**Table 1**.

*Dictionary Summary of C-MFD 2.0*

|  |  |  |
| --- | --- | --- |
| **Moral Foundations** | **Number of words** | **Examples** |
| care/harm | 1,477 | 同情, 恐吓, 辱骂, 情逾骨肉, 庇佑, 抚慰, 残民害物, 捕杀, 害死, 恩情 |
| authority/subversion | 1,415 | 级别, 越权, 统治, 品第, 提干, 将帅, 头目, 号令, 官学, 势族 |
| loyalty/betrayal | 698 | 公忠体国, 叛变, 绝情寡义, 殊荣, 奸险, 信守承诺, 背叛, 人民公敌, 种族主义, 投靠 |
| fairness/cheating | 1,023 | 形式主义, 惩罚, 童叟无欺, 委屈, 偏心眼，信义, 讹诈, 真心实意, 言而无信, 正大光明 |
| sanctity/degradation | 894 | 殉死, 湮没, 性服务, 邪荡, 病人, 功德无量, 虔诚, 浊世, 如蝇逐臭, 吸毒 |
| liberty/oppression | 50 | 独立自主, 解放, 自由, 恶霸, 五四运动, 阶级斗争, 剥削, 努力, 封建礼教, 压迫 |
| waste/efficiency | 83 | 勤俭节约, 铺张浪费, 节能, 高效, 以逸待劳, 白费, 糟蹋, 节水, 骄奢淫逸, 滥伐 |
| altruism/selfishness | 85 | 冲锋陷阵, 勇敢, 义勇, 无私, 利他主义, 舍己, 舍身取义, 侠肝义胆, 勇于献身, 抱头鼠窜 |
| diligence/laziness | 41 | 好吃懒做, 打拼, 拼命, 自强, 好逸恶劳, 将勤补绌, 熟能生巧, 敬业, 勤奋, 自律 |
| resilience/weakness | 63 | 坚忍不拔, 恒心, 曲意俯就, 毅力, 秉持信念, 忍让, 坚韧, 坚毅顽强, 果敢坚毅, 坚定不移 |
| modesty/arrogance | 75 | 谦逊, 目空一切, 傲慢, 谦虚有礼, 恭让, 不识抬举, 收敛锋芒, 颐指气使, 诚恳, 有恃无恐 |
| general | 234 | 八荣八耻, 礼乐崩坏, 美德, 卑劣, 民族大义, 公序良俗, 真善美, 缺德, 仁义, 安定和谐 |
| Total | 6,138 |  |

**Table 2**.

*Word Density Analysis of C-MFD 2.0 on the Crowd-sourced Texts*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | **Accuracy** | **Precision** | **Recall** | **F1 score** | **Coverage** |
| Five Foundation Model | |  |  |  |  |  |
|  | *care/harm* | 0.88 | 0.73 | 0.42 | 0.53 | 0.84 |
|  | *authority/subversion* | 0.90 | 0.79 | 0.46 | 0.58 | 0.92 |
|  | *loyalty/betrayal* | 0.93 | 0.66 | 0.62 | 0.64 | 0.90 |
|  | *fairness/cheating* | 0.94 | 0.60 | 0.71 | 0.65 | 0.85 |
|  | *sanctity/degradation* | 0.95 | 0.62 | 0.76 | 0.69 | 0.88 |
| Foundation Candidates | |  |  |  |  |  |
|  | *liberty/oppression* | 0.95 | 0.50 | 0.97 | 0.66 | 0.75 |
|  | *waste/efficiency* | 0.95 | 0.53 | 0.92 | 0.68 | 0.74 |
|  | *altruism/selfishness* | 0.93 | 0.29 | 0.86 | 0.44 | 0.82 |
|  | *diligence/laziness* | 0.94 | 0.32 | 0.97 | 0.48 | 0.50 |
|  | *resilience/weakness* | 0.93 | 0.31 | 0.93 | 0.46 | 0.78 |
|  | *modesty/arrogance* | 0.94 | 0.39 | 0.94 | 0.55 | 0.74 |
| Average | |  |  |  |  |  |
|  | *Five Foundation Model* | 0.92 | 0.68 | 0.59 | 0.62 | 0.88 |
|  | *Foundation Candidate* | 0.94 | 0.39 | 0.93 | 0.54 | 0.72 |

*Note*: Accuracy is the ratio of true predictions (including true positive and true negative) over all predictions. Precision is the ratio of true positives among all positives. Recall measures the ratio of true positives over the sum of true positives and false negatives. The F1 score measures the harmonic mean of the precision and recall, providing a weighted figure to evaluate the dictionary performance. And coverage measures how many labelled texts in the pool are successfully detected by the dictionary.

**Table 3**.

*Benchmark the performance of C-MFD 2.0 against C-MFD 1.0 (Wu et al., 2019)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Accuracy | | Coverage | | F1 Score | |
|  | C-MFD 1.0 | C-MFD 2.0 | C-MFD 1.0 | C-MFD 2.0 | C-MFD 1.0 | C-MFD 2.0 |
| care/harm | 0.76 | 0.88 | 0.69 | 0.84 | 0.43 | 0.53 |
| authority/subversion | 0.80 | 0.90 | 0.79 | 0.92 | 0.54 | 0.58 |
| loyalty/betrayal | 0.86 | 0.93 | 0.77 | 0.90 | 0.59 | 0.64 |
| fairness/cheating | 0.87 | 0.94 | 0.67 | 0.85 | 0.56 | 0.65 |
| sanctity/degradation | 0.89 | 0.95 | 0.67 | 0.88 | 0.61 | 0.69 |
| Average | 0.84 | 0.92 | 0.72 | 0.88 | 0.55 | 0.62 |

*Note*:

1. Accuracy means the ratio of one moral foundation (e.g., care) being correctly identified (including both true-positive and true-negative) to all crowd-generated texts. Higher accuracy means the dictionary is able to detect both care and not-care texts. For example, .88 means if a text is (not) categorized as "care", the decision is 88% likely to be correct.
2. Coverage refers to how many moral-relevant texts can be identified by C-MFD 2.0. For example,.88 coverage means that if a text expresses moral values, there is 88% chance that our dictionary can detect it from the word use.
3. F1 score measures the harmonic mean of the precision and recall (see Table 2 for their definitions), providing a weighted figure to evaluate the dictionary performance.
4. Wu, S., Yang, C., & Zhang, Y. (2019). The Chinese Version of Moral Foundations Dictionary: A Brief Introduction and Pilot Analysis. ChinaXiv:201911.00002. (ChinaXiv:201911.00002).