



## Gestão de Dados de Políticas Públicas: o caso da avaliação de cardápios da Alimentação Escolar usando mineração de textos

Daniel Osaku<sup>1</sup>; Rafaella Guimarães Moraes Camargo<sup>2</sup>; Isabela Prado Martins<sup>5</sup>; Evandro Marcos Saidel Ribeiro<sup>3</sup>; Alexandre Cláudio Botazzo Delbem<sup>1</sup>; Semiramis Martins Álvares Domene<sup>4</sup>; Ricardo Marcondes Marcacini<sup>1</sup>; Solange Oliveira Rezende<sup>1</sup>

<sup>1</sup>Instituto de Ciências Matemáticas e de Computação – Universidade de São Paulo (USP) – São Carlos, SP – Brazil.

<sup>2</sup>Programa de Pós Graduação Interdisciplinar em Ciências da Saúde – Universidade Federal de São Paulo (UNIFESP) – Santos, SP – Brazil

<sup>3</sup>Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo (USP) – Ribeirão Preto, SP – Brazil

<sup>4</sup>Departamento de Políticas Públicas e Saúde Coletiva, Instituto Saúde e Sociedade, Universidade Federal de São Paulo – Campus Baixada Santista, SP – Brazil

<sup>5</sup>Programa de Pós-Graduação em Nutrição – Universidade Federal de São Paulo (UNIFESP) – São Paulo, SP – Brazil

{danosaku, esaidel}@usp.br,  
{guimaraes.rafaella, isabela.prado, semiramis.domene}@unifesp.br,  
{acbd, ricardo.marcacini, solange}@icmc.usp.br

**Abstract.** *The Brazilian School Feeding Program contributes to the improvement and development of schools through the transfer of funds. However, there is no feedback to monitor the menu quality to guide the actions of the program to promote food security and nutrition. In this paper, the construction of a classification method for generating evaluation components from school menu data is presented. All evaluation components were obtained with negligible error rates and in a viable time.*

**Resumo.** *O Programa Nacional de Alimentação Escolar contribui para o aperfeiçoamento e desenvolvimento dos escolares por meio da transferência de fundos. No entanto, não existe apreciação para monitorar a qualidade dos cardápios para guiar as ações do programa para promover a segurança alimentar e nutricional. Neste artigo, está apresentada a construção de um método de classificação para geração dos componentes de avaliação a partir de dados de*

*cardápios escolares. Foram obtidos todos os componentes de avaliação com insignificante taxa de erro e em tempo viável.*

## **1. Introdução**

As políticas públicas estão relacionadas à atuação do governo que, a fim de alcançar objetivos pré-definidos, adota ações e estratégias a partir das quais se desenvolvem as inter-relações entre Estado, economia e sociedade [Nadal et al. 2021], também podem ser definidas como um campo dentro da política que analisa o governo à luz de grandes questões públicas [Mead 1995].

Considerado uma inovação em políticas públicas, o Programa Nacional de Alimentação Escolar (PNAE) prevê que as refeições ofertadas devem estar associadas aos hábitos alimentares culturais e saudáveis dos estudantes, assim como atreladas às iniciativas de compras sustentáveis, articuladas com o fortalecimento da agricultura familiar, contribuindo assim para a Segurança Alimentar e Nutricional (SAN) dos escolares [Kroth et al. 2020]. Para isso, o PNAE visa contribuir para o crescimento, o desenvolvimento, a aprendizagem, o rendimento escolar dos estudantes e a formação de hábitos alimentares saudáveis, por meio de ações de educação alimentar e nutricional e pela oferta de refeições [Brasil 2009]. No entanto, tais ações são realizadas sob demanda sem que haja um acompanhamento efetivo como, por exemplo, um instrumento de avaliação de cardápios escolares que possa contribuir para o monitoramento do PNAE e, consequentemente, melhorar a qualidade das refeições.

[Camargo et al. 2021] propuseram o Índice de Qualidade para Cardápios da Alimentação Escolar (IQCAE) para avaliar a qualidade das refeições escolares em âmbito nacional, com base na legislação vigente, de modo a contribuir para o monitoramento do PNAE [Brasil 2009]. No entanto, a avaliação é realizada de forma manual, que pode ser afetada por diversos fatores humanos como fadiga, cansaço e falta de atenção.

Dessa maneira, o uso de técnicas de mineração de textos pode contribuir para a geração automática de componentes necessários para realizar a avaliação dos cardápios. Essa técnica pode ser vista como uma especialização da mineração de dados, sendo que a diferença entre elas está na natureza dos dados, pois enquanto a mineração de dados lida com dados estruturados, a mineração de textos lida com dados não estruturados [Sinoara et al. 2021]. Também é definida como um conjunto de técnicas e processos utilizados em diversas áreas como Inteligência Artificial, Aprendizado de Máquina, Base de Dados e Estatística, para descoberta de conhecimento inovador a partir de dados textuais [Rezende 2003].

Assim, este trabalho visa descrever a construção do método de classificação de dados textuais de cardápios por meio do uso de técnicas de mineração de textos para a concepção dos componentes de avaliação e composição do cardápio escolar.

O estrutura do artigo está organizada da seguinte forma: A seção 2 descreve a visão geral da metodologia aplicada no presente estudo. A Seção 3 apresenta e discute os resultados obtidos. A Seção 4 apresenta considerações finais.

## **2. Material e métodos**

Esta seção visa descrever a base de dados empregada e os métodos de mineração de textos aplicados no presente estudo.

## 2.1. Aquisição dos dados

A base de dados é composta por dados secundários 2125 cardápios diários que correspondem a 425 cardápios semanais de municípios de todas as regiões do país<sup>1</sup>. Os cardápios foram avaliados com base em 14 componentes (ver Quadro 1), conforme proposto por [Camargo et al. 2021].

**Quadro 1. Componentes de avaliação dos cardápios empregados no método IQCAE-R. Fonte: (Brazil, 2023)**

ID	Componente de avaliação	ID	Componente de avaliação
CT	Cereais e tubérculos	AU	Alimentos ultraprocessados
L	Leguminosas	CV	Carnes vermelhas
LV	Legumes e verduras	PRD	Preparações regionais doces
F	Frutas	AP	Alimentos processados
AOP	Aves, ovos e peixes	AR	Alimentos regionais
LD	Leites e derivados	AS	Alimentos da socio-biodiversidade
DPD	Doces e preparações doces	CH	Compatibilidade de horário

## 2.2. Mineração de textos

O processo de mineração de textos foi realizado de acordo com as seguintes etapas [Rezende 2003]: i) identificação do problema; ii) pré-processamento; iii) extração de padrões; iv) pós-processamento; e v) utilização do conhecimento.

Para a etapa de identificação do problema, os cardápios foram analisados para verificar quais dados eram necessários para gerar cada componente de avaliação. A maioria dos componentes depende apenas da descrição do cardápio. No entanto, o componente AR depende de dados da sua região de origem, assim como o AS depende da região e do estado brasileiro para poder ser classificado corretamente. Por fim, o componente CH depende de classificação prévia quanto ao tipo de refeição a qual o cardápio se refere e do horário que a refeição foi ofertada. Dessa maneira, diferentes componentes de avaliação requerem diferentes modelos, uma vez que os dados de entrada de alguns componentes necessitam de dados referentes à estado, região ou horário para que sejam adequadamente modelados. Quanto à saída do modelo, uma vez que os cardápios podem estar presentes em diferentes componentes de avaliação, o processo de classificação pode ser realizado utilizando um modelo com saída multi-rótulos ou treinar um modelo para cada componente de avaliação, realizando classificação binária.

Na etapa de pré-processamento, foram utilizadas técnicas de processamento de linguagem natural como, remoção de palavras irrelevantes (*stopwords*) e radicalização para diminuir a variação entre as palavras (*stemming*). Em seguida, foi realizada a vetorização dos dados (*word embedding*) utilizando o método *Bag of Words* (BoW), que efetua a contagem dos n-gramas e gera uma representação dos dados em forma de atributo-valor.

Na etapa de extração de padrões, diferentes abordagens de classificação poderiam ter sido empregadas, que vão desde métodos tradicionais como Naïve Bayes [Koch 2006]

<sup>1</sup>Índice de Qualidade para Cardápios da Alimentação Escolar: Revisão e Validação. N° Certificado de Apresentação de Apreciação Ética (CAAE): 18020619.9.0000.5505.

à métodos que modelam o contexto das palavras como o método Representações de Codificador Bidirecional de Transformadores (*Bidirectional Encoder Representations from Transformers* - BERT) [Devlin et al. 2019]. Uma vez que os cardápios são formados por preparações que não possuem relações entre si, não foram empregados modelos que exploram o contexto das palavras, muito embora sejam os modelos estado-da-arte em mineração de textos. Dessa maneira, foram empregados os métodos de aprendizado de máquina baseados em Redes Neurais Artificiais (RNA) *Multi-Layer Perceptron* (MLP) [Haykin 1994] e Máquinas de Vetores de Suporte (*Support Vector Machines* - SVM) [Cortes and Vapnik 1995] para realizar a geração de componentes de avaliação.

Uma rede MLP é uma RNA que simula o funcionamento de um neurônio, na qual os sinais de entrada são transformados e propagados para os neurônios seguintes [Rodrigues 2019], sendo composta por uma camada de entrada, uma ou mais camadas intermediárias e uma camada de saída. Já o classificador SVM tenta separar as classes pela definição de hiperplanos, que mapeia o espaço de entrada em um espaço de mais alta dimensão [Rezende 2003], de forma que o problema se torne linearmente separável em um espaço de características de mais alta dimensão, na qual a margem de separação é controlada pelo usuário por meio de ajuste de parâmetros.

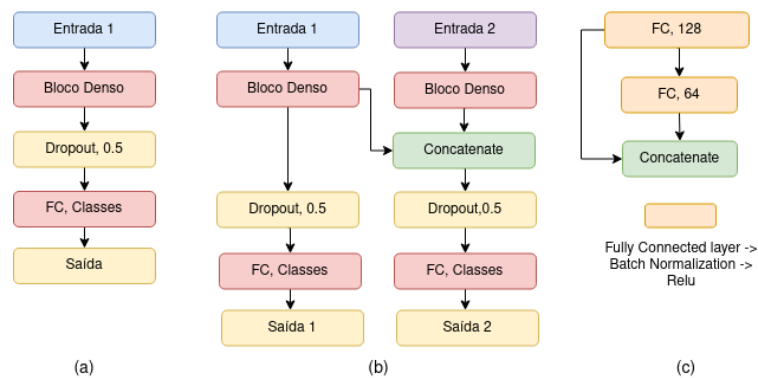
A Figura 1a apresenta a arquitetura da rede MLP utilizada para efetuar a classificação utilizando a descrição ou sentenças dos cardápios com o objetivo de obter os componentes de avaliação, exceto o componente de CH.

Para obter AR, foram gerados os vetores de atributos da descrição ou sentenças e dos dados da região separadamente. Em seguida, os vetores foram concatenados para serem usados como entrada. O mesmo procedimento foi utilizado para obter o componente de avaliação AS, contudo, gerando o vetor de atributos relativo aos dados do estado.

Já a Figura 1b apresenta a arquitetura utilizada para a classificação do componente de avaliação CH. Neste caso, foram utilizadas como entradas a descrição ou sentenças dos cardápios (Entrada 1) e o horário das refeições (Entrada 2), onde as camadas intermediárias são concatenadas para realizar simultaneamente a meta-classificação do tipo de refeição (saída 1) e a classificação do componente de avaliação CH (saída 2).

Além disso, a Figura 1c apresenta a sequência de operações realizadas no bloco denso, que consiste em duas camadas *Fully Connected* com 128 e 64 neurônios escolhidos empiricamente, seguidos por aplicação de normalização e de função de ativação *ReLU*. Suas ativações são concatenadas para gerar o vetor de características, o qual foi utilizado para classificar os cardápios em diferentes componentes de avaliação. Além disso, foi utilizada camada de *Dropout* e função de ativação sigmoide para gerar a saída da rede neural com valores entre 0 e 1.

A avaliação dos métodos de classificação foi realizada de forma quantitativa por meio das métricas de acurácia e coeficiente de *Cohen Kappa* [Cohen 1960]. Enquanto a primeira fornece uma medida geral de desempenho do modelo, a segunda indica a capacidade do modelo de identificar exemplos da classe minoritária, penalizando modelos que não são capazes de reconhecê-la. Uma vez que a distribuição entre o número de ausentes e presentes é altamente heterogênea em grande parte dos componentes de avaliação, é importante utilizar ambas métricas para avaliar os modelos.



**Figura 1. Arquiteturas utilizadas para a classificação dos componentes de avaliação em (a), exceto o componente de avaliação CH, que utiliza a arquitetura apresentada em (b) e a sequência de operações do bloco denso em (c). Fonte: (Brazil, 2023).**

### 3. Resultado experimentais

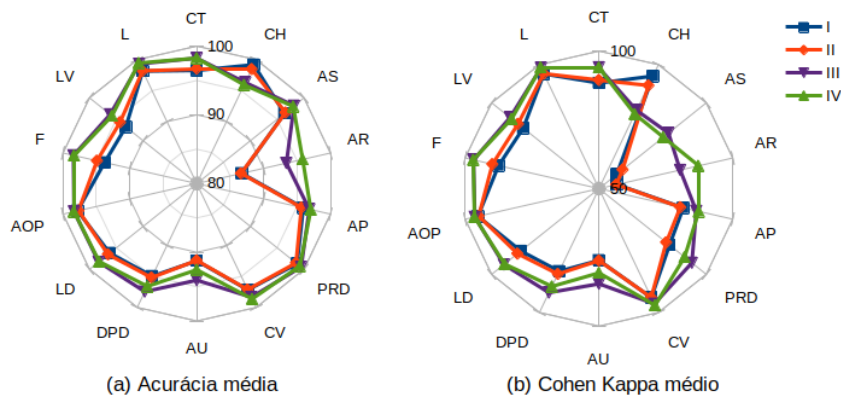
Os resultados experimentais da classificação dos cardápios estão apresentados segundo a classificação usando os métodos MLP e SVM. Para realizar a classificação, utilizou-se a descrição (D) ou sentença (S) como dados de entrada e foram gerados os componentes de avaliação com classificação multi-rótulos (M) ou binária (B). Dessa maneira, foram avaliadas as combinações D-M em (I), D-B em (II), S-M em (III) e S-B em (IV).

Os experimentos foram realizados com a validação cruzada, que utilizou a base de dados dividida em conjunto de treinamento e de teste, respectivamente, na proporção de 80% e 20%, que foram escolhidas empiricamente. Os resultados apresentados correspondem à média dos resultados de 5 rodadas. Foram selecionados unigramas, bigramas e trigramas, isto é, sequência de 1, 2 e 3 palavras, respectivamente, extraídos dos dados textuais para a geração dos vetores de características do texto de entrada.

#### 3.1. Geração dos componentes de avaliação segundo MLP

Os resultados experimentais com a utilização de uma rede MLP foram obtidos com as seguintes configurações definidas empiricamente: 50 épocas; *mini-batch* de 16; taxa de aprendizado de  $10^{-3}$ ; otimizador Adam; e métrica de perda erro quadrático médio. O modelo escolhido foi o que obteve menor perda durante o treinamento.

As Figuras 2a e 2b apresentam, respectivamente, as acurácias e os coeficientes de *Cohen Kappa* médios obtidas pela classificação dos dados dos cardápios nos diferentes componentes de avaliação. Foi possível observar que o componente CH apresentou melhores resultados com a utilização da descrição do cardápio, evidenciando a necessidade de sua utilização para determinar o tipo de refeição, sub-tarefa necessária para definir tal componente. Por outro lado, o componente AR obteve melhor resultado com a utilização das sentenças do cardápio, uma vez que nem todos os alimentos do cardápio são regionais e que podem atuar como uma espécie de ruído aos dados de entrada. Em geral, os melhores resultados foram obtidos utilizando como entrada as sentenças dos cardápios, sendo facilmente observado quando são analisados os resultados em ambas métricas. Isso demonstra que as preparações que compõem o cardápio possuem pouca relação uns com os outros, conforme foi identificado durante a etapa de identificação do problema.



**Figura 2. Avaliação dos resultados do segundo MLP usando as métricas: (a) Acurácia média e (b) Cohen Kappa médio. Fonte: (Brazil, 2023).**

Com relação ao tipo de classificação, não houve diferença estatística significativa entre classificação multi-rótulos e binária, como pode ser visto nos resultados de classificação dos componentes de avaliação CT, L, F e AOP. As maiores diferenças entre as abordagens ocorreram nos componentes de avaliação DPD, AU, PRD, AR e AS, tendo melhores resultados com classificação multi-rótulos, o que indica que possa existir dependência ou relação entre os componentes de avaliação nestes casos.

Uma vantagem de efetuar classificação de sentenças é a possibilidade de identificar quais preparações estão presentes nos componentes de avaliação, uma vez que cada preparação é classificada individualmente e, dessa forma, pode contribuir para a explicabilidade do modelo.

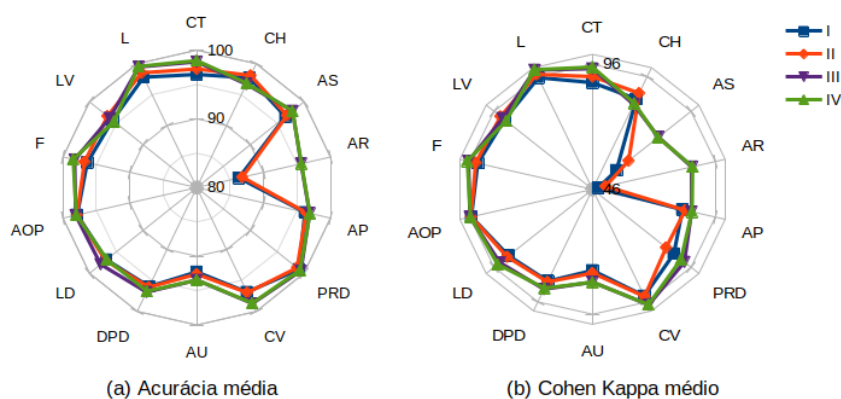
### 3.2. Geração dos componentes de avaliação segundo SVM

Os resultados de classificação dos dados dos cardápios segundo SVM foram obtidos a partir da versão *one-vs-rest* com *kernel* RBF e aplicado aos mesmos conjuntos de treinamento e de teste do experimento apresentado anteriormente. Os parâmetros  $C$  e  $\gamma$  foram otimizados com a utilização do método de busca em grade com  $C = [0.1, 1, 10, 100]$  e  $\gamma = [1, 0.1, 0.01, 0.001]$ .

Os dados de entrada do classificador são os *word embeddings* gerados pelo método BoW, referentes à descrição/sentenças dos cardápios. Para classificação do componente de avaliação AR, os *word embeddings* da descrição do cardápio foram concatenados com os *word embeddings* gerados a partir das regiões dos estados. O mesmo procedimento foi aplicado aos componentes de avaliação AS e CH, concatenando os *word embeddings* dos estados e do horário das refeições, respectivamente.

Os resultados de classificação em termos de acurácia média e *Cohen Kappa* médio estão apresentados nas Figuras 3a e 3b, respectivamente. Novamente, pode-se observar que a separação da descrição dos cardápios em sentenças geraram melhores resultados, o que demonstra a independência entre as sentenças considerando os diferentes componentes de avaliação. Embora os resultados da abordagem IV tenham sido estatisticamente similares aos obtidos com MLP, o classificador SVM foi mais susceptível aos problemas de maior complexidade, isto é, àqueles que necessitam de modelos mais complexos, como pode ser vista na comparação dos resultados obtidos nos componentes de avaliação AU,

AP, AR e AS, que estão apresentados nas Figuras 2b e 3b, demonstrando que o modelo baseado em MLP teve maior capacidade de reconhecer a classe minoritária em problemas mais complexos.



**Figura 3. Avaliação dos resultados do classificador SVM usando as métricas: (a) Acurácia média e (b) Cohen Kappa médio. Fonte: (Brazil, 2023).**

A utilização da descrição dos cardápios para efetuar a classificação (abordagens I e II), apresentou desempenho inferior quando comparado aos modelos treinados com as sentenças extraídas da descrição do cardápio, que reforça a importância da definição de um vocabulário/dicionário de alimentos e preparações de alimentos dos cardápios para auxiliar na anotação de cada sentença nos diferentes componentes de avaliação. Contudo, esta atividade também apresenta limitações, uma vez que um alimento pode ser classificado de maneira inadequada a depender do seu nível de processamento.

#### 4. Considerações finais

Neste artigo, foi retratada a construção do método de produção de componentes de avaliação a partir dos dados de cardápios escolares por meio de técnicas de mineração de textos. Diferentes abordagens foram avaliadas com o objetivo de buscar a configuração que maximiza a taxa de acerto no processo de classificação dos componentes de avaliação. Observou-se que os melhores resultados foram obtidos separando as preparações presentes na descrição dos cardápios em sentenças e classificando-as separadamente. Em termos de *Cohen Kappa* e acurácia médios, o modelo baseado em MLP e o classificador SVM obtiveram resultados similares entre si. Com acurácias acima de 93% com classificação multi-rótulos de sentenças, o método de geração de componentes de avaliação apresentou resultados promissores com relação à avaliação automática dos cardápios.

Além disso, a geração automática dos componentes de avaliação permitiu que a avaliação do cardápio fosse realizada de forma mais rápida, ampliando a capacidade de analisar mais dados de cardápios. Assim, é possível afirmar que o desempenho obtido com o uso de técnicas de mineração de textos se equipararam ao dos especialistas, no sentido de classificar os componentes de avaliação com taxa de erro admissível, mas com maior capacidade de avaliar grandes volumes de dados. Isso pode dar maior agilidade na geração de indicadores que possam auxiliar no aprimoramento e nas ações do PNAE, visando a melhoria da qualidade da alimentação escolar ofertada.

## 5. Agradecimentos

Este trabalho foi executado no Centro de Inteligência Artificial (C4AI-USP) com apoio da Fundação de Apoio à Pesquisa do Estado de São Paulo (processo FAPESP #2019/07665-4) e da IBM Corporation. Os autores também agradecem à Fundação de Apoio à Pesquisa do Estado de São Paulo (processo FAPESP #2019/25010-5), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq processo 309575/2021-4) e ao ICMC USP São Carlos pelo apoio financeiro e técnico, que contribuíram para que este trabalho tenha sido realizado.

## Referências

- Brasil, P. d. R. (2009). Lei nº 11.947, de 16 de junho de 2009. Dispõe sobre o atendimento da alimentação escolar e do Programa Dinheiro Direto na Escola aos alunos da educação básica.
- Camargo, R. G. M., Caivano, S. d. A., and Domene, S. M. Á. (2021). Avaliação qualitativa de cardápios escolares oferecidos em municípios brasileiros. *Ciência & Saúde Coletiva*, 26:2207–2213.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46.
- Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3):273–297.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 NAACL (Long and Short Papers)*, pages 4171–4186.
- Haykin, S. (1994). *Neural networks: a comprehensive foundation*. Prentice Hall PTR.
- Koch, K.-R. (2006). *Bayesian inference with geodetic applications*, volume 31. Springer, Germany.
- Kroth, D. C., Geremia, D. S., and Mussio, B. R. (2020). Programa nacional de alimentação escolar: uma política pública saudável. *Ciência & Saúde Coletiva*, 25:4065–4076.
- Mead, L. (1995). Public policy: Vision, potential, limits. *Policy Currents (Newsletter of the Public Policy Section, APSA)*, 68(3).
- Nadal, K., Kuasoski, M., Mascarenhas, L. P. G., Maganhotto, R. F., and Doliveira, S. L. D. (2021). Políticas públicas ambientais: uma revisão sistemática. *Revista Ibero-Americana de Ciências Ambientais*, 12(1):680–690.
- Rezende, S. O. (2003). *Sistemas inteligentes: fundamentos e aplicações*. Editora Manole Ltda.
- Rodrigues, W. G. (2019). Predição de diâmetros e cálculo de volume de clones de eucalipto: uma abordagem com redes multi layer perceptron e long-short term memory. Master's thesis, Universidade Federal de Goiás.
- Sinoara, R. A., Marcacini, R. M., and Rezende, S. O. (2021). Mineração de textos e semântica: desafios, abordagens e aplicações. *Revista de Sistemas de Informação da FSMA*, 27(ja-ju 2021):41–53.