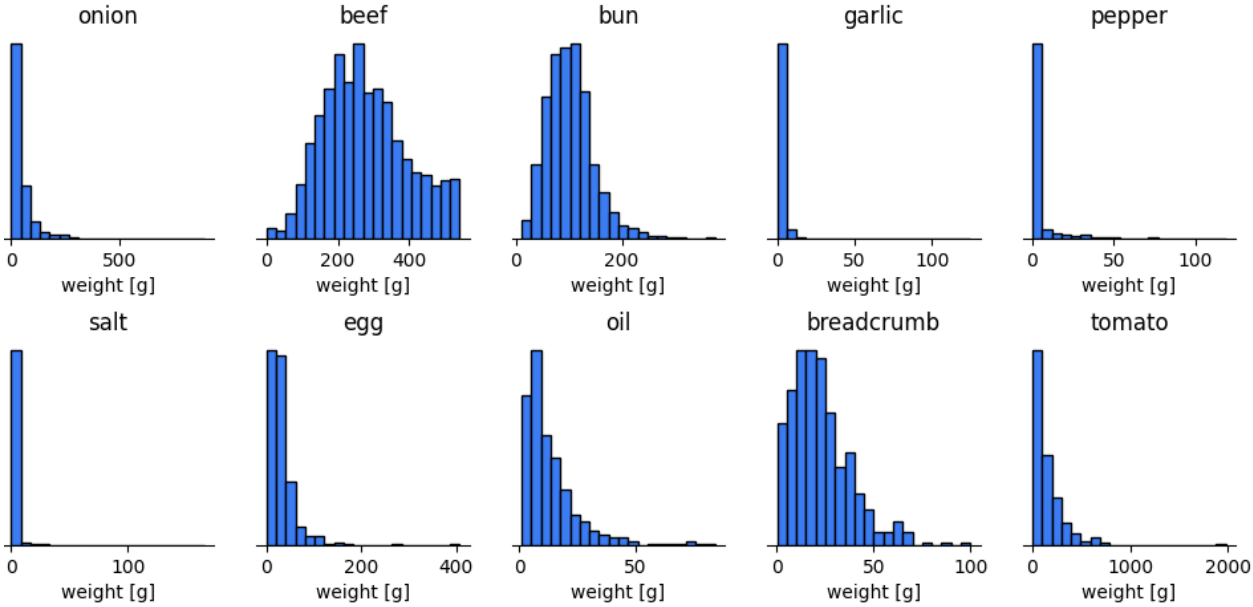

图2 使用频率最高的10种食材的重量分布。在总计146种食材中，使用频率最高的食材，见2,216份食谱，来自我们的训练数据，按使用频率排序（图 1a），并报告其相对分布。

掩码模型

在多项式扩散 [8], 扩散过程是根据类别分布定义的，其具有一个类别，

β_t 均匀重新采样的概率

$$q(\mathbf{m}_t|\mathbf{m}_{t-1}) = \mathcal{C}(\mathbf{m}_t|(1-\beta_t)\mathbf{m}_{t-1} + \beta_t/K)\,. \tag{1}$$

因此，当使用较小的 β_t 时，重新采样到与$\mathbf{m}_{t-1}$相同类别的概率较大，而均匀选择不同类别的概率较小。由于这些分布构成马尔可夫链，任意 t 处的分布都可以用初始数据$\mathbf{m}_0$ 表示为

$$q(\mathbf{m}_t|\mathbf{m}_0) = \mathcal{C}(\mathbf{m}_t|\bar{\alpha}_t\mathbf{m}_0 + (1-\bar{\alpha}_t)/K)\,, \tag{2}$$

其中 $\bar{\alpha}_t = \prod_{\tau=1}^{t}(1-\beta_\tau)$。后验分布的形式如下：

$$q(\mathbf{m}_{t-1}|\mathbf{m}_t, \mathbf{m}_0) = \mathcal{C}(\mathbf{x}_{t-1}|\boldsymbol{\theta}_{\text{post}}(\mathbf{m}_t, \mathbf{m}_0)) \quad \text{with} \quad \boldsymbol{\theta}_{\text{post}} = \boldsymbol{\theta} / \sum_{k=1}^{K} \theta\,, \tag{3}$$

其中， $\boldsymbol{\theta}_{\text{post}}$ 是权重向量 $\boldsymbol{\theta}$ 的归一化向量，该权重向量融合了噪声状态$\mathbf{m}_t$ 和原始数据$\mathbf{m}_0$,中的信息

$$\boldsymbol{\theta} = [(1-\beta_t)\mathbf{m}_t + \beta_t/K] \odot [\bar{\alpha}_{t-1}\mathbf{m}_0 + (1-\bar{\alpha}_{t-1})K]\,, \tag{4}$$

其中， $\odot$表示逐元素哈达玛积。与传统扩散模型直接预测用于预测$\mathbf{m}_{t-1}$ 的参数不同，我们转而对噪声进行建模，从而为后验均值提供显式形式，并得到条件良好的损失函数 [9]。采用这种方法，我们使用神经网络$\hat{\mathbf{m}}_0 = \mu(\mathbf{m}_t, t)$，根据$\mathbf{m}_t$估计$\mathbf{m}_0$ 。然后，后验分布变为

$$p(\mathbf{m}_0|\mathbf{m}_1) = \mathcal{C}(\mathbf{m}_0|\hat{\mathbf{m}}_0) \quad \text{and} \quad p(\mathbf{m}_{t-1}|\mathbf{m}_t) = \mathcal{C}(\mathbf{m}_{t-1}|\boldsymbol{\theta}_{\text{post}}(\mathbf{m}_t, \hat{\mathbf{m}}_0))\,. \tag{5}$$

我们在扩散过程中使用 1,000 个时间步，并采用线性递增序列，使 β_t 从 $\beta(t=0) = 10^{-4}$ 增至 $\beta(t=1000) = 5 \cdot 10^{-3}$。

值模型

我们关注的是在给定 N样本 $\{\mathbf{m}、\mathbf{w}\}^N$的条件下，对分布 $p(\mathbf{w}|\mathbf{m})$进行建模，其中$\mathbf{m}$ 是食谱掩码，$\mathbf{w}$ 是食材重量向量。为简洁起见，以下推导中不明确写出以$\mathbf{m}$ 为条件，但除非另有说明，这一条件均隐含成立。基于分数的生成模型是另一类扩散模型，其加噪和去噪过程通过随机微分方程实现。对于此处的加噪过程，

$$\mathrm{d}\mathbf{w}_t = f(\mathbf{w}_t, t)\mathrm{d}t + g(t)\mathrm{d}B_t\,, \tag{6}$$

而对于去噪过程，

$$\mathrm{d}\mathbf{w}_t = [g^2(t)\nabla_w \log p_t(\mathbf{w}) - f(\mathbf{w}_t, t)]\mathrm{d}t + g(t)\mathrm{d}\tilde{B}_t\,, \tag{7}$$

其中， B_t 是 K维布朗运动， B_t 是其时间反演，而 f 和 g 是用于计算漂移和扩散过程 [10, 11]的函数。尽管这两个随机微分方程有多种选择，但我们采用奥恩斯坦-乌伦贝克过程，因为它较为简单 [10, 12]。对随机微分方程进行缩放后，得到以下方程组：对于加噪过程，

$$\mathrm{d}\mathbf{w}_t = -\frac{1}{2}\beta_t \mathbf{w}_t \mathrm{d}t + \sqrt{\beta_t}\mathrm{d}\mathbf{B}_t \tag{8}$$

而对于去噪过程，

$$\mathrm{d}\mathbf{w}_t = \frac{1}{2}\beta(t)\mathbf{w}_t \mathrm{d}t + \beta(t)\nabla \log p_t(\mathbf{w}_t)\mathrm{d}t + \sqrt{t\beta(t)}\mathrm{d}\mathbf{B}_t\,, \tag{9}$$

其中， $\beta(t) = \beta_{\min} + t(\beta_{\max} - \beta_{\min})$。由此可得条件分布的以下简洁表达式，

$$p(\mathbf{w}_t|\mathbf{w}_0) = \mathcal{N}(\mathbf{w}_t|\mu(t)\mathbf{w}_0, \sigma(t)\mathbf{I}) \tag{10}$$

均值为 $\mu(t)$，方差为 $\sigma(t)$，表示为

$$\mu(t) = \exp(-\alpha(t)/2) \quad \text{and} \quad \sigma(t) = 1 - \exp(-\alpha(t)) \quad \text{with} \quad \alpha(t) = \int_0^t \beta(s)\mathrm{d}s\,. \tag{11}$$

表达式 $\nabla_w \log p_t(\mathbf{w})$ 称为分数函数。我们使用由参数 θ 参数化的神经网络 s 来近似分数函数，对于生成模型，其形式为 $s_{\boldsymbol{\theta}}(\boldsymbol{w},\ t) \approx \nabla_{\mathbf{w}} \log\ p_t(\boldsymbol{w})$。提醒一下，$\mathbf{x}$ 以掩码 $\mathbf{m}$为条件。明确纳入这一条件后，神经网络近似的形式为 $s_{\boldsymbol{\theta}}(\mathbf{w}\ t,\ \mathbf{m}) \approx \nabla_{\mathbf{w}} p_t(\mathbf{w}|\mathbf{m})$。为了训练神经网络，我们最小化以下损失函数

$$L(\boldsymbol{\theta}) = \int_0^T \mathbb{E}_{p_t}\left[||\nabla_{\mathbf{w}} \log p_t(\mathbf{w}) - s_{\boldsymbol{\theta}}(\boldsymbol{w}, t)||^2\right], \tag{12}$$

其中，$p_t(\mathbf{w})$ 使用数据 [10],进行近似

$$\hat{p}_t(\mathbf{w}) = \frac{1}{N}\sum_{i=1}^{N} p(\mathbf{w}|\mathbf{w}_0^i)\,. \tag{13}$$

我们在加噪和去噪过程中各使用 1000 个步骤，并选择缩放参数 $\beta_{\min} = 0.001$ 和$\beta_{\max} = 3$ [13]。

训练和验证

掩码模型 由一个神经网络构建而成，该网络包含一个嵌入层，为时间变量设置了 1000 个嵌入，并包含 3 个全连接层，每层有 512 个神经元，用于预测掩码模型的逻辑值。随后使用 Softmax 函数将逻辑值转换为原始概率。训练时采用小批量处理，批次大小为 $n = 1000$，并使用 Adam 优化算法，以 $5 \cdot 10^{-4}$ 的学习率训练 100,000 个训练轮次。

值模型 由一个前馈神经网络构建而成，包含 4 个隐藏层，每层有 256 个神经元。将 $\mathbf{x}$ t 和 $\mathbf{m}$ 的向量拼接，作为 $s_{\boldsymbol{\theta}}$ 的输入。我们使用 400 的批次大小和 $1 \cdot 10^{-3}$ 的学习率，采用 Adam 优化算法训练 20,000 个训练轮次。训练数据按 80% 用于训练、20% 用于验证的比例进行划分。训练在 Nvidia H100 和 L40S GPU 上进行。

巨无霸的再发现®

原版巨无霸®重 210 克，含 563 千卡。由于其食谱属于专有信息，我们参考开源复刻版本 [14–17] 对其进行近似还原。我们取四个掩码的交集，找出其中共有的食材，

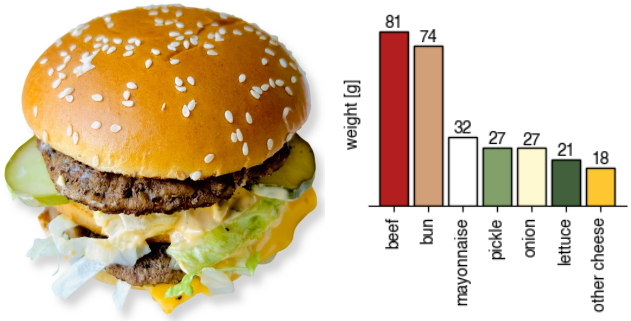

图 3 经典巨无霸®作为再发现的基准。由于巨无霸®（左图）的原始食谱属于专有信息，我们参考在线食谱估算其食材及重量：碎牛肉 81 克、汉堡面包 74 克、蛋黄酱 32 克、腌黄瓜 27 克、洋葱 27 克、生菜 21 克、奶酪 18 克（右图）。

并将其用量按卡路里标准化后取平均值。由此得出的巨无霸®食谱中，500 千卡份量由七种食材组成，重量分别为：81 克碎牛肉、74 克汉堡面包、32 克蛋黄酱、27 克腌黄瓜、27 克洋葱、21 克生菜和 18 克奶酪（图1b 正文稿，图 3）。

实质差异分数

为了量化食谱之间的相似程度并将相似食谱分组，我们引入实质差异分数（实质差异分数）。我们将两份食谱 r_1 和 r_2 之间的实质差异定义为 d_i 在数据库中所有食材上的二元距离之和，

$$\mathrm{SDS}(r_1, r_2) = \sum_{i=1}^{n_{\mathrm{ing}}} d_i(r_1, r_2), \quad \text{with} \quad d_i(r_1, r_2) = \begin{cases} 1 & \text{if } r_{1i} + r_{2i} \neq 0 \text{ and } r_{1i} \cdot r_{2i} = 0 \\ 1 & \text{if } \max(r_{1i}, r_{2i}) / \min(r_{1i}, r_{2i}) \geq 2 \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

实质差异分数为零，实质差异分数 $= 0$，即两份食谱含有相同的食材，且其用量之比始终不超过2。实质差异分数为零的两份食谱 r_1 和 r_2 称为重复食谱或匹配。在再发现过程中，我们使用为零的实质差异分数实质差异分数 $= 0$，来识别人工智能生成的食谱与人工设计的食谱之间的匹配。在发现过程中，我们使用大于零的数值实质差异分数 > 0，来量化人工智能生成的食谱相对于训练集中人工设计食谱的新颖性。正文中的美味汉堡1是重复次数最多的食谱，其实质差异分数 $= 3$ （图2c、f，正文）。它含有151克碎牛肉、79克番茄、75克莳萝腌黄瓜、47克黄洋葱、30克布里欧修面包、19克蛋黄酱、12克美式奶酪、7克生菜、7克番茄酱和1克盐。正文中的美味汉堡2是重复次数最多的食谱，其实质差异分数 $= 6$ （图2d、g，正文）。它含有62.0克碎牛肉、35.5克汉堡面包、28.0克番茄片、12.5克雷穆拉德酱、8.0克Gruy`ere奶酪、3.5克黄洋葱、2.0克伍斯特酱、1.0克红糖、0.5克盐、0.2克大蒜和0.2克新鲜百里香叶。

环境影响评分

生命周期评估通过调查全球生产者，估算农产品在生产和分销链中的总环境影响 [18]。在此，我们使用环境数据库 [19], 量化可持续性，该数据库是一项荟萃分析，整合了来自 $n = 570$ 项研究的数据。由于该数据库未包含蘑菇，我们补充了美国农业部提供的栽培蘑菇平均土地利用数据 [20], ，以及来自三个欧洲国家的淡水富营养化潜势、按水资源稀缺性加权的用水量和温室气体排放数据 [21]。为便于优化和比较，我们将环境影响归纳为一个环境影响评分（环境影响评分） [22]。我们通过对四个类别的影响进行归一化并取平均值，计算环境影响评分：土地利用（LU）、水生富营养化潜势（EU）、按水资源稀缺性加权的用水量（WU）和温室气体排放（GHG），计算方式如下

$$\mathrm{EIS} = \frac{1}{4}\left[\frac{\mathrm{LU}}{\mathrm{LU}_{\max}} + \frac{\mathrm{EU}}{\mathrm{EU}_{\max}} + \frac{\mathrm{WU}}{\mathrm{WU}_{\max}} + \frac{\mathrm{GHG}}{\mathrm{GHG}_{\max}}\right] \tag{15}$$

其中，$\mathrm{LU}_{\max}$、$\mathrm{EU}_{\max}$、$\mathrm{WU}_{\max}$和$\mathrm{GHG}_{\max}$ 的数值直接取自 [22] ，分别为62.2 m^2、20.2 kg CO_2当量、174 g PO_4 当量和43,600 L。根据这些信息，我们计算了本研究中每种食材的环境影响评分（图4）。我们将各食材环境影响评分的贡献值按食材用量加权后求和，以此量化食谱的环境影响。我们手动将模型中的食材（表2）匹配到环境数据库中的某一类别（表3） [19]。本研究中的大多数食材都是完整的农产品；不过，少数加工食材（如调味品）由多种成分构成。由于缺乏一致的环境数据，且这些成分按质量计的贡献可以忽略不计，因此我们在环境影响计算中排除了调味品、浓汤块、肉汤、酵母和醋。

通过找出在特定范围的环境影响评分 中反复出现的食谱，确定环境可持续汉堡。例如，我们可以生成环境影响评分 $0.1 \leq$ < 0.2 和 $0.2 \leq$ 环境影响评分 < 0.3的汉堡食谱（图5a、b）。正文中的可持续汉堡1是一个普通蘑菇汉堡，环境影响评分为0.06，也是环境影响评分 < 0.1下重复次数最多的食谱。（正文图3c、f）。它含有226克波多贝罗蘑菇、78克蛋黄酱、46克布里欧修面包、21克芝麻菜、9克棕色辣芥末、8克特级初榨橄榄油、3克迷迭香、2克新鲜大蒜和1克犹太洁食盐。正文中的可持续汉堡2 是牛肉蘑菇混合汉堡，环境影响评分为0.51，也是同时含有牛肉和蘑菇的食谱中重复次数最多的一个（正文图3d、g）。它含有70克碎牛肉、42克

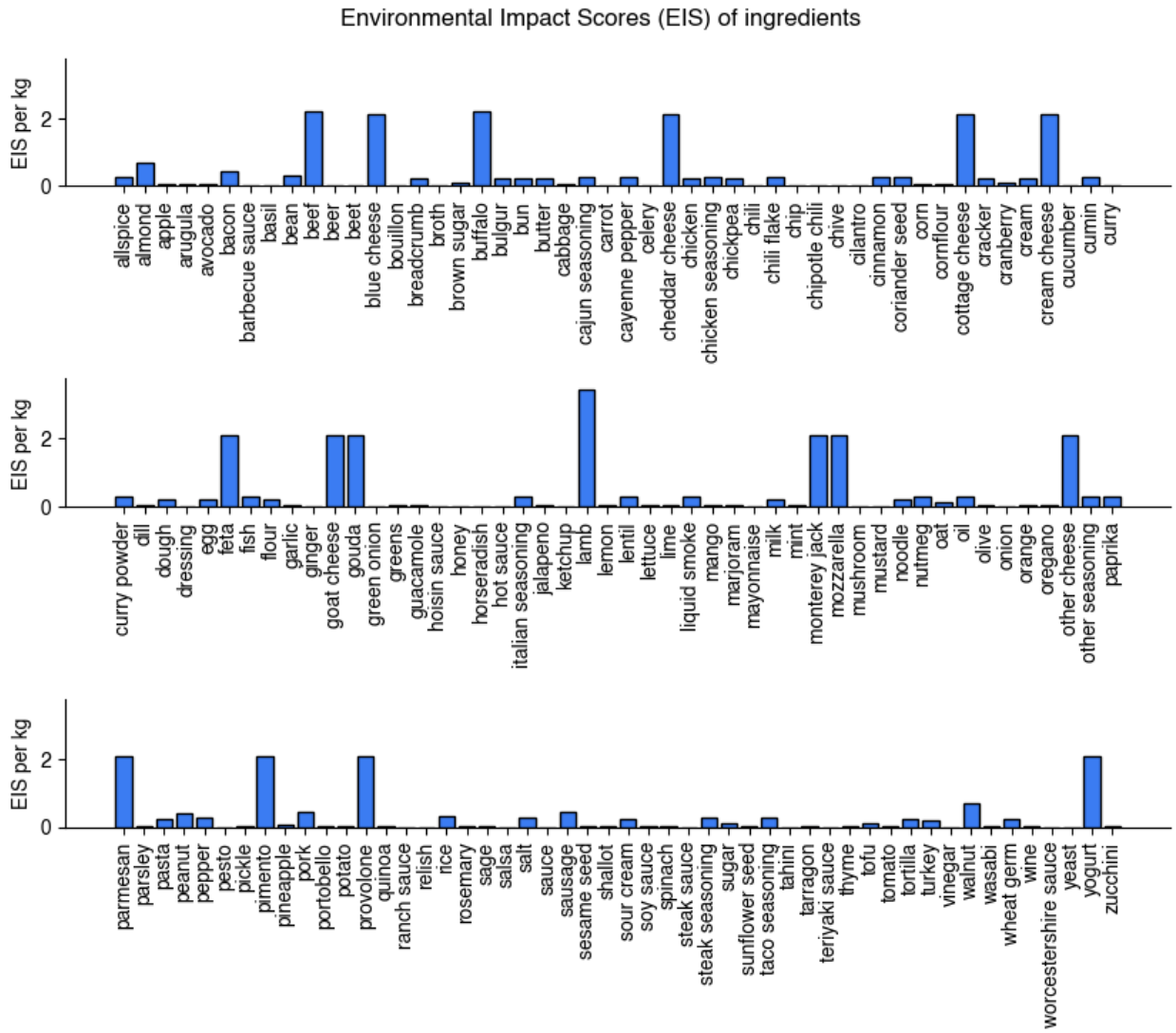

图 4 食材的环境影响评分。报告的环境影响评分衡量了数据库中146种食材在四个类别中的每千克综合影响：土地利用、水生富营养化潜势、按稀缺程度加权的用水量，以及温室气体排放。

布里欧修面包、39克切达奶酪、19克培根、17克番茄酱、14克褐蘑菇、9克黄洋葱、1克黑胡椒和1克犹太洁食盐。

健康饮食指数

营养画像模型通过将营养素和食物组组成与膳食指南进行比较，量化食物和膳食的营养质量。常见的营养画像模型包括健康饮食指数（HEI） [23], Nutri-Score [24], 和健康星级评级 [25]。健康饮食指数由美国农业部开发，用于评估与《美国人膳食指南》的符合程度 [26] ，采用0–100分制，评分越高表示遵循程度越高。该指数主要强调食物组充足性，例如水果、蔬菜、全谷物和乳制品，而非单个营养素，这与全球转向基于食物的膳食指南的趋势一致 [27]。相比之下，其他营养画像模型明确强调受到关注的营养素，包括蛋白质、膳食纤维、钾、钙、铁和维生素D [28—30], 在某些情况下，还包括黄酮类化合物等生物活性化合物 [31]。在此，我们使用健康饮食指数作为主要指标，评估生成的汉堡食谱的营养质量。我们

表 3 我们的 146 种汉堡配料与环境数据库 LCA 类别的对应关系。 为计算环境影响评分，我们将模型中的 146 种配料（表 2）分别对应到环境数据库中的一个 LCA 类别。

LCA 类别	汉堡配料
小麦与黑麦（面包）	面包屑、碎小麦、汉堡面包、薄脆饼干、面团、面粉、面条、意大利面、玉米饼、小麦胚芽
玉米（粉）	玉米, 玉米粉
大麦（啤酒）	beer
燕麦片	oat
Rice	rice
马铃薯	薯片, 土豆
糖	红糖，糖
其他豆类	bean, 扁豆
Peas	鹰嘴豆
Nuts	杏仁 ，核桃
落花生	花生
豆浆	酱油
Tofu	tofu
Oil	oil
番茄	番茄
洋葱与韭葱	葱、细香葱、洋葱、火葱
根茎类蔬菜	甜菜、胡萝卜、姜、辣根、芥末
十字花科蔬菜	芝麻菜、卷心菜、生菜
其他蔬菜	罗勒、芹菜、辣椒、烟熏墨西哥辣椒、香菜、黄瓜、莳萝、大蒜、绿叶菜、墨西哥辣椒、马郁兰， 薄荷、牛至、欧芹、腌黄瓜、藜麦、迷迭香、鼠尾草、芝麻、菠菜, 葵花籽、龙蒿、百里香、西葫芦
柑橘类水果	柠檬、青柠、橙子
苹果	苹果
浆果和葡萄	蔓越莓
Wine	wine
其他水果	牛油果、芒果、鳄梨酱、橄榄、菠萝
Beef	beef
羊肉	lamb
猪肉	培根、猪肉、香肠
禽肉	鸡肉、火鸡肉
Milk	milk
奶酪	蓝纹奶酪、切达奶酪、茅屋奶酪、奶油奶酪、菲达奶酪、山羊奶酪、豪达奶酪、蒙特雷杰克奶酪、马苏里拉奶酪、其他奶酪、帕尔马干酪、甜椒、普罗卧干酪、酸奶
Eggs	egg
养殖鱼	fish
甜味剂和蜂蜜	蜂蜜
兴奋剂 & 香料	多香果、卡真调味料、卡宴辣椒、鸡肉调味料、辣椒碎、肉桂,
其他	芫荽籽、孜然、咖喱粉、意大利综合香料、液体烟熏剂、肉豆蔻、其他海盐、红椒粉、胡椒、盐、牛排调味料、塔可调味料
水牛肉	水牛肉
黄油 ，奶油和酥油	黄油 ，奶油 ，酸奶油
蘑菇	波多贝罗蘑菇、蘑菇
调味品及其他	烧烤酱、咖喱、沙拉酱、海鲜酱、辣酱、番茄酱、蛋黄酱、芥末, 青酱、牧场沙拉酱、调味酱、莎莎酱、酱汁、牛排酱、芝麻酱、照烧酱、伍斯特酱酱汁、肉汤、肉汤、酵母、醋

从生成模型中抽取大量食谱，并在健康饮食指数分布的第90百分位数和第95百分位数样本中，选取出现频率最高的食谱，以确定具有代表性的营养汉堡。选择重复率高的样本可确保最终食谱对应于生成模型下高概率的区域，因此更有可能适口且受欢迎。主文稿中的营养汉堡是一款健康饮食指数为63.12的豆类汉堡，也是健康饮食指数第95百分位数样本中重复次数最多的食谱（主文稿图4d）。其配料包括71克罐装芸豆、50克布里欧修面包、15克压片燕麦、10克鸡蛋、4克玉米粉、4克特级初榨橄榄油、2克墨西哥辣椒˜、2克香菜、2克白洋葱、0.8克犹太洁食盐、0.2克孜然、0.2克黑胡椒，以及

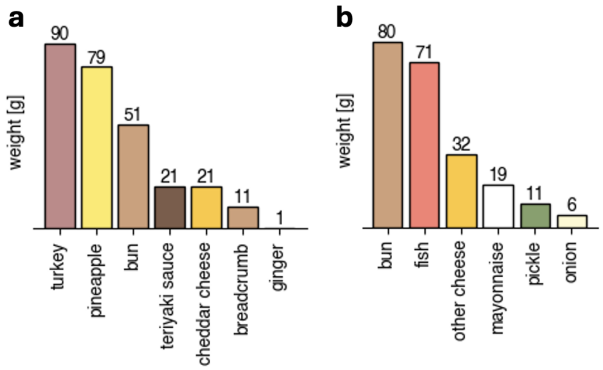

图5 替代性可持续汉堡。 环境影响评分为 $0.1 \leq \text{EIS} < 0.2$ (a) 和$0.2 \leq \text{EIS} < 0.3$ (b) 的最常见食谱，其排名介于正文中的可持续汉堡1（$\text{EIS} = 0.03$）与可持续汉堡2（$\text{EIS} = 0.51$）之间。

0.05克墨西哥牛至。作为比较，健康饮食指数第90百分位数中最常见的食谱是一种食材和营养成分概况各不相同的豆类-玉米混合物（图6a、b）。

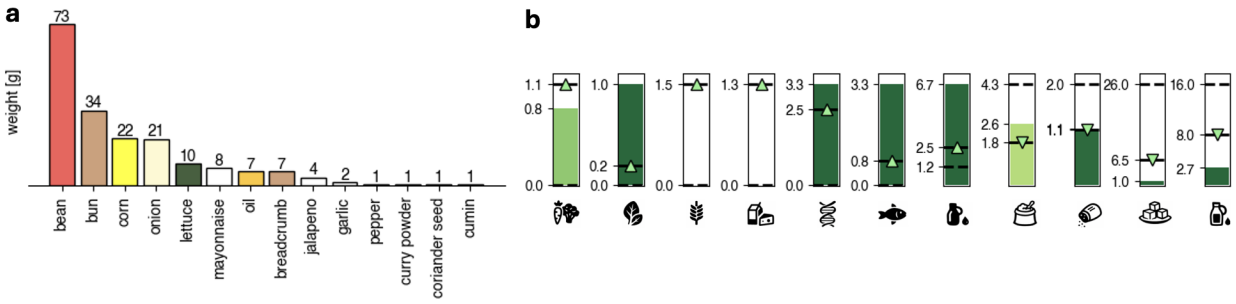

图6 替代性营养汉堡。 健康饮食指数（HEI）第90百分位数中最常见的食谱：一种豆类-玉米混合物，其食材见(a)，营养成分概况见(b)。

此外，我们使用平衡混合营养素密度评分（平衡混合营养密度评分） [28, 32]，这一结合食物组与营养素的营养分析模型，生成替代性营养汉堡配方。为进行比较，我们根据平衡混合营养密度评分，生成了第95和第90百分位数的营养汉堡食谱（图 7a、b）。

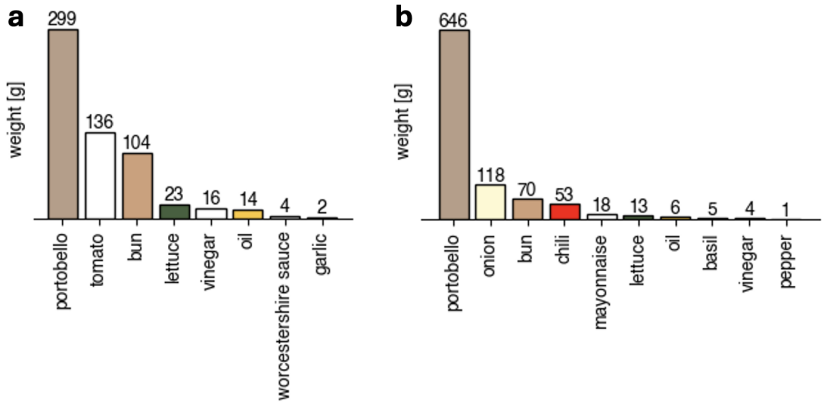

图7 替代性营养汉堡。 按平衡混合营养素密度评分（bHNDS）排名的第95百分位数(a)和第90百分位数(b)中最常见的食谱。

个性化营养

营养需求会因年龄、性别、身体组成、身体活动和生理状态而在个体之间存在显著差异。例如，19至50岁成年女性的铁膳食参考摄入量，是同年龄组男性的两倍以上 [33]。为考虑这种异质性，我们采用个性化营养素概况模型 [34] ，据此生成符合个人营养需求的食谱。我们根据个人特征计算0–100分的个性化营养评分，这些特征包括年龄、性别、体重、身高和身体活动水平。我们依据膳食参考摄入量和可接受宏量营养素供能比范围 [33, 35]，并结合世界卫生组织关于钠、游离糖和饱和脂肪最高摄入量的指南 [36–38]，推导出推荐摄入范围。这些建议共同界定各营养素的特定目标范围，并汇总为每款汉堡的单一个性化营养评分。运用这一框架，我们为四种具有代表性的人口统计特征生成个性化汉堡食谱：一名30岁、体重80千克、身高180厘米的久坐男性；一名30岁、体重60千克、身高160厘米的久坐女性；一名15岁、体重55千克、身高160厘米的活动量极大的男性；以及一名70岁、体重70千克、身高170厘米的中等活动量女性（图 8）。

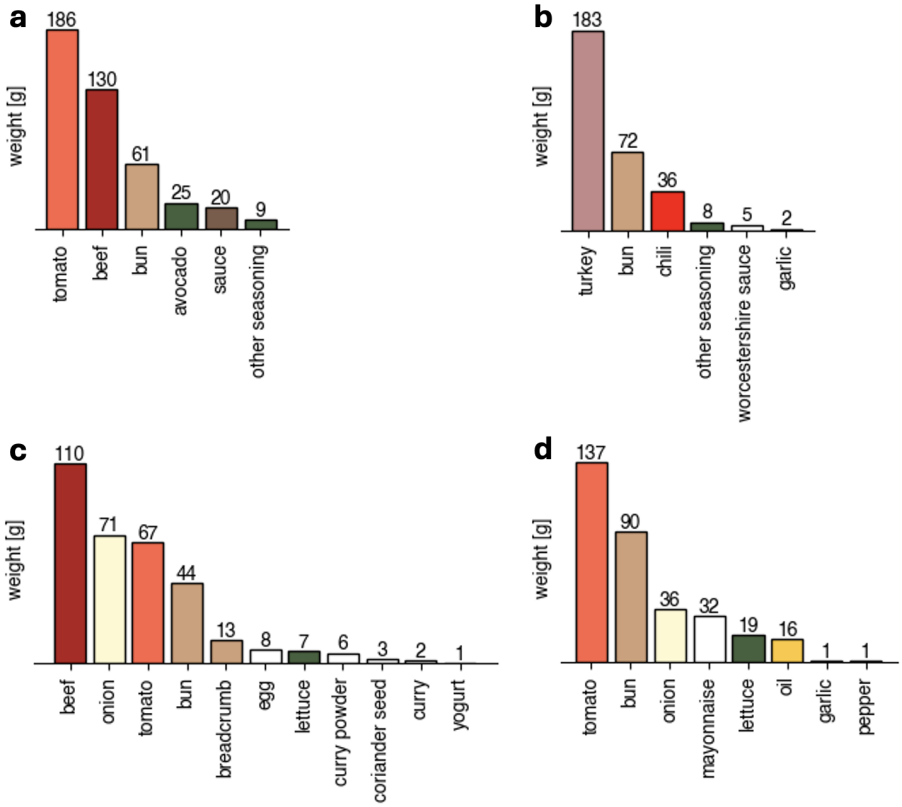

图 8 个性化营养汉堡。为久坐、体重 80 kg、身高 180 cm 的 30 岁男性设计的个性化食谱（a）；为久坐、体重 60 kg、身高 160 cm 的 30 岁女性设计的个性化食谱（b）；为非常活跃、体重 55 kg、身高 160 cm 的 15 岁男性设计的个性化食谱（c）；以及为中等活跃、体重 70 kg、身高 170 cm 的 70 岁女性设计的个性化食谱（d）。

参考文献

[1] Alvin。Food.com - 食谱与评论（2020年）。

[2] Wei, Alexander。Food.com 食谱、食材与标签（2023年）。

[3] Jiang, A. Q.等。 Mistral 7B（2023年）。 arXiv:2310.06825。

[4] Charrondiere, U. Ruth、Haytowitz, David 和 Stadlmayr, Barbara。FAO/INFOODS 数据库（2012年）。

[5] Aqua-Calc。Aqua-Calc：水的物理性质在线计算器。www.aqua-calc.com（2025年）。访问日期：2025年12月13日。[6] kg-m3。kg/m3：关于材料密度的一切。www.kg-m3.com（2025年）。访问日期：2025年12月13日。[7] 美国农业部（USDA）。*FoodData Central*：*Foundation Foods*（农业研究局，2024年）。[8] Hoogeboom, E.、Nielsen, D.、Jaini, P.、Forr´e, P. 和 Welling, M. Ranzato, M.、Beygelzimer, A.、Dauphin, Y.、Liang, P. 和 Vaughan, J. W.（编）*Argmax* 流与多项式扩散：学习分类分布。（编者：Ranzato, M.、Beygelzimer, A.、Dauphin, Y.、Liang, P. 和 Vaughan, J. W.）神经信息处理系统进展，第34卷，12454–12465（2021年）。[9] Ho, J.、Jain, A. 和 Abbeel, P. Larochelle, H.、Ranzato, M.、Hadsell, R.、Balcan, M. 和 Lin, H.（编）去噪扩散概率模型。（编者：Larochelle, H.、Ranzato, M.、Hadsell, R.、Balcan, M. 和 Lin, H.）神经信息处理系统进展，第33卷，6840–6851（2020年）。[10] Ta˙c, V.、Rausch, M. K.、Bilionis, I.、Sahli Costabal, F. 和 Tepole, A. B. 基于物理信息概率扩散场的生成式超弹性。计算机工程**41**，51–69（2024年）。[11]Song, Y. 等通过随机微分方程进行基于分数的生成建模。*arXiv*doi:10.48550/arXiv.2011.13456（2021年）。[12] Pidstrigach, J. Koyejo, S. 等（编）基于分数的生成模型可检测流形。（编者：Koyejo, S.等）神经信息处理系统进展，第35卷，35852–35865（2022年）。[13] Karras, T.、Aittala, M.、Aila, T. 和 Laine, S. Koyejo, S. 等（编）阐明基于扩散的生成模型的设计空间。（编者：Koyejo, S. 等）神经信息处理系统进展，第35卷，26565–26577（2022年）。[14] Taylor Stinson。《最佳自制巨无霸食谱》（2024年）。

[15] Theodora Kaloudis。《巨无霸》（2024年）。

[16] Krystle Smith。《巨无霸食谱》（2024年）。[17] Ron Dimpflmaier。《自制巨无霸》。[18] Hellweg, S. & Mil`a I Canals, L. 生命周期评价中的新兴方法、挑战与机遇。《科学》**344**, 1109–1113 (2014年)。[19] Poore, J. & Nemecek, T. 通过生产者和消费者减少食品的环境影响。《科学》**360**, 987–992 (2018年)。[20] 蘑菇。技术报告。ISSN：1949-1530，国家农业统计局（NASS）（2016年）。[21] Goglio, P. 等。 欧洲各地双孢蘑菇（((J.E.Lange) Imbach)）生产系统的环境评估。《欧洲农学杂志》**155**, 127108 (2024年)。[22] Clark, M. 等。 估算57,000种食品的环境影响。《美国国家科学院院刊》**119**, e2120584119 (2022年)。[23] Krebs-Smith, S. M.等。 健康饮食指数更新：HEI-2015。《营养与饮食学会杂志》**118**, 1591–1602 (2018年)。[24] Julia, C., Etil´e, F. & Hercberg, S. 法国的包装正面 Nutri-Score 标签：一项循证政策。《柳叶刀·公共卫生》**3**, e164 (2018年)。

[25] 澳大利亚联邦。健康星级评级系统。https://www.healthstarrating.gov.au/ (2014年)。网址 https://www.healthstarrating.gov.au/。访问日期：2025年12月10日。[26] 2015-2020年《美国人膳食指南》。技术报告。第8版，美国卫生与公众服务部（USDHHS）和美国农业部（USDA）（2015年）。[27] Herforth, A.等。 基于食物的膳食指南全球综述。《营养进展》**10**, 590–605 (2019年)。[28] Drewnowski, A., Dwyer, J., King, J. C. & Connie M Weaver. 一项纳入食物类别和营养素、以更好契合膳食指导的营养素密度评分提案。《营养学评论》**77**, 404–416 (2019年)。[29] Fulgoni, V. L., Keast, D. R. & Drewnowski, A. 营养丰富食品指数的开发与验证：一种衡量食物营养质量的工具。《营养学杂志》**139**, 1549–1554 (2009年)。[30] Drewnowski, A. & Fulgoni, V. L. 营养素密度：原理与评估工具。《美国临床营养学杂志》**99**, 1223S–1228S (2014年)。[31] Drewnowski, A. & Burton-Freeman, B. 一种全新的类别特异性营养丰富食品（NRF9f.3）评分，将黄酮类化合物纳入水果营养素密度评估。《食品与功能》**11**, 123–130 (2020年)。[32] Drewnowski, A., Gonzalez, T. D. & Rehm, C. D. 采用受试者工作特征曲线分析，将平衡混合营养素密度评分与 Nutri-Score 和健康星级评级进行比较。《营养学前沿》**9**, 867096 (2022年)。[33] 钠和钾膳食参考摄入量审查委员会、食品与营养委员会、健康与医学部及美国国家科学、工程与医学研究院。钠和钾的膳食参考摄入量（美国国家科学院出版社，华盛顿特区，2019年）。[34] Mainardi, F., Drewnowski, A. & Green, H. 饮食模式的个性化营养素画像：将 Nestl´e 的营养算法应用于美国国家健康与营养调查（NHANES）的膳食摄入量。《营养素》**11**, 379 (2019年)。[35] 膳食参考摄入量框架审查常务委员会、食品与营养委员会、健康与医学部及美国国家科学、工程与医学研究院。重新思考*21*世纪可接受宏量营养素分布范围：信函报告（美国国家科学院出版社，华盛顿特区，2024年）。[36] 世界卫生组织。世卫组织全球减少钠摄入量报告第1版（世界卫生组织，日内瓦，2023年）。[37] 世界卫生组织。指南：成人和儿童糖摄入量（世界卫生组织，日内瓦，2015年）。[38] 世界卫生组织。成人和儿童饱和脂肪酸与反式脂肪酸摄入量：世卫组织指南（世界卫生组织，日内瓦，2023年）。

补充文件 2

生成式人工智能创造美味、可持续且营养丰富的汉堡

Vahidullah Tac[1*]、Christopher Gardner[2] 和Ellen Kuhl[1]

[1]斯坦福大学机械工程系，美国斯坦福。[2]斯坦福大学医学院预防研究中心，美国斯坦福。

*通讯作者。电子邮箱：vtac@stanford.edu；共同作者：cgardner@stanford.edu；ekuhl@stanford.edu；

感官评价

我们从一般人群中招募了 $n = 101$ 名自愿参与者，在加利福尼亚州旧金山的一家营业餐厅，对全部六种汉堡进行盲法感官评价。本研究遵循斯坦福大学机构审查委员会的指南。我们请参与者回答七个有关其背景的问题，然后针对每种汉堡，在7点李克特量表上评价总体喜好度、风味和质地，并从12种风味属性和15种质地属性中勾选所有适用项。



图1 感官评价。 一般人群中的自愿参与者在一家营业中的餐厅参与了全部六种汉堡的盲法感官评价。这些照片记录了全部六种汉堡的加工过程，以及品尝前标记为413的可持续汉堡1、标记为674的美味汉堡1和标记为491的营养汉堡。

参与者信息

您的年龄是多少?

选择一项

- 18岁以下
- 18-25
- 26-35
- 36-45
- 46-55
- 55岁以上

您的性别是?

请选择一项

- 男性
- 女性
- 非二元性别
- 不愿透露

您已完成的最高学历或教育程度是什么?

请选择一项

- 高中未毕业
- 高中
- 上过大学
- 学士学位
- 硕士学位
- 博士学位或更高学历
- 职业学校

您多久吃一次汉堡?

请选择一项

- 每天
- 2每周 2-3 次
- 每周一次
- 2每月 2-3 次
- 每 1-2 个月一次
- 4-每年 5 次
- 2-每年 3 次
- 从不或很少

您的饮食偏好是什么?

选择一项

- 纯素
- 素食
- 鱼素
- 弹性素食者（即偶尔吃肉类）
- 杂食者（即经常吃肉类）

您通常在哪里购买日常杂货?

选择所有适用项

- 天然食品杂货店（例如全食超市）
- 折扣杂货店（例如 Trader Joe’s、Lidl、Aldi）
- 传统杂货店（例如 Stop & Shop、Safeway）
- 大型零售商（例如 Target、沃尔玛）
- 杂货配送服务（例如 Instacart、Amazon Fresh）
- 预制菜套装（例如 Blue Apron）
- 散装食品杂货店（例如好市多、山姆会员店）
- 社区小型杂货店
- 直接从品牌/制造商网站购买

决定吃什么时，您最看重什么?

请选择所有适用项

- 健康
- 价格
- 口味
- 便利性
- 熟悉度（我习惯的）
- 环境
- 动物福利

感官反馈

您对 Burger X 的总体喜好度如何评价?

请选择一项

- 非常喜欢
- 喜欢
- 有些喜欢
- 既不喜欢也不讨厌
- 有些讨厌
- 讨厌
- 非常讨厌

您如何评价 Burger X 的风味?

请选择一项

- 非常喜欢
- 喜欢
- 有些喜欢
- 既不喜欢也不讨厌
- 有些讨厌
- 讨厌
- 非常讨厌

请勾选所有描述 Burger X 风味的词语或短语。

选择所有适用项

- 肉味
- 怪异余味
- 异味
- 清淡无味
- 泥土味
- 烟熏味
- 浓烈
- 甜味

- 咸
- 油腻
- 鲜美
- 回味好

您如何评价 Burger X 的质地?

请选择一项

- 非常喜欢
- 喜欢
- 有些喜欢
- 既不喜欢也不讨厌
- 有些讨厌
- 讨厌
- 非常讨厌

请勾选所有能够描述 Burger X 质地的词语或短语。

选择所有适用项

- 有嚼劲
- 酥脆 / 香脆
- 易碎 / 粗糙
- 紧实 / 硬
- 柔软 / 糊状
- 不易散开
- 湿润
- 干燥
- 坚韧
- 油腻
- 纤维状 / 有丝状感
- 易碎
- 黏软
- 有弹性的
- 黏性的

人口统计学特征

该 $n = 101$ 研究参与者代表了一般人群的均衡横截面（图2）。在$n = 101$参与者中，47.5%为男性，47.5%为女性，3%为非二元性别，2%不愿透露。22%年龄为18至25岁，26%为26-35岁，19%为36-45岁，18%为46-55岁，16%年龄超过55岁。65%为杂食者，35%为弹性素食者。其最高学历分别为：4%高中学历、24%大学、50%学士学位、11%硕士学位、8%博士学位或更高学历，以及3%职业学校学历。4%每天吃汉堡，20%每周吃2至3次，31%每周吃1次，27%每月吃2至3次，16%每1至2个月吃1次，3%每年吃4至5次。在杂货购买习惯方面，37%偏好天然食品杂货店，55%偏好折扣杂货店，58%偏好传统杂货店，12%偏好大型零售商，12%偏好杂货配送服务，30%偏好散装食品杂货店，26%偏好社区小型杂货店，2%偏好制造商网站。在购买优先考虑因素方面，50%列出了健康，61%列出了价格，76%列出了口味，25%列出了便利性，16%列出了熟悉度，20%列出了环境，18%列出了动物福利。

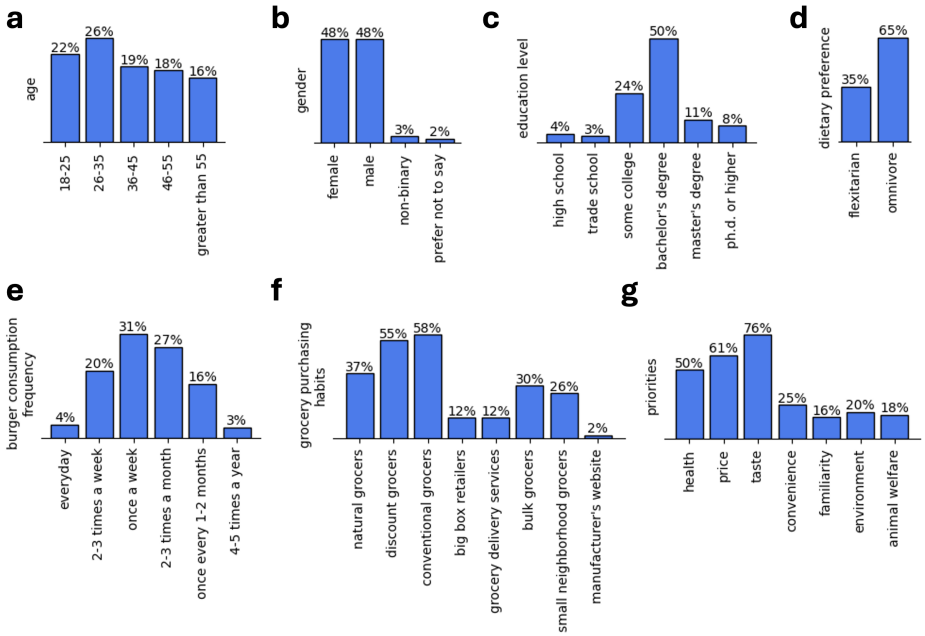

图2 人口统计学特征。 问卷受访者的人口统计特征和饮食偏好概览。根据前七个调查问题的回答，展示参与者年龄、性别、教育程度、饮食偏好、汉堡消费频率、杂货购买习惯和优先事项。

补充文件 3
生成式人工智能打造美味、可持续且营养的汉堡

Christopher Gardner[2] 与 Ellen Kuhl[1]

[1]斯坦福大学机械工程系，斯坦福，美国。[2]斯坦福大学医学院预防研究中心，斯坦福，美国。

贡献作者：cgardner@stanford.edu；ekuhl@stanford.edu；

食谱开发指南

本文档已交给一位行政主厨，其中包含人工智能生成的配料清单及对应的用量，并说明了应如何解读这些清单。

食谱

本研究包含 5 款汉堡。以下列出了这 5 款汉堡的食材及其用量 .
厨师须遵循的 指南，见下一节。

美味汉堡1

- 牛肉 151 g
- 汉堡面包 30 g
- 番茄酱 7 g
- 生菜 7 g
- 蛋黄酱 19 g
- 洋葱 47 g
- 其他奶酪 12 g
- 腌黄瓜 75 g
- 番茄 79 g

美味汉堡2

- 牛肉 124 g
- 红糖 2 克
- 汉堡面包 71 克
- 大蒜 0.4 克
- 洋葱 7 克
- 其他奶酪 16 克
- 其他酱汁 25 克
- 百里香 0.3 g
- 番茄 56 g
- 伍斯特酱 4 克

可持续汉堡1

- 芝麻菜 21 克
- 汉堡面包 46 克
- 大蒜 2 克
- 蛋黄酱 78 克
- 芥末 9 克
- 油 8 克
- 波多贝罗蘑菇 226 g
- 迷迭香 3 g

可持续汉堡2

- 培根 19 g
- 牛肉 70 g
- 汉堡面包 42 g
- 切达奶酪 39 克
- 番茄酱 17 克
- 蘑菇 14 克
- 洋葱 9 克
- 黑胡椒 1 克

营养汉堡

- 豆类 71 g
- 汉堡面包 50 g
- 香菜 2 g
- 玉米粉 4 g
- 孜然 0.2 g
- 鸡蛋 10 g
- 墨西哥辣椒 2 克
- 燕麦 15 克
- 油 4 克
- 洋葱 2 克
- 牛至 0.05 克
- 黑胡椒 0.2 克

指南

1. 上一节中的清单列出了制作每种汉堡应使用的食材用量。厨师应使用这些食材和用量设计食谱，不得添加或删去任何食材。2. 食材清单只列出了食材，所有加工步骤均由厨师决定。3. 所有食材名称均以单数形式列出，例如用 oat 而不是 oats、用 onion 而不是 onions 等。这只是为了方便。4. 厨师应决定食材清单中的哪些部分用于肉饼，哪些部分铺在肉饼上。这可能包括将某些食材分成两组或更多组。例如，如果食谱中提到洋葱，厨师可以将一部分切碎后加入肉饼，再将另一部分切片后铺在肉饼上。在这种情况下，请厨师记录各部分的重量。5. 每份食谱所列的用量均经过调整，使每份食材清单的总热量达到 500 卡路里。厨师可以自行调整用量，使这些食材改为制作一个汉堡，只要食材比例保持不变即可。例如，如果厨师认为某个汉堡所列的食材足以制作两个汉堡，可以将每种食材的用量都减半。6. 这里的重量通常指干燥未烹饪重量。如果厨师认为某种食材所列的重量不适合作为其干燥未烹饪重量，例如豆类，请联系我们以进一步确认。7. 如果香草或香料的用量极少，厨师可以将其理解为非常少的量。例如，0.05 克牛至可以理解为几片牛至叶。

8. 盐：食材清单中没有明确列出盐。厨师应根据口味调整盐量，并记录实际用量。9. 所列食材均用于制作汉堡，但并非所有食材都必须留在汉堡上。例如：

- 蔬菜和水果可以先称重，之后再去除果皮和茎。
- 如果食谱列出了油，而且厨师认为油最适合用于油炸或煎炒，那么煎炒后油可以留在平底锅中
- 有些香草（如迷迭香）可用于为其余食材增添香气，之后再丢弃。

10. 有些食谱可能会做出馅料溢出的汉堡，这是可以接受的，但仍请厨师尽力将汉堡制作得外观得体，并保持结构完整性。

11. 食材清单不会注明食材的具体品种。厨师可以自行决定。示例：

- 牛肉可以理解为整块肉或碎牛肉，具体取决于厨师认为哪种最适合该食谱。
- 蔬菜可以是厨师喜欢的任何品种，例如红洋葱、白洋葱等，但不能是青葱，因为青葱是另一种产品。

12. 其他奶酪包括不属于以下类别的任何奶酪：蓝纹奶酪、切达奶酪、茅屋奶酪、奶油奶酪、菲达奶酪、山羊奶酪、豪达奶酪、蒙特雷奶酪、蒙特雷杰克奶酪、马苏里拉奶酪、帕尔马干酪、普罗卧干酪。我们请厨师选择他们认为最适合其余食材的奶酪。13. 其他酱汁包括不属于以下类别的任何酱汁：烧烤酱、水牛城辣酱、奶油、咖喱酱、海鲜酱、蜂蜜、辣椒酱、番茄酱、芥末、青酱、牧场沙拉酱、调味酱、莎莎酱、酱油、牛排酱、照烧酱、醋、芥末、伍斯特酱。我们请厨师选择他们认为最适合其余食材的酱汁。14. 在可持续汉堡2中，我们请厨师将肉类和蘑菇混合在一起制作肉饼。这样一来，这款汉堡就更接近市面上可以买到的牛肉蘑菇混合物。15. 我们诚挚地请各位厨师发挥创意和艺术才能，利用这些食材清单制作出尽可能美味的汉堡，并衷心感谢他们对本研究的帮助！

补充文件 4
生成式人工智能打造美味、可持续且营养的汉堡

Vahidullah Tac[1*]、ChristopherGardner[2] 和Ellen Kuhl[1]

[1]斯坦福大学机械工程系，美国斯坦福。[2]斯坦福大学医学院预防研究中心，美国斯坦福。

*通讯作者。电子邮箱：vtac@stanford.edu；参与作者：cgardner@stanford.edu；ekuhl@stanford.edu；

含食材与制作说明的食谱

本文档由一位行政主厨编写。其中包含人工智能生成的配料清单、用量，以及主厨设计的组装汉堡的加工步骤和照片。为便于比较，我们还列入了巨无霸的配料清单®，尽管我们为调查购买了巨无霸，以提供最真实的体验。

巨无霸®

配料表

- 81克碎牛肉
- 74克汉堡面包
- 32克蛋黄酱
- 27克腌黄瓜
- 27克洋葱
- 21克生菜
- 18克奶酪

购买的产品

图 1 展示了巨无霸®，作为感官调查中的基准对比。巨无霸®并非由厨师制作，而是购买而来，以提供最真实的感官体验。

美味汉堡1

配料表

- 151克碎牛肉（80/20）
- 30 克布里欧修面包，修去大量部分以达到重量要求
- 7克番茄酱
- 7 克生菜，小宝石生菜



图1 巨无霸®。巨无霸®是经典的牛肉汉堡，包含碎牛肉、蛋黄酱、腌黄瓜、洋葱、生菜和奶酪等七种主要食材，夹在汉堡面包中食用。

- 19克蛋黄酱
- 47 克黄洋葱，切成圈
- 12 克美式奶酪，切片
- 75 克莳萝腌黄瓜，切片
- 79 克番茄，切片
- 1 克犹太洁食盐

制作说明

图2 展示了按照以下说明制作的美味汉堡1：使用一个比汉堡面包稍大的环形模具，将混合物塑形成肉饼，并在模具内压成厚薄均匀的一层。在肉饼两面撒上盐。用中火预热不粘煎锅。在干燥的平底锅中轻轻烘烤汉堡面包，直至呈金黄色，然后取出备用。将肉饼放入热锅中，每面煎3.5至4分钟，并在最后一分钟加入奶酪使其融化。在一个小碗中，将蛋黄酱和番茄酱混合均匀，然后将酱汁均匀涂抹在汉堡面包的两面。将生菜、番茄、腌黄瓜和洋葱仔细叠放在肉饼上，组装成汉堡。



图2 美味汉堡1。美味汉堡 1 是一款经典牛肉汉堡，以九种食材提升美味度：牛肉、番茄、腌黄瓜、洋葱、蛋黄酱、奶酪、生菜和番茄酱，并夹在汉堡面包中。

美味汉堡2

配料表

- 62克碎牛肉（80/20）
- 1 克红糖
- 35.5 克汉堡面包
- 0.2 蒜瓣，磨成细末
- 3.5 克黄洋葱，刨成薄片
- 8 克格吕耶尔奶酪，切片
- 12.5 克瑞姆拉德酱
- 0.15 克新鲜百里香叶，切碎
- 28 克番茄片
- 2克伍斯特酱
- 0.5 克犹太洁食盐

制作说明

图 3 展示了按以下说明制作的美味汉堡 2：在碗中，将牛肉、大蒜、百里香、红糖和伍斯特酱充分混合。用一个比汉堡面包稍大的环形模具，将混合物压入模具内，塑成厚薄均匀的肉饼。在中高火上预热不粘锅。在干燥的锅中将汉堡面包稍微烤至金黄色，然后取出备用。用盐给肉饼调味。将洋葱片铺在肉饼上，然后将肉饼放入热锅中，洋葱一面朝下。煎2分钟后翻面，加入奶酪，再煎2分钟。盖上锅盖，帮助奶酪融化。组装时，将番茄片放在底部汉堡面包上，再放上肉饼。将雷穆拉德酱抹在顶部汉堡面包上，然后合上汉堡



图3 美味汉堡2。 美味汉堡 2 是一款创新牛肉汉堡，以十种食材调配，旨在提升美味度：牛肉、番茄、雷穆拉德酱、Gruy`ere 奶酪、洋葱、伍斯特酱、红糖、大蒜和百里香叶，夹在汉堡面包中享用。

可持续汉堡1

配料表

- 21克芝麻菜

- 46 克布里欧修面包（修整至所需重量）
- 2 克新鲜大蒜，用细磨器磨碎
- 78 克蛋黄酱，分开使用（Best Foods）
- 9 克棕色辣芥末
- 8 克特级初榨橄榄油
- 226 克波多贝罗蘑菇（去除茎和菌褶）
- 3克迷迭香，切碎
- 0.5 克犹太洁食盐

制作说明

图 4 展示了按照以下说明制作的可持续汉堡1：将波多贝罗蘑菇切成横跨菌盖、厚度为 0.5 英寸的片。在碗中混合大蒜、芥末、迷迭香、盐和一半的蛋黄酱。加入蘑菇，轻轻拌匀，使其均匀裹上调料。腌制 20 分钟。用中高火预热不粘锅。在干燥的锅中轻轻烘烤汉堡面包，直至呈金黄色，然后取出备用。将橄榄油加入同一口锅中。油热后，加入腌好的蘑菇片，每面煎约 2.5 分钟，直至深度焦黄。将蘑菇转移到纸巾上，吸去多余的油。在汉堡面包的两面涂上剩余的蛋黄酱。将蘑菇铺在下半片面包上，再放上新鲜的芝麻菜。



图4 可持续汉堡1。 可持续汉堡1是一款以可持续性为优化目标的蘑菇汉堡，包含八种食材：波多贝罗蘑菇、蛋黄酱、芝麻菜、芥末、油、迷迭香和大蒜，夹在汉堡面包中食用。

可持续汉堡2

配料表

- 19 克培根，切成两片
- 70 克碎牛肉（80/20）
- 42 克布里欧修面包，稍微修整以达到所需重量
- 39 克切达奶酪，切片
- 17克番茄酱
- 14 克切碎的褐蘑菇
- 9 克刨成薄片的黄洋葱
- 1 克研磨黑胡椒

- 0.8 克犹太洁食盐

制作说明

图 5 展示了按照以下说明制作的可持续汉堡2：将切碎蘑菇和碎牛肉混合，反复翻拌至均匀。将混合物捏成球状，然后在盐和胡椒中滚动，使其均匀裹上调味料。用中火预热不粘煎锅。将汉堡面包放入干燥的平底锅中稍微烘烤至金黄，然后取出备用。将火调至中小火，加入培根，煎至脂肪析出且培根变脆。将培根转移到铺有纸巾的盘子上沥油，同时将析出脂肪留在锅中。将火调至中高火。把洋葱片放在平底锅中央，然后将调味好的肉球放在上面。盖上一张烘焙纸或蜡纸，用锅铲用力压扁肉饼，直到其略宽于汉堡面包。煎1.5分钟，然后翻面，放上奶酪，再煎1.5分钟，或直到肉饼熟透且奶酪融化。在烤好的汉堡面包两面涂上番茄酱。将汉堡肉饼放在底部那半片面包上，铺上培根，再盖上另一半面包。



图 5 可持续汉堡2。 可持续汉堡2是一种由八种食材制成的蘑菇牛肉混合物，包括碎牛肉、切达奶酪、培根、番茄酱、褐蘑菇、洋葱和胡椒，搭配汉堡面包食用。

营养汉堡

配料表

- 71 克罐装芸豆（沥干并冲洗）
- 50 克布里欧修面包，轻微烤过
- 2 克香菜（叶片和嫩茎）
- 4克玉米粉
- 0.2克孜然，磨碎
- 10 克鸡蛋（打散一整颗鸡蛋并量取10克）
- 2 克墨西哥辣椒jalape˜no，切成薄圆片
- 15克压片燕麦
- 4克特级初榨橄榄油
- 2 克白洋葱，切碎
- 0.05 克干墨西哥牛至
- 0.2 克研磨黑胡椒

- 0.8 克犹太洁食盐，分为两份，一份加入肉饼混合料，另一份用于烹饪前给成型的肉饼调味

制作说明

图 6 展示了按以下说明制作的营养汉堡：用叉子轻轻压碎豆类，保留一些颗粒以保留质地。拌入洋葱、孜然、一半的盐、黑胡椒和牛至。加入燕麦和玉米粉，搅拌至均匀混合。最后，加入鸡蛋并充分搅拌，形成黏稠均匀的混合物。使用与汉堡面包大小相同的环形模具，在一张烘焙纸上将混合物塑形成肉饼。冷藏30分钟，让淀粉吸水，使肉饼定型后再烹煮。用中高火预热不粘锅，并在干锅中轻轻烘烤汉堡面包。汉堡面包烤好后，将橄榄油倒入锅中。用剩余的盐为冷藏后的肉饼调味，每面煎约3.5分钟，直到表面呈金黄色且内部热透。组装汉堡：将煎好的肉饼放在烤好的汉堡面包上，再撒上剩余的香菜和墨西哥辣椒片。



图 6 营养汉堡。 营养汉堡是一种经过优化、营养丰富的豆类汉堡，含有十二种食材：芸豆、压片燕麦、鸡蛋、玉米粉、油、香菜、jalape˜ño、洋葱、孜然、胡椒和牛至，夹在汉堡面包中食用。