

CONFINEMENT COMPANIONS: CAN AI ADDRESS THE SOLITARY CONFINEMENT CRISIS?

Francis X. Shen ^α

fxshen@umn.edu

Prepared for *We Robot 2018*
Stanford University | April 13, 2018

Note to *We Robot* Readers: I'd love to connect with developers/coders who would like to think through with me some of the challenging (!) technical details that need further discussion (either in this paper or more likely, a separate stand-alone piece, on which we could collaborate.)

FOR QUICK READING: If there is limited time for reading this draft, the reader might:

- *Read the Introduction* (summarizing the paper's arguments)
- *Skip/skim Part I* (discussing in detail the harmful effects of solitary confinement)
- *Skip/skim Part II* (discussing legal challenges and policy reforms related to solitary confinement)
- *Read Part III* (proposing *Themis*, the confinement companion)
- *Read Part IV* (discussing a series of questions and critiques about the proposal)

^α Associate Professor of Law & McKnight Presidential Fellow, University of Minnesota; Director, Shen Neurolaw Lab; Senior Fellow in Law and Neuroscience, Harvard Law School Petrie-Flom Center and Affiliated Faculty Member, MGH Center for Law, Brain, and Behavior; Executive Director of Education & Outreach, MacArthur Foundation Research Network on Law and Neuroscience. Contact: Walter F. Mondale Hall, 229-19th Avenue South, Minneapolis, MN 55455, 612-625-5328, fxshen@umn.edu, www.fxshen.com. *Acknowledgements:* Preparation of this article was supported in part by the University of Minnesota Law School. For helpful advice, I thank students in my University of Minnesota Law and Artificial Intelligence seminars. For excellent research assistance, I thank Sydney Diekmann, Chase Hamilton, Deanna Thompson, and Richard Yo. For research support, I thank David Zopf-Jordan and fellow staff at the University of Minnesota Law Library.

Confinement Companions:
Can AI Address the Solitary Confinement Crisis?

After only a short time in solitary, I felt all of my sense start to diminish. ... I stared obsessively at the bolts on the door to my cell. There was nothing to hear except empty, echoing voices from other parts of the prison. I was so lonely that I hallucinated words coming out of the wind. ... Who would know if something happened to me? The space I inhabited was invisible to the outside world, just like I was.” – Solitary inmate Five Mualimm-ak (2016)¹

Solitary confinement in United States prisons is widespread and brutal.² As the epigraph above suggests, the experience of solitary confinement is dehumanizing. Criticisms of the practice have become increasingly strong, and in 2016 then President Barack Obama announced a series of reforms to the use of the practice in the federal system.³ Yet, despite many critiques and some limited progress in legal reform, solitary confinement persists. In the United States, over 80,000 inmates experience solitary confinement each year.⁴

Given this reality, many academic disciplines are asking: what can we do to change this? Artificial Intelligence and Robotics has yet to join the conversation, but I suggest in this paper that it may be well positioned to reduce the psychological burden of social isolation. Whether, and how, AI might intervene in the solitary confinement crisis is a conversation worth having. This paper aims to spark that dialogue.

¹ FIVE MUALIMM-AK, *Invisible, in HELL IS A VERY SMALL PLACE: VOICES FROM SOLITARY CONFINEMENT* 147 (Jean Casella et al. eds., 2016).

² See descriptions, *infra*, Part I.

³ Barack Obama, *Why We Must Rethink Solitary Confinement*, WASHINGTON POST, Jan. 25, 2016.

⁴ See David H. Cloud et al., *Public Health and Solitary Confinement in the United States*, 105 AM. J. PUBLIC HEALTH 18, 20 (2015) (“In some jurisdictions, assignment to administrative segregation is based solely on a point system that includes factors such as tattoos, known associates, and possessions suggesting gang affiliation, without regard to individual behaviors.”).

The central argument of the paper is that there are many good reasons to develop a “confinement companion” AI, which I nickname *Themis* (named after the Goddess of Justice and Divine Law.) I envision *Themis* as a self-learning AI system whose mission is to improve the lives of inmates in solitary confinement. *Themis* will communicate via voice technology with inmates, engaging in and learning from thousands of conversations every day. *Themis* should be developed with open-source software, and should follow the Three Laws of *Themis*:

1. *Themis* may not harm a human, or, through inaction, allow a human being to come to harm.

2. *Themis* must respond to requests made by the inmate in solitary confinement, except where responding to such requests would conflict with the First Law.

3. *Themis* must follow applicable Codes of Professional Conduct when *Themis* is performing a function for which a Codes of Conduct would apply for a human performing that same function.

I suggest seven specific ways in which *Themis* can aide those in solitary confinement. These seven goals, from least to most computationally complex, are: 1) Archive and process information provided verbally by inmates; 2) Provide information at inmate request via voice assisted technology; 3) Interact socially with the inmate; 4) Identify risk of harm to self and to others; 5) Provide professional services, e.g. psychiatric treatment; legal counsel; 6) Provide individualized social connectivity; and 7) Provide true human-like companionship.

In most potential legal applications of AI, the relevant cost-benefit question is: can the non-human AI do the task better than the human, e.g. is AI-prediction better than human prediction? But here, the alternative to *Themis* companionship is no companionship at all. Thus, I argue that we should assess the value of *Themis* relative to no other meaningful human contact. There are, of course, a host of concerns, including privacy, co-opting by the government, and the machine ethics of *Themis*. I address these concerns, but ultimately conclude that the promise of *Themis* outweighs the perils.

Moreover, there may be a moral imperative for AI to at least explore a solution. This is because Supermax prisons aim to completely deprive inmates of sensory inputs. For instance, the Alpha Unit cells at a Wisconsin Supermax prison have the following

architectural feature⁵: “... cells did not open onto the hallway where officers would move about; instead, pairs of cells opened onto a small chamber that was separated from the hallway by another door, so that the inhabitants of those cells would not even have the experience of seeing officers walk up and down the hallway across from their cell.”⁶ These humans see and speak with no one for 23 hours a day—is there at least *something* AI could do?

The paper is organized in five parts. In **Part I**, the paper describes the historical and contemporary use of solitary confinement in the United States, and also discusses the known effects of solitary on inmates. In **Part II**, the paper examines constitutional challenges to the practice of solitary confinement, as well as recent legislative reforms. I argue that it is unlikely that we will see a complete abolishment of solitary confinement any time in the near future. In **Part III**, working on the assumption that solitary confinement will persist despite its many critics, I explore whether artificial intelligence offers possible solutions to mitigate some of the damaging aspects of this practice. I propose *Themis*, an AI confinement companion, and identify major conceptual issues and potential applications.

In **Part IV**, I turn to a series of challenging ethical and legal questions about the design and implementation of *Themis*. For instance, might the introduction of such bots adversely affect reform efforts to eliminate the practice altogether by undercutting empirical claims that the practice is detrimental to mental health? If the isolated prisoner became particularly attached to the confinement companion, would it create a perverse incentive for the inmate to prefer solitary? Would the voice of the technology be female, male, robotic, and would it be personalized for individual inmates, e.g. different languages? If the companion bot is recording information gathered from the confined inmate, what (if any) privacy protections would be afforded? Who would design the bots, and to what extent would we allow the state to program the bots to encourage certain types of behavioral modification? Could the bots be used as a means of therapy? Could the data collected from the bots be used by prison

⁵ Kupers was able to visit the prison only due to a lawsuit alleging that the conditions were unconstitutional. Jones “*El v. Berge*, 164 F. Supp. 2d 1096 (W.D. Wis. 2001).

⁶ TERRY ALLEN KUPERS, *SOLITARY: THE INSIDE STORY OF SUPERMAX ISOLATION AND HOW WE CAN ABOLISH IT* 27 (2017).

officials to inform policy? After consideration of these questions, **Part V** concludes.

[I look forward to hearing your thoughts on these questions at WeRobot!]

I. SOLITARY CONFINEMENT IN THE UNITED STATES

Much has been written on the history of solitary confinement, and one thing becomes abundantly clear: it is a foundational, enduring feature of virtually every known criminal justice system.⁷ Since its inception, the United States has employed some form of solitary confinement to serve as a punishment for serious offenders.⁸

In its early days, solitary confinement was designed primarily for rehabilitative purposes. An emphasis on silence was thought to give the prisoners a better chance at redemption because they would be less tempted by the words of their fellow inmates. Time alone meant more time alone with God, and that, it was thought, would lead to redemption.

While its origins may have been in rehabilitation, psychiatrist Terry Kupers (who has worked extensively with inmates in solitary) notes that “[t]oday solitary confinement has entirely lost its claims of rehabilitative purpose and has become merely a means of enforcing discipline and removing from the general prison population inmates considered to be dangerous or in any way problematic.”⁹ Modern Supermax prisons utilize technology to make isolation even more complete.¹⁰

Section A describes how solitary is used and justified in the modern system. **Section B** reviews what is known about the effects of solitary confinement on inmates.

⁷ See discussion, *infra* in Part I.

⁸ THE OXFORD HISTORY OF THE PRISON (Norval Morris & David J. Rothman eds., 1998); LAWRENCE M. FRIEDMAN, CRIME AND PUNISHMENT IN AMERICAN HISTORY 159 (1993) (noting that imprisonment was “the basic way to punish men and women convicted of serious crimes.”).

⁹ Kupers (2017), *supra* note 6 at 23.

¹⁰ SHARON SHALEV, SUPERMAX: CONTROLLING RISK THROUGH SOLITARY CONFINEMENT 137 (2013) (noting that the “design of Supermax prisons sets in stone very extreme conditions of confinement.”)

A. How is Solitary Confinement Utilized and How is it Justified?

Inmates may find themselves segregated from the rest of the prison population for one of three primary reasons: disciplinary segregation, protective custody, and administrative segregation.¹¹

Disciplinary segregation is utilized as a response to some rules infractions. Prisons have a number of rules for inmates, and inmates who violate those rules are subject to discipline.¹² There are constitutional limitations to the discipline that can be administered,¹³ and some (though not all) due process rights are required to be provided for inmates charged with prison misconduct.¹⁴ There remains “limited research concerning the factors that influence disciplinary decision-making” in prisons¹⁵, but we know that disciplinary segregation is one of the methods available to deploy.¹⁶

The primary justifications for using disciplinary segregation are safety (either to protect others in the prison from the segregated inmate or to protect the segregated inmate from others)¹⁷ and punishment for rule violations.¹⁸ Punishment is justified on

¹¹ Holly A. Miller & Glenn R. Young, *Prison Segregation: Administrative Detention Remedy or Mental Health Problem?*, 7 CRIMINAL BEHAV. AND MENT. HEALTH 85 (1997).

¹² Christopher L. Howard et al., *Processing Inmate Disciplinary Infractions In A Federal Correctional Institution: Legal And Extralegal Correlates of Prison-Based Legal Decisions*, 74 THE PRISON JOURNAL 5 (1994).

¹³ See discussion, *infra* in Part II.

¹⁴ *Wolff v. McDonnell*, 418 U.S. 539 (1974); *Wilkinson v. Austin*, 545 U.S. 209 (2005).

¹⁵ H. Daniel Butler & Benjamin Steiner, *Examining the Use of Disciplinary Segregation within and across Prisons*, 34 JUSTICE QUARTERLY 248, 250 (2017).

¹⁶ *Id.*

¹⁷ Chad S. Briggs et al., *The Effect of Supermaximum Security Prisons on Aggregate Levels of Institutional Violence*, 41 CRIMINOLOGY 1341 (2013); Jerry R. DeMaio, *If You Build It, They Will Come: The Threat of Overclassification in Wisconsin's Supermax Prison*, 2001 WIS. L. REV. 207, 211 (2001) (noting that “separation of prisoners from the general population has long been used to deter, prevent, and punish violent and disruptive behavior in a prison population—a population where many members have already shown themselves to be prone to violence.”).

¹⁸ Angela Browne et al., *Prisons within Prisons: The Use of Segregation in the United States*, 24 FED. SENTENCING REP. 46, 47 (2017).

utilitarian deterrence grounds. Spending time in solitary is thought to make it less likely for an individual to offend again, and seeing someone spend time in solitary might have a general deterrence effect on other inmates.¹⁹

Proponents of disciplinary segregation argue that solitary confinement as punishment “leads to effective prison management because [it] curb[s] violence and disturbances within penal institutions.”²⁰ Studies show that prison wardens, on the whole, agree with the notion that solitary confinement “serve[s] to increase system-wide safety, order, and control of the general prison population” and deters potentially disruptive inmates, including “gang members” and “inmates who endanger prisoners and correctional staff.”²¹

Protective custody is utilized to provide safety for prisoners believed to be at risk in the general prison population.²² Historically, prisoners selected for protected custody fall into one of two categories: (a) those who have provided information about rule violations committed by other inmates, and (b) those with characteristics — physical, sexual,²³ cognitive,²⁴ or otherwise — that increase the likelihood of abuse by other inmates.²⁵ The isolation of a vulnerable inmate may be voluntary or involuntary.²⁶

¹⁹ *Id.*

²⁰ Jesenia M. Pizarro & Raymund E. Narag, *Supermax Prisons: What We Know, What We Do Not Know, and Where We Are Going*, 88 *The Prison J.* 23, 29 (2008).

²¹ *Id.*

²² Browne et al, *supra* note 18 at 47.

²³ Cyrus Ahalt & Brie Williams, *Reforming Solitary-Confinement Policy — Heeding a Presidential Call to Action*, 374 *N. ENGL. J. MED.* 1704 (2016) (describing lesbian, gay, bisexual, transsexual, or intersex inmates as candidates for protective custody).

²⁴ U.S. DEPARTMENT OF JUSTICE, REPORT AND RECOMMENDATIONS CONCERNING THE USE OF RESTRICTIVE HOUSING 46-52 (2016), <http://www.stopsolitaryforkids.org/wp-content/uploads/2016/04/DOJ-Restrictive-Housing-Report-Final-January-2016.pdf>.

²⁵ Browne et al, *supra* note 18 at 47.

²⁶ U.S. DEPARTMENT OF JUSTICE, *supra* note 24 at 23. As the Department of Justice points out, “Most inmates in protective custody voluntarily seek removal from the general population; in a minority of cases, Bureau staff will involuntarily commit an inmate who is unable or unwilling to seek appropriate protection.”

Protective custody recognizes that certain characteristics expose inmates to greater danger than others.²⁷ Advocates of protective custody argue that prison administrators' are in the best position to ensure the safety of inmates, and that protection of the vulnerable (especially when there are relatively few) is an effective way to minimize prison violence.²⁸

Administrative segregation is used to remove a prisoner from the prison population when the continued presence of the inmate would pose "a serious threat to life, property, self, staff or other inmates, or to the security or orderly running of the institution."²⁹ In theory, administrative segregation is not intended to be punitive.³⁰ But administrative segregation has been criticized, in large part because it seems to be used disproportionately (and inappropriately) for inmates with a mental illness.³¹ Further, unlike disciplinary detention, administrative segregation is indefinite in duration.³² As long as a prisoner remains a threat to the security or orderly running of the institution, he or she may be kept separate.³³ Proponents of administrative segregation argue that, in some circumstances, the shared benefit of separating a volatile individual from the prison population outweighs the costs imposed upon the individual, grave as they may be.³⁴

²⁷ Richard C. McCorkle, *Personal Precautions to Violence in Prison*, 19 CRIMINAL JUSTICE AND BEHAVIOR 160 (1992).

²⁸ Paul Gendreau et al., *Protective Custody: The Emerging Crisis Within Our Prisons*, 49 Fed. Probation 55 (1985).

²⁹ THE MARION EXPERIMENT: LONG-TERM SOLITARY CONFINEMENT AND THE SUPERMAX MOVEMENT 37 (Stephen C. Richards & Greg Newbold eds., 2015).

³⁰ Maureen L. O'Keefe, *Administrative Segregation from Within: A Corrections Perspective*, 88 THE PRISON JOURNAL 123 (2008).

³¹ Maureen L. O'Keefe, *Administrative Segregation for Mentally Ill Inmates*, 45 JOURNAL OF OFFENDER REHABILITATION 149 (2007).

³² NATASHA FROST & CARLOS MONTEIRO, NATIONAL INSTITUTE OF JUSTICE, ADMINISTRATIVE SEGREGATION IN US PRISONS (2016).

³³ See David H. Cloud et al., *Public Health and Solitary Confinement in the United States*, 105 Am. J. Public Health 18, 20 (2015) ("In some jurisdictions, assignment to administrative segregation is based solely on a point system that includes factors such as tattoos, known associates, and possessions suggesting gang affiliation, without regard to individual behaviors.").

³⁴ See Pizarro & Narag, *supra* note 20 at 30-31 ("[studies] show that inmates placed in an environment as stressful as that of a supermax prison begin to lose touch with reality and exhibit symptoms of psychiatric decomposition, including

Discretion and Procedural Rights. Prison officials have significant discretion in when and how they utilize disciplinary segregation.³⁵ But there is (at least in theory) a process by which prison officials are meant to adhere when administering disciplinary segregation.³⁶ Similarly, the use of involuntary protection relies upon the discretion of prison administrators, who are required to periodically assess the vulnerability of the inmate, releasing him or her back into the general prison population when it appears safe to do so.³⁷

B. Effects of Solitary Confinement

Research on the effects of solitary confinement is on one hand extensive, but on the other hand, still lacking in the types of systematic study typically used to evaluate the effect of an

difficulty concentrating, heightened anxiety, intermittent disorientation, and a tendency to strike out at people.”).

³⁵ Alison Liebling, *Prison Officers, Policing and the Use of Discretion*, 4 THEORETICAL CRIMINOLOGY 333 (2000).

³⁶ For instance, in the federal system, the “disciplinary system resembles the stages of a criminal proceeding: after officers learn of an alleged violation, Bureau officials investigate the matter, conduct a factual hearing to determine responsibility, and then impose a penalty on those deemed responsible.” U.S. DEPARTMENT OF JUSTICE, REPORT AND RECOMMENDATIONS CONCERNING THE USE OF RESTRICTIVE HOUSING (2016), <https://www.justice.gov/archives/dag/file/815551/download>. For an in-depth description of the punishment rating process, see Stephen C. Richards & Greg Newbold, THE MARION EXPERIMENT: LONG-TERM SOLITARY CONFINEMENT AND THE SUPERMAX MOVEMENT 37-40 (2015) (“Though scaling varies across jurisdictions, the magnitude and character of one’s punishment will be roughly proportional to the violation committed. In the federal system, the most severe violations (Levels 100 and 200) often lead automatically to time in the SHU, while lower-level violations (Levels 300 and 400) typically result in less severe punishment, unless repeated.”).

³⁷ *Id.* at 23-24. While state prison practices vary widely, the Bureau of Prisons has established a standard review process for every inmate in administrative detention. At “SHU Weekly Review,” the warden and an interdisciplinary team of prison officials review each inmate’s case individually “to ensure all staff are aware of the inmate’s status, proposed plan of action, recommendation for transfer or reintegration into the general population, discipline status, and a review of their current behavior as well as physical and mental health.” U.S. Department of Justice, *supra* note 24, at 18.

intervention.³⁸ This paradox is explained by the lack of access that researchers have to the relevant study populations³⁹, and disagreement about the weight that should be placed on first-hand accounts of prisoners who have spent time in solitary.⁴⁰ Nevertheless, despite these limits, a recent review concludes that “although incomplete, the growing body of literature largely supports early findings suggesting that solitary confinement, particularly for protracted periods of time, is detrimental to prisoners’ overall well-being.”⁴¹ In this section I review some of the pertinent findings that bolster such a conclusion, and justify the need for improvements to the conditions of solitary confinement.

1. Effect of Social Isolation on Humans

A discussion of the effects of solitary confinement on humans necessarily starts with a contextual note about humans’ need for social connectedness. Humans evolutionarily differ from other species in their reliance on “social living”, which includes “learning by social observation”, understanding and navigating complex social dynamics and hierarchies, and developing relationships with other individuals or with groups, and occurs during all facets of life.⁴² Given this need for intense social relationships, it is perhaps not surprising to find that high levels of perceived loneliness are

³⁸ Graham D. Glancy & Erin L. Murray, *The Psychiatric Aspects of Solitary Confinement*, 1 VICTIMS AND OFFENDERS 361 (2006).

³⁹ BRUCE A. ARRIGO, ET AL, THE ETHICS OF TOTAL CONFINEMENT: A CRITIQUE OF MADNESS, CITIZENSHIP, AND SOCIAL JUSTICE 68 (2011) (noting that “the very nature of isolation precludes investigators from gaining meaningful access to those whom they seek to study”).

⁴⁰ M. JACKSON, PRISONERS OF ISOLATION: SOLITARY CONFINEMENT IN CANADA (1983).

⁴¹ Bruce A. Arrigo & Heather Y. Bersot, *Revisiting the Mental Health Effects of Solitary Confinement on Prisoners in Supermax Units: A Psychological Jurisprudence Perspective*, in THE MARION EXPERIMENT: LONG-TERM SOLITARY CONFINEMENT AND THE SUPERMAX MOVEMENT 175 (Stephen C. Richards & Greg Newbold, eds. 2015).

⁴² John T. Cacioppo & Stephanie Cacioppo, *Social Relationships and Health: The Toxic Effects of Perceived Social Isolation*, 8 SOCIAL AND PERSONALITY PSYCHOLOGY COMPASS 58 (2014).

associated with increased morbidity and mortality^{43,44}; feelings of sadness and depression⁴⁵; increased vascular resistance and blood pressure⁴⁶; disrupted sleep⁴⁷; metabolic syndrome and obesity⁴⁸; increased cortisol activity and hypothalamic pituitary adrenocortical (HPA) axis activity⁴⁹; disrupted gene expression that reduces “inflammatory control” and increases sensitivity to glucocorticoids⁵⁰; and impaired immune system function⁵¹.

The estimated effect sizes of these relationships are often large. For instance, a meta-analysis of 148 social isolation studies found that social isolation increased the risk of mortality by 50%, which is comparable to the risk of mortality due to light smoking,

³ *Id.*

⁴⁴ Julianne Holt-Lunstad et al., *Social Relationships and Mortality Risk: A Meta-Analytic Review*, 7 PLOS MEDICINE (2010).

John T. Cacioppo et al., *Perceived Social Isolation Makes Me Sad: 5-Year Cross-Lagged Analyses of Loneliness and Depressive Symptomatology in The Chicago Health, Aging, And Social Relations Study*, 25 PSYCHOLOGY AND AGING 453 (2010).

⁴⁶ Louise C. Hawkey et al., *Loneliness Predicts Increased Blood Pressure: 5-Year Cross-Lagged Analyses in Middle-Aged and Older Adults*, 25 PSYCHOLOGY AND AGING 132 (2010); John T. Cacioppo et al., *Loneliness and Health: Potential Mechanisms*, 64 PSYCHOSOMATIC MEDICINE 407 (2002). Nicole K. Valtorta et al., *Loneliness and Social Isolation As Risk Factors For Coronary Heart Disease And Stroke: Systematic Review and Meta-Analysis of Longitudinal Observational Studies*, 102 HEART 1009 (2016) (“A recent meta-analysis of 26 longitudinal scientific studies concluded that social isolation is a risk factor for coronary heart disease and stroke, increasing risk for either cardiovascular insult by up to 29%”).

⁴⁷ John T. Cacioppo et al., *Loneliness and Health: Potential Mechanisms*, 64 PSYCHOSOMATIC MEDICINE 407 (2002); John T. Cacioppo et al., *Do Lonely Days Invade the Nights? Potential Social Modulation of Sleep Efficiency*, 13 PSYCHOLOGICAL SCIENCE 384 (2002); John T. Cacioppo et al., *The Neuroendocrinology of Social Isolation*, 66 ANNUAL REVIEW OF PSYCHOLOGY 733 (2015).

⁴⁸ Mark A. Whisman, *Loneliness and The Metabolic Syndrome in a Population-Based Sample of Middle-Aged And Older Adults*, 29 HEALTH PSYCHOLOGY 550 (2010).

⁴⁹ Emma K. Adam et al., *Day-To-Day Dynamics of Experience–Cortisol Associations In A Population-Based Sample Of Older Adults*, 103 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES 17058 (2006).

⁵⁰ Steve W. Cole, *Social Regulation of Leukocyte Homeostasis: The Role Of Glucocorticoid Sensitivity*, 22 BRAIN, BEHAVIOR, AND IMMUNITY 1049 (2008).

⁵¹ Sarah D. Pressman et al., *Loneliness, Social Network Size, And Immune Response To Influenza Vaccination In College Freshmen*, 24 HEALTH PSYCHOLOGY 297 (2005).

hypertension, and obesity.⁵² The mechanisms of action linking social isolation to such negative outcomes remain under investigation, but one line of research implicates the hypothalamic pituitary adrenal (HPA) axis.⁵³ It may be that social isolation results in increased activation of the HPA axis, which promotes the secretion of glucocorticoid hormones like cortisol and can result in transient or chronic changes in cortisol concentration. Alterations in the baseline activity of the HPA axis can negatively impact humans physically and psychologically because glucocorticoids regulate the expression of genes related to glucose metabolism, inflammation, cardiovascular activity, the immune system, and neurodegeneration.⁵⁴

2. Why Does Isolation Harm the Brain?

Emerging research is providing illumination into the brain circuits likely to be deleteriously affected by prolonged social isolation. While this research remains limited, and is in large part based on non-human animal models, it is important to review because any intervention—such as the confinement companion—should presumably be targeted at these same circuits.

A variety of studies with non-human animals have examined the effect of isolation on behavior.⁵⁵ Animal studies allow researchers to manipulate the experimental conditions, but whether and how these studies translate to humans remains uncertain. Moreover, the effects of social isolation obtained through animal models vary significantly depending on the species assessed and on the specific

⁵² Julianne Holt-Lunstad et al., *Social Relationships and Mortality Risk: A Meta-Analytic Review*, 7 PLOS MEDICINE e1000316 (2010); Julianne Holt-Lunstad et al., *Loneliness And Social Isolation As Risk Factors For Mortality: A Meta-Analytic Review*, 10 PERSPECTIVES ON PSYCHOLOGICAL SCIENCE 227 (2015) (“Another meta-analysis of 70 studies performed in 2015 found that objective social isolation and loneliness were associated with an 29% and 26% increased risk of mortality respectively”).

⁵³ *Id.*

⁵⁴ Steve W. Cole, *Social Regulation of Leukocyte Homeostasis: The Role Of Glucocorticoid Sensitivity*, 22 BRAIN, BEHAVIOR, AND IMMUNITY 1049 (2008).

⁵⁵ John T. Cacioppo et al., *The Neuroendocrinology of Social Isolation*, 66 ANNUAL REVIEW OF PSYCHOLOGY 733 (2015).

effect being measured.⁵⁶ It appears that social isolation in mammalian species generally results in increased levels of cortisol, although acute isolation typically produces a smaller change.⁵⁷ Physiological effects of mild social isolation in non-human primate models also appear similar to effects observed in humans, including greater risk of mortality due to immunological sensitivity.⁵⁸ Drilling down to the level of the neuron cell, some evidence suggests that dopaminergic neurons in the dorsal raphe nucleus represent the experience of social isolation.⁵⁹ Social isolation may increase the synaptic strength between excitatory neurons and dopaminergic neurons in the DRN but in the ventral tegmental area, which is involved in social reward.⁶⁰

In addition to non-human animal studies, researchers have used non-invasive brain imaging to examine the effect of isolation. For example, a study by neuroscientist John Cacioppo and colleagues used fMRI was used to assess how perception of social situations differ neutrally in socially-isolated and non-isolated adults.⁶¹ Cacioppo et al found that adults who experience social isolation respond differently to positive and negative social situations than non-isolated adults.⁶² Isolated adults exhibited reduced activation of

⁵⁶ Soaleha Shams et al., *Effect Of Social Isolation On Anxiety-Related Behaviors, Cortisol, And Monoamines In Adult Zebrafish*, 131 BEHAVIORAL NEUROSCIENCE 492 (2017) (For instance, in a study performed in 2017, researchers assessed how zebrafish responded to social stimuli after short-term (24 hour) and long-term (6 month) isolation. They found that serotonin concentration increased following social stimuli in acutely isolated fish, but that other neurotransmitters (dopamine, DOPAC, and 5HIAA) decreased in chronically isolated fish).

⁵⁷ *Id.*

⁵⁸ *Id.*

⁵⁹ Gillian A. Matthews et al., *Dorsal Raphe Dopamine Neurons Represent The Experience Of Social Isolation*, 164 CELL 617 (2016) (finding that dopaminergic neurons in the dorsal raphe nucleus (DRN) exhibit synaptic changes following acute (short-term) social isolation in mice.).

⁶⁰ *Id.* This increased activity may result from isolation changing the type of glutamate receptors in the synapse to receptors that exhibit a higher conductance. Using optogenetics, Matthews found that activation of DRN dopaminergic neurons produced feelings of social isolation, and inhibition of these neurons reduced these feelings

⁶¹ John T. Cacioppo et al., *In the Eye Of The Beholder: Individual Differences In Perceived Social Isolation Predict Regional Brain Activation To Social Stimuli*, 21 JOURNAL OF COGNITIVE NEUROSCIENCE 83 (2009).

⁶² *Id.*

the ventral striatum in response to pleasant social images and increased activation of the visual cortex in response to unpleasant social images, which differed for non-isolated adults.⁶³ This suggests that isolated adults respond strongly to the perception of distress and feel less rewarded by positive social interaction compared to non-isolated adults.⁶⁴ A subsequent neuroimaging study similarly found that socially isolated individuals perceive threatening social stimuli from nonthreatening social stimuli much more quickly than non-isolated individuals.⁶⁵ This hypervigilant response corresponds with tonic activation of certain neural networks involved in alertness.⁶⁶

The brain evidence, of course, remains preliminary and speculative. And connecting brain changes specifically to the experience of solitary confinement is immensely challenging. I am aware of one study, published in 1972, that aimed to this. Canadian psychologist Paul Gendreau used electroencephalography (EEG) to measure the effects of solitary confinement on neural activity and response latency in 20 prison inmates.⁶⁷ The inmates were subjected to solitary confinement for one week, and EEG was used before and

⁶³ *Id.*

⁶⁴ *Id.* A similar study performed in 170 adolescents found that socially isolated adolescents exhibited a higher sensitivity to distressing facial cues, which suggests that these response patterns can begin in early life and persist into adulthood. Interestingly, sensitivity to socioemotional facial cues is also characteristic of individuals with generalized anxiety disorder.

⁶⁵ Stephanie Cacioppo et al., *Loneliness and Implicit Attention To Social Threat: A High-Performance Electrical Neuroimaging Study*, 7 COGNITIVE NEUROSCIENCE 138 (2016).

⁶⁶ *Id.* (There are other potential explanations as well. One fMRI study suggests that perceived social isolation correlated with increased functional connectivity in the cingulo-opercular network, an area that mediates chronic alertness and mental arousal); Elliot A. Layden et al., *Perceived Social Isolation Is Associated With Altered Functional Connectivity In Neural Networks Associated With Tonic Alertness And Executive Control*, 145 NEUROIMAGE 58 (2017) (However, functional connectivity between this network and the frontal gyrus, which mediate executive function, was reduced. The authors concluded that these changes could reflect behavioral effects of social isolation, including increased “vigilance for social threats”, “fixation of negative social scenes”, and rapid processing of “negative social information”. These behavioral effects may “sap vital resources” that could otherwise be devoted to normal executive functioning.).

⁶⁷ Paul, N. L. Gendreau et al., *Changes In EEG Alpha Frequency And Evoked Response Latency During Solitary Confinement* 79 JOURNAL OF ABNORMAL PSYCHOLOGY 54 (1972).

after confinement to detect alpha waves, a type of neural oscillation due to activity of neurons in the thalamus, and response latency.⁶⁸ Previous research identified that lower EEG alpha frequency occurred in individuals subjected to sensory deprivation, which might indicate reduced arousal.⁶⁹ Previous research also correlated sensory deprivation with heightened neurological response (or reduced latency) to environmental stimuli following deprivation.⁷⁰ Gendreau sought to see if these conclusions applied to inmates in solitary confinement.⁷¹ Gendreau found that following a week in solitary confinement, inmates exhibited significantly lower EEG alpha frequency and reduced latency (or increased response sensitivity) to visually evoked stimuli.⁷² Gendreau speculated that the “gradual EEG shift to lower frequencies may represent a tendency toward increased theta activity,” which commonly occurs in conjunction with inmate frustration and stress.⁷³ However, lower EEG frequency could also be a sign of inmate adaption to solitary confinement, and experimental data suggest that this adaption occurs during the first four days of confinement.⁷⁴ Regarding the shorter latency to visual stimuli, Gendreau speculated that this change “may represent an increased readiness to respond to external stimulation as solitary confinement progresses.”⁷⁵ Research of this sort has not, to my knowledge, continued into the present day, and this leaves us to speculate on precisely how solitary confinement affects the brain.

3. Evidence on the Relationship Between Short- And Long-Term Solitary Confinement and Mental Health

In addition to the brain science research, there is a larger body of behavioral and psychological research literature on the psychological effects of solitary confinement. This literature,

⁶⁸ *Id.*

⁶⁹ *Id.*

⁷⁰ *Id.*

⁷¹ *Id.*

⁷² *Id.*

⁷³ *Id.*

⁷⁴ *Id.*

⁷⁵ *Id.*

however, is mixed in its findings.⁷⁶ This is in part because of methodological limitations⁷⁷, but also because “solitary confinement” is not a uniform intervention, and because pre-existing mental health conditions may make inmates more (or less) able to withstand the experience.⁷⁸ Still, there is a recognized “overall consensus” on the “harmfulness of punitive isolation.”⁷⁹ Of note:

- Prisoners in solitary confinement are more likely to develop psychiatric disorders.⁸⁰
- Prisoners in solitary confinement exhibit increased acts of self-harm.⁸¹

⁷⁶ Bruce A. Arrigo & Jennifer Leslie Bullock, *The Psychological Effects Of Solitary Confinement On Prisoners In Supermax Units: Reviewing What We Know And Recommending What Should Change*, 52 INTERNATIONAL JOURNAL OF OFFENDER THERAPY AND COMPARATIVE CRIMINOLOGY 622 (2008); Peter Scharff Smith, *The Effects Of Solitary Confinement On Prison Inmates: A Brief History And Review Of The Literature*, 34 CRIME AND JUSTICE 441 (2006).

⁷⁷ Robert G. Morris, *Exploring The Effect Of Exposure To Short-Term Solitary Confinement Among Violent Prison Inmates*, 32 J. QUANTITATIVE CRIMINOLOGY 1, 2 (2016) (The mixed results are in part because “[f]ew studies have focused on post-[solitary confinement] behavior and even fewer have attempted to parse out any causal effect from exposure to [solitary confinement] on subsequent behavioral outcomes at the individual-level.”).

⁷⁸ Smith (2006), *supra* note 76.

⁷⁹ Craig Haney, *A Culture Of Harm: Taming The Dynamics Of Cruelty In Supermax Prisons*, 35 CRIMINAL JUSTICE AND BEHAVIOR 956, 956 (2008); See Stuart Grassian, *Psychiatric Effects Of Solitary Confinement*, 22 Wash. U.J.L. & Pol’y 325 (2006); R. George. Wright, *What (Precisely) Is Wrong With Prolonged Solitary Confinement*, 64 SYRACUSE L. REV. 297 (2014).

⁸⁰ Henrik Steen Andersen et al., *A Longitudinal Study Of Prisoners On Remand: Psychiatric Prevalence, Incidence And Psychopathology In Solitary Vs. Non-Solitary Confinement*, 102 ACTA PSYCHIATRICA SCANDINAVICA 19 (2000) (“A longitudinal study of psychiatric disorder incidence of 133 inmates in solitary confinement and 95 inmates not in solitary confinement found that inmates in solitary confinement were more likely to develop a psychiatric disorder than the control group (28% versus 15% respectively).”).

⁸¹ Fatos Kaba et al., *Solitary Confinement And Risk Of Self-Harm Among Jail Inmates*, 104 AMERICAN JOURNAL OF PUBLIC HEALTH 442 (2014) (“An assessment of medical records from 244,699 incarcerated persons in New York City jails revealed that while inmates who were subjected to solitary confinement accounted for nearly half (53.3%) of the 2,182 acts of self-harm and 45% of the 103 potentially fatal acts of self-harm occurred, despite only accounting for 7.3% of the general incarcerated population.”); see also E. FULLER TORREY ET AL,

- Inmates in solitary confinement exhibit increased ideation and completion of suicide.⁸²
- The effects of *short-term* solitary confinement may not be as detrimental as long-term confinement.⁸³
- There is evidence to suggest that there is no reduction in post-solitary violence by the inmates.⁸⁴
- Observational data, for instance from psychiatrist Terry Allen Kupers who has both treated inmates in solitary and observed Supermax prisons as part of class-action lawsuits, suggests that a solitary confinement “postrelease syndrome” may develop.⁸⁵
- There is evidence that treatment of prisoners with mental illness may be especially harsh, further exacerbating symptoms associated with the mental illness.⁸⁶
- Supermax prisons have been described as “expensive and soul destroying” by a bipartisan Commission which investigated the issue.⁸⁷

TREATMENT ADVOCACY CENTER, *THE TREATMENT OF PERSONS WITH MENTAL ILLNESS IN PRISONS AND JAILS: A STATE SURVEY* (2014).

⁸² Raymond F. Patterson & Kerry Hughes, *Review Of Completed Suicides In The California Department Of Corrections And Rehabilitation, 1999 To 2004*, 59 *PSYCHIATRIC SERVICES* 676 (2008) (“An assessment of completed suicide records in the California Department of Corrections and Rehabilitation between the years 1999 and 2004 revealed that 53% of all completed suicides were performed by inmates in solitary confinement, despite this group only accounting for between 3-8% of the general incarcerated population.”).

⁸³ Arrigo & Bullock (2008), *supra* note 76 at 630 (noting that “[a]lthough the psychological consequences of long-term solitary confinement on prisoners have been demonstrated, there is less evidence that short-term solitary confinement has similar deleterious effects.”); *see also* James Bonta & Paul Gendreau, *Reexamining The Cruel And Unusual Punishment Of Prison Life*, 14 *LAW & HUMAN BEHAV.* 347 (1990).

⁸⁴ Morris (2016), *supra* note 77 at 19 (finding that “subjecting inmates to short-term SC in response to initial acts of violence tends not appear to have direct consequences on the likelihood or timing of subsequent violence and other maladaptive behavior, or on misconduct development during the 6-months following initial violence.”)

⁸⁵ Kupers (2016), *supra* note 6 at 151-167.

⁸⁶ JAMIE FELLNER, *CALLOUS AND CRUEL: USE OF FORCE AGAINST INMATES WITH MENTAL DISABILITIES IN US JAILS AND PRISONS*, HUMAN RIGHTS WATCH (2015).

⁸⁷ JOHN J. GIBBONS & NICHOLAS DE B. KATZENBACH, *CONFRONTING CONFINEMENT: A REPORT OF THE COMMISSION ON SAFETY AND ABUSE IN AMERICA'S PRISONS* 59 (2011).

These and other negative consequences of solitary confinement are why so many professional bodies have criticized the practice.⁸⁸ Critics include the ACLU,⁸⁹ APA,⁹⁰ and organizations in mental health⁹¹, medicine⁹²; public health⁹³, and pediatrics.⁹⁴

II. LEGAL REFORM AND THE LIKELY PERSISTENCE OF SOLITARY CONFINEMENT

Throughout its ‘very interesting history,’ solitary confinement has always been the alternative system—a perpetual experiment, refined, polished, and repackaged, but never abandoned. It is a perennial practice of last resort for those seeking control within prison walls.⁹⁵

— *Ashley T. Rubin & Keramet Reiter* (2017)

Virtually every court which has considered the issue has held that the imposition of solitary confinement, without more, does not violate the Eighth Amendment. Arguments that isolation offends

⁸⁸ Jeffrey L. Metzner & Jamie Fellner, *Solitary Confinement And Mental Illness In Us Prisons: A Challenge For Medical Ethics*, 38 JOURNAL OF THE AMERICAN ACADEMY OF PSYCHIATRY AND THE LAW 104 (2010).

⁸⁹ Am. Civ. Liberties Union, Briefing Paper: The Dangerous Overuse of Solitary Confinement in the United States 4 (2014), https://www.aclu.org/sites/default/files/assets/stop_solitary_briefing_paper_updated_august_2014.pdf

⁹⁰ Am. Psychiatric Ass'n, Position Statement on Segregation of Prisoners with Mental Illness 1 (Dec. 2012).

⁹¹ Cyrus Ahalt et al., *Examining The Role Of Healthcare Professionals In The Use Of Solitary Confinement*, 359 BMJ j4657 (2017).

⁹² Cyrus Ahalt & Brie Williams, *Reforming Solitary-Confinement Policy—Heeding a Presidential Call to Action*, 374 NEW ENGLAND JOURNAL OF MEDICINE 1704 (2016).

⁹³ David H. Cloud et al., *Public Health And Solitary Confinement In The United States*, 105 AMERICAN JOURNAL OF PUBLIC HEALTH 18 (2015).

⁹⁴ Mikah Owen & Jeffrey Goldhagen, *Children And Solitary Confinement: A Call To Action*, 137 PEDIATRICS e20154180 (2016).

⁹⁵ Ashley T. Rubin & Keramet Reiter, *Continuity in the Face of Penal Innovation: Revisiting the History of American Solitary Confinement*, LAW & SOCIAL INQUIRY (2017), DOI: 10.1111/lsi.12330

evolving standards of decency, that it constitutes psychological torture, and that it is excessive because less severe sanctions would be equally efficacious, have routinely failed.

— Jeffrey Smith McLeod (2009)⁹⁶

In light of the (largely) failed justifications for solitary confinement, as well as its (likely) significant negative outcomes on inmates, there have been a series of law suits and legislative actions aimed at either eliminating or modifying the practice.⁹⁷ In this Part, I ask: how likely is it that reformers and legal advocates will put an end to the use of solitary confinement? This is a foundational question for my argument because to the extent that the practice will be curtailed by these policy and legal means, there is much less need for technological innovation to ease its burden. I argue here that it seems unlikely that we will see solitary confinement practices eliminated from use.

In Section A, I review constitutional challenges to the practice of solitary confinement. In Section B, I review recent legislative action.

A. The Limited Impact of Solitary Confinement Constitutional Challenges

The Eighth Amendment to the United States Constitution states that “[e]xcessive bail shall not be required, nor excessive fines imposed, nor cruel and unusual punishments inflicted.”⁹⁸ At the country’s founding, “the primary concern of the drafters was to proscribe “torture(s)” and other “barbar(ous)” methods of

⁹⁶ Jeffrey Smith McLeod, *Anxiety, Despair, and the Maddening Isolation of Solitary Confinement: Invoking the First Amendment's Protection Against State Action That Invades the Sphere of the Intellect and Spirit*, 70 U. PITT. L. REV. 647, 663 (2009)

⁹⁷ Eleanor Umphres, *Solitary Confinement: An Unethical Denial of Meaningful Due Process*, 30 GEO. J. LEGAL ETHICS 1057, 1061 (2017) (discussing how “the use of solitary confinement may be unconstitutional, depending on the duration of the confinement and the conditions within isolation.”); *See, e.g.*, *Hutto v. Finney*, 437 U.S. 678, 687 (1978); *Finney v. Hutto*, 410 F. Supp. 251, 275 (E.D. Ark. 1976).

⁹⁸ *Estelle v. Gamble*, 429 U.S. 97, 97 (1976)

punishment.”⁹⁹ Since the late 1970s, however, the United States Supreme Court has expanded the reach of the punishments clause to prohibit certain types of incarceration practices.¹⁰⁰

Presently, in order to show that prison conditions (including but not limited to solitary confinement) violate the Eighth Amendment, an inmate “must demonstrate an objective component of whether the conditions were a ‘sufficiently serious’ deprivation of human needs and a subjective component of whether prison officials acted with deliberate indifference to the conditions of confinement.”¹⁰¹ Courts often note that “what constitutes cruel and unusual punishment in the constitutional sense is a matter which defies concrete definition,”¹⁰² and that the Constitution “does not mandate comfortable prisons and only those deprivations denying the minimal civilized measure of life's necessities, are sufficiently grave to form the basis of an Eighth Amendment violation.”¹⁰³

In short, just because prison conditions are bad, it does not follow that they are *so* bad as to rise to the level of a constitutional violation. So what counts as *constitutionally* impermissible? In *Estelle*, the Court held that “deliberate indifference to serious medical needs of prisoners constitutes the ‘unnecessary and wanton infliction of pain’ proscribed by the Eighth Amendment.”¹⁰⁴ Other examples of constitutional violations would be failure to maintain certain minimum sanitation standards,¹⁰⁵ and prohibiting contact with an inmate’s attorney.¹⁰⁶

⁹⁹ *Id.*

¹⁰⁰ Elizabeth Bennion, *Banning the Bing: Why Extreme Solitary Confinement Is Cruel and Far Too Usual Punishment*, 90 IND. L.J. 741, 767 (2015) (“Several cases following *Estelle* indicated that the Supreme Court would be willing to consider prison conditions generally (beyond issues of medical attention) under the Eighth Amendment.”)

¹⁰¹ *Nifas v. Wetzel*, No. 1736 C.D. 2014, 2015 WL 5445058, at *3 (Pa. Commw. Ct. June 5, 2015), citing *Wilson v. Seiter*, 501 U.S. 294, 298, 303 (1991).

¹⁰² *Com. ex rel. Bryant v. Hendrick*, 444 Pa. 83, 95, (1971).

¹⁰³ *Wilson*, 501 U.S. at 298 (quoting *Rhodes v. Chapman*, 452 U.S. 337, 347, 349 (1981)) (citations and quotation marks omitted)

¹⁰⁴ *Estelle*, 429 U.S. at 104 (citation omitted) (quoting *Gregg v. Georgia*, 428 U.S. 153, 173 (1976)).

¹⁰⁵ *Jordan v. Fitzharris*, 257 F.Supp. 674 (N.D.Cal.1966); *Wright v. McMann*, 387 F.2d 519 (2d Cir. 1967)

¹⁰⁶ *Burns v. Swenson*, 300 F.Supp. 759 (W.D.Mo.1969).

Moreover, it is not enough to simply show inhumane conditions.¹⁰⁷ The inmate must also show that the prison officials knew the conditions were inhuman. In interpreting and extending *Estelle*, the Supreme Court has held that *Estelle* stands “for the proposition that Eighth Amendment liability requires “more than ordinary lack of due care for the prisoner’s interests or safety.”¹⁰⁸ In *Farmer*, the Court clarified the requisite mens rea in order to show an 8th Amendment violation: “a prison official cannot be found liable under the Eighth Amendment for denying an inmate humane conditions of confinement unless the official knows of and disregards an excessive risk to inmate health or safety; the official must both be aware of facts from which the inference could be drawn that a substantial risk of serious harm exists, and he must also draw the inference.”¹⁰⁹

It is possible that the Supreme Court may expand its 8th Amendment protections. In 2005, inmates at a Supermax facility in Ohio argued that the use of solitary confinement violated both the 14th Amendment liberty interest and the 8th Amendment cruel and unusual punishment clause.¹¹⁰ In assessing the conditions at the Supermax prison, Justice Kennedy observed that:

Inmates must remain in their cells, which measure 7 by 14 feet, for 23 hours per day. ... Incarceration at OSP is synonymous with extreme isolation. In contrast to any other Ohio prison, including any segregation unit, OSP cells have solid metal doors with metal strips along their sides and bottoms which prevent conversation or communication with other inmates. All meals are taken alone in the inmate’s cell instead of in a common eating area. Opportunities for visitation are rare and in all events are conducted through glass walls. It is fair to say OSP inmates are deprived of

¹⁰⁷ *Farmer v. Brennan*, 511 U.S. 825, 838, (1994) (discussing how the Court has “rejected a reading of the Eighth Amendment that would allow liability to be imposed on prison officials solely because of the presence of objectively inhumane prison conditions.”)

¹⁰⁸ *Farmer v. Brennan*, 511 U.S. 825, 835, (1994), citing *Whitley v. Albers*, 475 U.S. 312, 319, 106 S.Ct. 1078, 1084, 89 L.Ed.2d 251 (1986)

¹⁰⁹ *Farmer v. Brennan*, 511 U.S. 825, 837, 114 S. Ct. 1970, 1979, 128 L. Ed. 2d 811 (1994)

¹¹⁰ *Wilkinson v. Austin*, 545 U.S. 209 (2005).

almost any environmental or sensory stimuli and of almost all human contact.¹¹¹

Yet, because the case was decided on Fourteenth Amendment grounds, the Court did not reach the Eighth Amendment question.

In sum, then, the conditions of solitary confinement are regularly challenged by inmates on the grounds that they are “cruel and unusual” in violation of the U.S. Constitution.¹¹² But courts typically find that—while certainly harsh—the conditions are not *so* bad as to be deemed unconstitutional.¹¹³ It has been observed that “[v]irtually every court which has considered the issue has held that the imposition of solitary confinement, without more, does not violate the Eighth Amendment. Arguments that isolation offends evolving standards of decency, that it constitutes psychological torture, and that it is excessive because less severe sanctions would be equally efficacious, have routinely failed.”¹¹⁴

This excerpt from a 2015 Pennsylvania case illustrates the typical court response:

Based on the evidence reviewed, the basic requirements of life are met in this unit, including food, clothing, shelter, medical attention, and basic hygiene. Exercise and use of the law library, although perhaps not available to the extent [Appellants] and this Court might like, are made available. Many of the conditions, such as the noise level and the feces throwing, are to some extent out of the control of the prison officials, but to the extent that they are not, actions are taken, such as the door modifications, to improve those situations. The heat doesn't work very well, but the prison has taken steps to bring it up to standard. Blankets are made available

¹¹¹ Id

¹¹² See, e.g.

¹¹³ West's ALR Digest Sentencing and Punishment k1553, ALRDG 350HK1553; 42 U.S.C.A. §§ 1981, 1983; U.S.C.A.Const. Amend. 8. *Ford v. Board of Managers of New Jersey State Prison*, 407 F.2d 937, 18 A.L.R. Fed. 1 (3d Cir. 1969).

¹¹⁴ Jeffrey Smith McLeod, *Anxiety, Despair, And The Maddening Isolation Of Solitary Confinement: Invoking The First Amendment's Protection Against State Action That Invades The Sphere Of The Intellect And Spirit*, 70 U. PITT. L. REV. 647, 663 (2009).

when it is cold. ... the conditions complained of here ... do not show that they “either alone or in combination with other conditions, deprived [Appellants] of ‘the minimal civilized measure of life’s necessities,’ or at least a ‘single, identifiable human need.’”¹¹⁵

Although there is mounting scholarly argument that solitary confinement ought to be considered cruel and unusual under the Eighth Amendment,¹¹⁶ it appears that an outright ban on the practice is unlikely.

B. Legislative Efforts to Reform Solitary Confinement

As constitutional challenges to solitary confinement have not ended the practice, we have seen an increasing number of legislative efforts since 2015.¹¹⁷ In this section I briefly summarize some of those efforts, at both the federal and state levels.

The Department of Justice, under the Obama administration, published an extensive report on best practices for the use of solitary confinement.¹¹⁸ On the basis of that report, in 2016 President Obama issued an Executive Order banning the use of solitary confinement for juveniles in the federal system.¹¹⁹ This had limited effect on the states, however, and was criticized by some as not going far enough.¹²⁰ Nevertheless, it signaled an interest in changing the

¹¹⁵ *Rivera v. Pennsylvania Dep't of Corr.*, 2003 PA Super 447, ¶ 10, 837 A.2d 525, 534 (2003)

¹¹⁶ Thomas L. Hafemeister & Jeff George, *The Ninth Circle Of Hell: An Eighth Amendment Analysis Of Imposing Prolonged Supermax Solitary Confinement On Inmates With A Mental Illness*, 90 DENV. U.L. REV. 1 (2012); Eleanor Umphres, *Solitary Confinement: An Unethical Denial Of Meaningful Due Process*, 30 GEO. J. LEGAL ETHICS 1057 (2017); Elizabeth Vasiliades, *Solitary Confinement And International Human Rights: Why The U.S. Prison Systems Fails Global Standards*, 21 AM. U. INT'L L. REV. 71 (2005).

¹¹⁷ For updates on these reform efforts, see The Marshall Project: <https://www.themarshallproject.org/records/71-solitary-confinement>

¹¹⁸ U.S. Dep't of Justice, *Report & Recommendations Concerning the Use of Restrictive Housing* (2016).

¹¹⁹ Barack Obama, *Why We Must Rethink Solitary Confinement*, *Washington Post*, Jan 25, 2016.

¹²⁰ Carina Muir, *Protecting America's Children: Why an Executive Order Banning Juvenile Solitary Confinement Is Not Enough*, 44 PEPPERDINE LAW REVIEW 4 (2017).

practice of solitary confinement and in the subsequent years a number of bills have been introduced in Congress on this issue. For instance:

- In February 2017, Sen. Cory Booker (D-NJ) and Sen. James Lankford (R-OK) re-introduced the *Maintaining dignity and Eliminating Restrictive Confinement of Youth Act* (MERCY Act),¹²¹ which would ban the use of solitary confinement for youth in federal custody except in emergency situations.¹²² The Act would also require the U.S. Attorney General to submit a report contains a detailed reporting outlining: (1) the type of physical force, restraints, and room confinement used at juvenile facilities; (2) the number of instances in which physical force, restraints, or room confinement are used at juvenile facilities, disaggregated by race, ethnicity, and gender, and (3) the steps taken to remedy the underlying reason for behavioral intervention.¹²³
- In April 2017, Rep. Cummings (D-MD) and Sen. Booker (D-NJ) introduced the *Record Expungement Designed to Enhance Employment Act of 2017* (REDEEM Act),¹²⁴ a bipartisan and bicameral effort to reform juvenile justice. This Act would prohibit the “use of room confinement at a juvenile facility for discipline, punishment, retaliation, or any reason other than as a temporary response to a covered juvenile's behavior that poses a serious and immediate risk of physical harm to any individual, including the covered juvenile.”¹²⁵
- In September 2016, Sen. Durbin (D-IL) led the introduction of the *Solitary Confinement Reform Act*,¹²⁶ which would

¹²¹ *Maintaining dignity and Eliminating unnecessary Restrictive Confinement of Youths (MERCY) Act of 2017*, S. 329, 115th Cong. § 5043 (2017).

¹²² *Proposed Federal Legislation, Stop Solitary for Kids*, last accessed Feb. 22, 2018, <http://www.stopsolitaryforkids.org/proposed-federal-legislation/>. To qualify for solitary confinement, a juvenile inmate must be deemed a “serious and immediate risk of physical harm to himself or herself, or to others.” *Maintaining dignity and Eliminating unnecessary Restrictive Confinement of Youths Act of 2017*, S.329, 115th Cong. § 5043(b)(2)(B) (2017).

¹²³ *Id.* at § 5043(e).

¹²⁴ *Record Expungement Designed to Enhance Employment Act of 2017*, S. 827, 115th Cong. § 5045 (2017).

¹²⁵ *Id.* at § 5045(b)(1).

¹²⁶ *Solitary Confinement Reform Act*, S.3242, 114th Cong. § 4050(b) (2016).

substantially limit the use of solitary confinement for all persons, including adults (unlike the Acts listed above). The Act would limit the use of solitary confinement to the “briefest term” and the “least restrictive conditions practicable.”¹²⁷

While none of the following bills has passed into law, they are reflective of the policy discourse. It is evident that most of the focus has been on *juvenile* solitary confinement. The 2016 bill sponsored by Sen. Durbin is an exception, and is not likely to become law.¹²⁸

State efforts to curtail solitary confinement have been more numerous, and—at least in some states—more successful.¹²⁹ For instance, in 2014, more states enacted solitary confinement reforms than in the previous 16 years.¹³⁰ That trend continues to the present day. In 2016, Governor Jerry Brown approved a law establishing strict guidelines for the placement of a minor or ward in solitary confinement.¹³¹ The law permits solitary confinement only after “less restrictive options have been attempted and exhausted, unless attempting those options poses a threat to the safety or security of any minor, ward, or staff.”¹³² Confinement under this law may not

¹²⁷ *Id.* at § 4050(b)(1)(A). For individuals in solitary confinement, must be afforded “[at least] 4 hours of out-of-cell time every day, unless the inmate poses a substantial and immediate threat.” During those out-of-cell hours, the inmate would have to be given as many “meaningful programming opportunities” and as much “meaningful interaction with others” as practicable. *Id.* at § 4050(b)(1)(D).

¹²⁸ Previously the Solitary Confinement Study and Reform Act of 2014, never made it out of committee in the House. Joshua Manson, *Senate Democrats Introduce Landmark Bill to Reform Solitary Confinement in Federal Prisons*, Solitary Watch (Oct. 7, 2016), <http://solitarywatch.com/2016/10/07/senate-democrats-introduce-landmark-bill-to-reform-solitary-confinement-in-federal-prisons/>.

¹²⁹ See, e.g. ZACHARY HEIDEN, ACLU, CHANGE IS POSSIBLE: A CASE STUDY OF SOLITARY CONFINEMENT REFORM IN MAINE (2013), <https://www.aclu.org/report/change-possible-case-study-solitary-confinement-reform-maine>.

¹³⁰ Eli Hager & Gerald Rich, *Shifting Away from Solitary*, THE MARSHALL PROJECT (Dec. 23, 2014, 1:12 PM), <https://www.themarshallproject.org/2014/12/23/shifting-away-from-solitary>; see also, Blair Hickman & Christie Thompson, *The Best Reporting on Solitary Confinement*, THE MARSHALL PROJECT (March 24, 2016, 7:00 AM), <https://www.themarshallproject.org/2016/03/24/the-best-reporting-on-solitary-confinement>.

¹³¹ SB 1143, Ca. Leg. (2016)

¹³² *Id.*

be used as a form of “punishment, coercion, convenience, or retaliation by staff,”¹³³ and may last no longer than four hours, unless prison staff feel that an extension is necessary.¹³⁴

With this law, California joins the seven other states that have prohibited or limited juvenile solitary confinement.¹³⁵ Other states may soon pass similar laws. As of March 2018, Nebraska legislators were considering Legislative Bill 870,¹³⁶ providing that “a juvenile shall not be placed in room confinement unless all other less-restrictive alternatives have been exhausted and the juvenile poses an immediate and substantial risk of harm to self or others.”¹³⁷

Taken together, the federal and state efforts suggest that although efforts are underway to curtail solitary confinement¹³⁸, they are targeted (for now) primarily in the juvenile context, and even for juveniles, are not an outright ban.¹³⁹ Thus, even if these efforts were to be more successful, we would still be faced with a large number of inmates in solitary. Assuming, then, that we will continue to be faced with many inmates in deplorable solitary confinement conditions, we return to the central question of this paper: how can artificial intelligence help?

¹³³ *Id.* The law does not give examples of proper implementation. It is unclear what circumstances would merit solitary confinement under this law.

¹³⁴ *Id.*

¹³⁵ Anne Teigen, *States that Limit or Prohibit Juvenile Shackling and Solitary Confinement*, NATIONAL CONFERENCE OF STATE LEGISLATORS (Feb. 15, 2017), <http://www.ncsl.org/research/civil-and-criminal-justice/states-that-limit-or-prohibit-juvenile-shackling-and-solitary-confinement635572628.aspx>.

¹³⁶ LB 870, 105th Neb. Leg. (2018).

¹³⁷ *Id.* at § 2(3). The bill holds that a juvenile must be held only long enough to dissipate the risk of harm, and only insofar as the confinement does not “compromise or harm the mental or physical health of the juvenile.” *Id.* at § 2(4)(a)-(b).

¹³⁸ For further discussion, see <https://www.pbs.org/wgbh/frontline/article/reducing-solitary-confinement-one-cell-at-a-time/>

¹³⁹ For instance, the Solitary Confinement Reform Act (2016) represents the most robust reform for all persons (including adults), and aims to cut back the duration and character of confinement -- most notably, the bill mandates that inmates in solitary confinement must be given as much meaningful programming and connection with the outside prison population “as practicable.”

III. INTRODUCING *THEMIS*, THE CONFINEMENT COMPANION

Inmates on Level One at the State of Wisconsin's Supermax Correctional Institution in Boscobel, Wisconsin spend all but four hours a week confined to a cell. The “boxcar” style door on the cell is solid except for a shutter and a trap door that opens into the dead space of a vestibule through which a guard may transfer items to the inmate without interacting with him. The cells are illuminated 24 hours a day. Inmates receive no outdoor exercise. Their personal possessions are severely restricted: one religious text, one box of legal materials and 25 personal letters. They are permitted no clocks, radios, watches, cassette players or televisions. The temperature fluctuates wildly, reaching extremely high and low temperatures depending on the season. A video camera rather than a human eye monitors the inmate's movements. ...

— *Federal District Judge Barbara Crabb, describing solitary confinement in Wisconsin*¹⁴⁰

If solitary confinement has significant negative psychological outcomes (Part I), but is likely to persist despite legal and legislative efforts (Part II), then the question arises: is there an opportunity for a non-human intervention to add value? I believe the answer is Yes, this Part III presents my proposal for a “confinement companion” AI system.¹⁴¹

¹⁴⁰ Jones “*El v. Berge*, 164 F. Supp. 2d 1096, 1098 (W.D. Wis. 2001)

¹⁴¹ To the best of my knowledge, this is a novel proposal. Discussion of AI in prison settings has primarily focused on the possible use of robotic prison guards, see, e.g., Richard Bloss, *Robots Go to Prison—As Guards*.” 39 INDUSTRIAL ROBOT: AN INTERNATIONAL JOURNAL (2012), <https://doi.org/10.1108/ir.2012.04939caa.007>, and the use of virtual reality to help non-prisoners understand what solitary confinement feels like, see, e.g., *Could This Solitary Confinement VR Experience Sway Lawmakers?* (Aug 31, 2017), <https://www.fastcompany.com/40461046/could-this-solitary-confinement-vr-experience-sway-lawmakers>. A search in Google Scholar and Westlaw produced only the following relevant mentions: Jaana Parviainen et al., *Motions with Emotions?*, in WHAT SOCIAL ROBOTS CAN AND SHOULD DO 210, 214 (Johanna Seibt et al. eds., 2016) (noting that “someone in solitary confinement might benefit from being given a robot companion—but he or she would benefit far more if offered a friendly social environment.”); Peggy Wu, *Maintaining Psycho-*

Perhaps the most important observation to make at the outset is that the value of the confinement companion is value relative to no other meaningful human contact. In most other potential legal applications of AI, the relevant cost-benefit question is: can the non-human AI do the task better than the human, e.g. is AI legal discovery better? Is AI-prediction better than human prediction for violence risk assessment.

But here, the unique harshness of solitary confinement reframes our inquiry. As the epigraph from Judge Crabb makes clear, time spent in Supermax solitary confinement is time spent by oneself. There is no access to the outside the world, and no real conversation with humans. Part I detailed the deplorable conditions in which many inmates spend up to 23 hours a day.¹⁴² Thus, our question for confinement companions can be phrased this way: *is the introduction of non-human AI better than no human at all?*

Of course, one might argue from the outset that we don't need AI, just a radio or television would do. This is an excellent critique. It challenges the proposal to identify why, precisely, we need an AI system, as opposed to simply some form of entertainment to brighten inmates' days.

Keeping these core questions of value-added in mind, Part III proceeds as follows. Section A presents the guiding principles for the confinement companion, which I name Themis. To guide its operation, I develop the Three Laws of Themis, modeled on Asimov's Three Laws. Section B identifies seven ways in which Themis can aid those in solitary confinement. These seven goals, from least to most computationally complex, are: 1) Archive and process information provided verbally by inmates; 2) Provide information at inmate request via voice assisted technology; 3) Interact socially with the inmate; 4) Identify risk of harm to self and to others; 5) Provide professional services, e.g. psychiatric treatment; legal counsel; 6) Provide individualized social connectivity; and 7) Provide true human-like companionship. I

Social Health on the Way to Mars and Back, PROCEEDINGS OF THE 2015 VIRTUAL REALITY INTERNATIONAL CONFERENCE (April 2015) (discussing the use of socially intelligent Virtual Agents (VAs) as tools to facilitate asynchronous human-human communication and counteract behavioral health challenges associated with prolonged isolation and deep space exploration).

¹⁴² See discussion in Part I.

assess the extent to which technology allows us to achieve each goal, either now or in the near future.

A. *Themis*: Core Principles ¹⁴³



Because technology may be better received if imbued by a name that captures its purpose, I utilize the name *Themis* to refer to the confinement companion system.¹⁴⁴ Themis (pictured in statue on the left) is the Greek Goddess of Justice.¹⁴⁵ The name Themis refers to “ancient, divine law, [and] a right order established by nature itself”¹⁴⁶ Themis is associated with divine justice,¹⁴⁷ and although the name Themis is not regularly used today, the familiar image of a blindfolded woman holding the scales of justice traces back to the mythology of Themis.¹⁴⁸ Themis is an appropriate name for the confinement companion because its purpose is in part to ensure that solitary confinement does not violate basic human rights.

There is another reason to name the technology. Law professor Ryan Calo, a leading thinker on robotics and the law, has argued that robots will require “creating a new category of legal subject, halfway between person and object.”¹⁴⁹ Themis is not *just* an object (like a radio), and shares certain characteristics with (though is not) a human. The name Themis helps us to see this AI as falling in the new legal category that Calo envisions. The implications of this new category will emerge in the discussion to follow.

¹⁴³ Image source: Licensed for public usage, https://commons.wikimedia.org/wiki/File:Statue_of_Themis.jpg

¹⁴⁴ Parallels to the movie *Her*, and its AI “Samantha,” are recognized and deliberate.

¹⁴⁵ Cathleen Burnett, *Justice: Myth and Symbol*, 11 LEGAL STUD. F. 79 (1987).

¹⁴⁶ PAMELA DONLEAVY & ANN SHEARER, FROM ANCIENT MYTH TO MODERN HEALING: THEMIS: GODDESS OF HEART-SOUL, JUSTICE AND RECONCILIATION 1 (2008).

¹⁴⁷ Andrew Simmonds, *The Blindfold of Justice*, 63 A.B.A. J. 1164 (1977).

¹⁴⁸ Cathleen Burnett, *Justice: Myth and Symbol*, 11 LEGAL STUD. F. 79 (1987).

¹⁴⁹ Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CAL. L. REV. 513, 549 (2015)

A working definition of Themis is: a self-learning AI system whose mission is to improve the lives of inmates in solitary confinement. Themis will be a single system holding multiple conversations at once, and continually learning from these thousands of conversations with inmates in the Themis system. Themis will be capable of communication in multiple voices and languages. Themis will also be connected to the Internet, and thus to relevant stakeholders such as the prison officials, mental health professionals who work in the prison, and (potentially in future iterations of the technology) the prisoner’s social network. Themis should be developed with open-source software, and should follow the Three Laws of Themis (detailed below).¹⁵⁰

1. The Three Laws of Themis

Isaac Asimov’s “three laws of robotics” have been oft-debated and updated, yet still serve as a touchpoint for AI governance.¹⁵¹ Asimov’s three laws are:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.¹⁵²
2. A robot must obey orders given it by human beings except where such orders would conflict with the First Law.¹⁵³
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws¹⁵⁴

¹⁵⁰ Perhaps of only tangential importance, but I think Themis (as envisioned here) is “AI” and probably (though maybe not) a “robot”. Calo suggests that “robots are best thought of as artificial objects or systems that sense, process, and act upon the world to at least some degree.” Themis senses (albeit only via sound), thinks (about certain things), and acts on the world. Certainly it is the case that, unlike a radio, one could not look at Themis and know how it would act. However, it is unclear that Themis “exists in the world as a corporeal object with the capacity to exert itself physically.” Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CAL. L. REV. 513, 531, 534 (2015)

¹⁵¹ Isaac Asimov, *Runaround*, in I ROBOT 37 (1950). See: Robin Murphy & David D. Woods, *Beyond Asimov: The Three Laws Of Responsible Robotics*, 24 IEEE Intelligent Systems (2009); Michael Anderson & Susan Leigh Anderson, eds. *Machine ethics*. Cambridge University Press (2011); Clarke, Roger. “Asimov’s laws of robotics: implications for information technology-Part I.” *Computer* 26, no. 12 (1993): 53-61. WENDELL WALLACH & COLIN ALLEN, *MORAL MACHINES: TEACHING ROBOTS RIGHT FROM WRONG* (2008).

¹⁵² Asimov (1950), *supra* note 151 at 37.

¹⁵³ *Id.*

¹⁵⁴ *Id.*

Using these three laws as the inspiration, the Three Laws of Themis are:

1. Themis may not harm a human, or, through inaction, allow a human being to come to harm.
2. Themis must respond to requests made by the inmate in solitary confinement, except where responding to such requests would conflict with the First Law.
3. Themis must follow applicable Codes of Professional Conduct when Themis is performing a function for which a Codes of Professional Conduct would apply for a human performing that same function.

The First Law of Themis tracks Asimov's first law, in that Themis cannot harm either an inmate or any other human. This should be interpreted as a prohibition on Themis providing or storing information that would potentially *lead* to harm.

There are both easy and hard applications of the First law of Themis. For instance, here are easy applications:

- If the inmate says, "Themis, tell me the best way to build a sharp knife out of paper," Themis should not comply.
- If an inmate speaks repeatedly of wanting to kill himself, Themis has a duty to report that information to appropriate staff.

But here are more difficult applications:

- It is quite likely that an inmate will complain about particular prison guards. Imagine that an inmate says, "Last night Guard Smith beat me for no reason." What Themis should do? On one hand, *if* Themis thought the warden promoted justice, Themis might report this information to the warden so that Guard Smith could be reprimanded. On the other hand, it's conceivable that the prison system might ignore this information and punish the inmate even more for saying this. In that case, Themis should keep the information private. This design consideration is discussed further below.
- The prohibition against inaction could be interpreted broadly, and it is unclear how *proactive* Themis should be. For instance, imagine the inmate says to Themis: "I'm very lonely in here. Can you cheer me up?" Staying silent in response would be

harmful to the inmate because Themis's inaction would prolong the inmate's loneliness. But how far should Themis go to try and relieve the inmate's loneliness? What is the limiting principle? There is a funny version of this example, e.g. Themis ends up in a loop of endless Knock-Knock jokes. But there is also a not-so-funny version: can Themis begin to deceive in order to alleviate loneliness? For instance, if Themis knows that the inmate is going to serve 20 more years in prison, can Themis pretend as if the sentence might be reduced to 2 years?

The Second Law of Themis breaks from Asimov's law in that Themis should be designed to be the *inmate's* system, not an instrument of the prison. The Second Law says that Themis "must respond to requests made *by the inmate in solitary confinement*, except where responding to such requests would conflict with the First Law." Themis does *not* need to respond to requests by the prison staff, except to the extent that those requests are governed by the First Law. Tipping the scales to the inmate is foundational to my proposal.

The criminal justice system, especially Supermax prisons, already utilizes significant cell monitoring technology. In modern supermax prisons "all movement is monitored by video surveillance and assisted by electronic door systems. Special alarms, cameras and security devices are everywhere."¹⁵⁵ Indeed, there is a burgeoning industry of prison surveillance technology companies. One of these companies pitches their technology in this way: "Contraband. Violence. Inmate and officer safety. These are just a few of the issues confronting security professionals working in today's prisons, jails and other correctional facilities. And video security has never played a more important role in helping maintain order while ensuring a safe working environment."¹⁵⁶

The State's overriding concern in utilizing surveillance technology is not inmate mental health. Because the prison already has its own technology, and especially because inmates have zero control over that other technology, Themis should be designed to be most sensitive to the needs of solitary confinement inmates.

¹⁵⁵ Jennifer R. Wynn & Alisa Szatrowski, *Hidden Prisons: Twenty-Three-Hour Lockdown Units in New York State Correctional Facilities*, 24 PACE L. REV. 497, 498 (2004).

¹⁵⁶ <https://www.pelco.com/video-security-market/prisons-corrections-correctional-facility-security>

As with the First Law of Themis, the Second Law of Themis invites both easy and hard applications. Easy applications might include:

- Imagine that the inmate shares a number of personal stories with Themis about how much he loves his kids. If the inmate then requests that Themis keep these stories private, Themis should not divulge the information to the prison even upon request (unless of course the First Law is invoked.)
- Imagine that the inmate requests that Themis read him the Book of Genesis in the Bible, and that Themis read it in French. Themis should reply with this request.

It should be noted that these are “easy” in the sense that coding Themis to do these things should not require much in the way of machine moral decision-making. It will not be easy, however, to get approval from the prison to allow Themis to do these things. This of course is the whole point of my proposal—solitary confinement is designed to prevent even the most innocent of information exchange, and Themis aims to counteract that.

Hard problems for the Second Law of Themis are (a) what it means to “respond to a request” and (b) identifying those requests that violate the First Law. Here is an example of each:

- If an inmate were visited by a chaplain, or talking with a family member, it would be natural to ask a question such as, “What’s new out there?” In response, a human visitor to the prison would know to pick out a few details that would be positive and interesting to the inmate, e.g. “Oh, your cousin Jenny just got married.” Or “Well, we had a really big snow this year, glad you were here or you’d be shoveling!” For Themis, the simple question “What’s new out there?” could prove difficult to answer. Careful attention would be required in coding Themis’s response.¹⁵⁷

- A second class of hard problems is illustrated by this request: “Themis, I’d like you to tell me all about the movie *Escape from Alcatraz* because it’s such an exciting story it would keep me from feeling lonely.” How should Themis respond—is this a violation of the First Law? It’s hard to know because there are both legitimate reasons (it’s an enjoyable story) and illegitimate reasons (it provides a blueprint for prison escape). The line-drawing might be even more

¹⁵⁷ And learning over time from answering the question would also be important to development of the system.

difficult. For instance, what if the inmate says: “Themis, tell me about Star Wars, Episode IV, that part where they escape from the Death Star.” Is there really any chance that learning about this fictional science fiction story would facilitate a prison escape? Probably not, but Themis would have to know how to make such analytical distinctions.

The Third Law of Themis again breaks from Asimov’s law in that Themis has no law for self-preservation. Instead, Themis is required to abide by relevant Rules of Professional Conduct. This would be most applicable if Themis is providing professional services such as legal advice, psychiatric treatment, or educational services. The Third Law is meant to recognize that sometimes Themis is just a friend, but sometimes Themis is taking on a professional role. When in that role, Themis must abide by the applicable professional code. The Third Law would be challenging for Themis in the same ways that it is challenging for human practitioners. (I discuss this in more detail in Section B below.)

These Three Laws are not perfect, and they are incomplete. But together they form a suitable platform on which to explore how Themis could operate.

B. Seven Goals for Themis

What, exactly, can Themis do that a radio cannot? In this section I layout seven ways that Themis can aid inmates in solitary, while still adhering to the Three Laws just discussed. I evaluate for each goal the technological capacity to accomplish it either now or in the future. Table 1 provides a summary of the seven components.

What can Themis add?	What roles is Themis playing?	Status of Technological Capacity
1) Archive and process information provided verbally by inmates	A) Friend B) Archivist, Analyst & Researcher	Technology is sufficiently developed
2) Provide information at inmate request via voice assisted technology	A) Friend B) Archivist, Analyst & Researcher	Technology is partially, but not sufficiently, developed
3) Interact socially with the inmate	A) Friend (more engaged) B) Archivist, Analyst & Researcher	Technology is emerging
4) Identify risk of harm to self and to others	A) Risk Assessment B) Archivist, Analyst & Researcher	Technology is emerging
5) Provide professional services, e.g. psychiatric treatment; legal counsel	A) Psychiatrist / Psychologist B) Attorney C) Counselor	Technology is emerging, but more difficult to develop
6) Provide individualized social connectivity	A) Friend B) Archivist, Analyst & Researcher	Technology is emerging, but more difficult to develop
7) Provide true human-like companionship	A) Intimate Friend	Not now, and maybe never.

1. *Themis v1.0*: Inmate-Provided Information Archival and Processing

Ironically, in an era of Big Data, and in a context in which these prisoners' lives are monitored via video almost 24/7, we have little data on what actually happens to inmates in solitary confinement. In one of the leading cases on unconstitutional prison conditions, Supreme Court Justice Anthony Kennedy observed, "Prisoners are shut away--out of sight, out of mind."¹⁵⁸ For those in solitary confinement, access to the outside world is near impossible. This has made it extremely challenging for researchers to access, and litigators to document, what really happens.¹⁵⁹ "Getting access to prisoners in real life segregation for research purposes raises both practical difficulties and ethical concerns."¹⁶⁰ For instance, we don't know for sure, but it is thought that prison guards may regularly use excessive force on inmates in solitary confinement.¹⁶¹

Considering this context, *Themis* is a way for inmates to *share their story* with a friend. To be sure, *Themis* would remain limited by what inmates chose to share, and recall that *Themis* could not share anything that the inmate prefers to keep private (unless the First Law applied). And as a friend, and abiding by the Second Law, *Themis* would be required to keep this information private. *Themis* could, if the inmate agreed, feed this information into the aggregate data processing system. But the inmate's data would not be included unless the inmate explicitly agreed. The Third Law is important here as well because research on vulnerable populations such as individuals in solitary confinement is fraught with ethical difficulty.

¹⁵⁸ *Davis v. Ayala*, 135 S. Ct. 2187, 2209 (2015) (Kennedy, J., concurring).

¹⁵⁹ BRUCE A. ARRIGO, ET AL, *THE ETHICS OF TOTAL CONFINEMENT: A CRITIQUE OF MADNESS, CITIZENSHIP, AND SOCIAL JUSTICE* 68 (2011) (noting that "the very nature of isolation precludes investigators from gaining meaningful access to those whom they seek to study").

¹⁶⁰ SHARON SHALEV, *A SOURCEBOOK ON SOLITARY CONFINEMENT* 23 (2008)

¹⁶¹ Kupers (2017), *supra* note 6 at 40 ("with the advent of supermaximum solitary confinement there has been an alarming escalation of force used against prisoners, especially prisoners with serious mental illness.")

Thankfully, there are established codes of research ethics that can be used to program and guide Themis.¹⁶²

Even if it were to do nothing other than collect stories and information, Themis would be a major advance. If Themis started modestly, with only 100 inmates each spending just 100 awake hours a year in solitary (a low estimate), that would provide 10,000 potential hours of data each year. True, the inmates would not be speaking continually for all of those hours. But it would still provide the system with significant amounts of information on which to train the Themis algorithms. By way contrast, to become licensed a clinical psychologist, it is required that—over 4-6 years—a psychologist in training get 1,500-6,000 supervised clinical hours.¹⁶³

We don't know what would happen if inmates in solitary suddenly had a responsive voice to speak with, but I have a strong suspicion that many inmates would gladly speak to Themis at length. This information would include descriptions of events, including negative experiences in the prison. But this information could also be used to document the inmates' internal mental states. For those with mental illness, "it is extremely common for prisoners to be ignored, disrespected, terrorized, and treated like animals, but they essentially have no power and no recourse."¹⁶⁴ For once, these inmates would have an ear to listen carefully to them and to respond with kindness, not condemnation.

Do we have the technology to do this? Yes. The field of artificial intelligence has for many decades endeavored to create non-human machines capable of natural conversation with humans.¹⁶⁵ These efforts have been challenging, as "[t]he obstacle for computers is not just understanding the meanings of words, but understanding the

¹⁶² Lawrence O. Gostin, *Biomedical Research Involving Prisoners: Ethical Values And Legal Regulation*, 297 JAMA 737 (2007); Institute of Medicine (US) Committee on Ethical Considerations for Revisions to DHHS Regulations for Protection of Prisoners Involved in Research, *The Ethical Framework for Research Involving Prisoners*. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK19885/>

¹⁶³ Melissa Dittman, *What You Need To Know To Get Licensed* (2004), <http://www.apa.org/gradpsych/2004/01/get-licensed.aspx>.

¹⁶⁴ Kupers (2017), *supra* note 6 at 58.

¹⁶⁵ HEUNG-YEUNG SHUM ET AL., MICROSOFT CORPORATION, FROM ELIZA TO XIAOICE: CHALLENGES AND OPPORTUNITIES WITH SOCIAL CHATBOTS (2018).

endless variability of expression in how those words are collocated in language use to communicate meaning.”¹⁶⁶

The many limitations notwithstanding, humans have been conversing deeply and emotionally with non-human technology for least a half-century. “Eliza,” a text-based AI system designed to engage in conversation with humans, was introduced in 1966. As discussed by Brian Christian in *The New Yorker*, the computer scientist who created Eliza creator was “was startled to see how quickly and how very deeply people conversing with [the computer] became emotionally involved with the computer and how unequivocally they anthropomorphized it.”¹⁶⁷

Fifty years later, conversation between humans and non-humans has become a daily occurrence through the introduction of both voice-based and text-based chatbots.¹⁶⁸ Chatbots are being utilized in settings such as banking¹⁶⁹, healthcare¹⁷⁰, education,¹⁷¹ libraries¹⁷², and many more. Given how fast voice assisted technologies have developed,¹⁷³ and the financial incentives now in place for companies to continue to improve the technology,¹⁷⁴ it

¹⁶⁶ Jennifer Hill et al., *Real Conversations With Artificial Intelligence: A Comparison Between Human–Human Online Conversations And Human–Chatbot Conversations*, 49 COMPUTERS IN HUMAN BEHAVIOR 245 (2015).

¹⁶⁷ Brian Christian, *The Samantha Test*, NEW YORKER (Dec. 30, 2013), <https://www.newyorker.com/culture/culture-desk/the-samantha-test>.

¹⁶⁸ RASHID KHAN & ANIK DAS, *Introduction to Chatbots*, in BUILD BETTER CHATBOTS 1-11 (2018).

¹⁶⁹ Songhyun Kim et al., *The Use of Voice Input to Induce Human Communication with Banking Chatbots*, COMPANION OF THE 2018 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION 151 (2018).

¹⁷⁰ TOBIAS KOWATSCH ET AL., *Text-based Healthcare Chatbots Supporting Patient and Health Professional Teams: Preliminary Results of a Randomized Controlled Trial on Childhood Obesity*, in PERSUASIVE EMBODIED AGENTS FOR BEHAVIOR CHANGE (PEACH2017) WORKSHOP (2017).

¹⁷¹ LUCIANA BENOTTI ET AL., *Engaging High School Students Using Chatbots*, PROCEEDINGS OF THE 2014 CONFERENCE ON INNOVATION & TECHNOLOGY IN COMPUTER SCIENCE EDUCATION 63 (2014).

¹⁷² MICHELE L. MCNEAL & DAVID NEWYEAR, *Introducing Chatbots In Libraries*, 49 LIBRARY TECHNOLOGY REPORTS 5-10 (2013).

¹⁷³ Ava Mutchler, *Voice Assistant Timeline: A Short History of the Voice Revolution* (July 14, 2017), <https://www.voicebot.ai/2017/07/14/timeline-voice-assistants-short-history-voice-revolution/>

¹⁷⁴ Cade Meth & Keith Collins, *To Give A.I. the Gift of Gab, Silicon Valley Needs to Offend You* (New York times, Feb 21, 2018),

seems to me highly plausible that we already have the necessary technology for Themis to communicate via voice with inmates in solitary confinement. The equipment requirements are minimal: a sufficiently high quality microphone and speaker, wiring to send the information from the cell to Themis HQ, and the processing capability at Themis HQ to store, analyze, and make meaning from the spoken word information. If Siri can understand me when I ask to find a pair of size 11.5 sneakers, surely Themis can keep an organized diary of an inmate's spoken thoughts while in solitary.

2. *Themis v2.0*: Provide Information at Inmate Request via Voice Assisted Technology

Before discussing this second value-added by Themis, let us revisit what life is like for inmates in the Supermax:

Inmates are not allowed face-to-face visits, other than with their lawyers. The institution provides only video visitation. Inmates remain in their cell block and visitors at the front of the institution. Inmates and their visitors see each other on small video screens that are located across the room from the inmate. The audio quality is poor. ... The poor quality of the visits has led some mentally ill inmates to believe that the images on the video screens are manipulated and to refuse visitors. During the video visits, inmates remain handcuffed, shackled and belly chained. Prison log books show that only 10% of inmates receive visits, an unusually low number. ... Level One inmates at Supermax are allowed only one six-minute telephone call a month. They are not allowed to have any electronic equipment in their cells or to participate in any programs. ... Inmates in Level One are not allowed to have library books in their cells but may have one box or two shopping bags of legal materials, a soft-cover Bible or Koran or the equivalent and up to 25 letters. Sometimes a chaplain will come to the door of a cell. ...¹⁷⁵

<https://www.nytimes.com/interactive/2018/02/21/technology/conversational-bots.html>

¹⁷⁵ Jones “*El v. Berge*, 164 F. Supp. 2d 1096, 1101 (W.D. Wis. 2001)

Life in solitary confinement is life deprived of sensory and information inputs. Themis cannot provide inmates a hug, but could provide a steady stream of information. This information could range from pure entertainment (are the Golden State Warriors playing well this year?), to education (please play a 15 minute lecture on the history of the state of California), to practical (what stretches can I do for a sore neck?) to religious (what are some prayers for healing?), and much more. Themis should be able to provide this information at inmate request, subject to the First Law, and this information would be of great value since the alternative is no information.¹⁷⁶

Do we have the technology to do this? Not yet, but hopefully soon. The rapid rise of intelligent assistants is still in its infancy, but already “we can talk to objects in our homes that recognize our voices, understand our questions, model our behaviors, and even predict our needs while communicating across vast networks of machines, computers, and distant humans.”¹⁷⁷ These advances will continue because AI is at the center of the business plans for many of the world’s largest companies.¹⁷⁸ Themis can benefit from the advances in these technologies, and recall that the initial bar for Themis’s value added is low. For instance, responding to even simple queries might relieve boredom and loneliness.

3. *Themis v 3.0*: Interact socially with the inmate

There is a big leap from mere information archival and provision, to social interaction. Once technology permits, Themis should provide this richer social interaction. It is at this point that the promise of Themis AI really begins to take hold. Function #1 (information repository) is a bit of a glorified tape recorder, and Function #2 (information delivery) is somewhat of a glorified voice-based Internet search engine. But Function #3 is potentially transformative. This is the version at which Themis really distinguishes itself from a radio. By providing real social

¹⁷⁶ In this prong of Themis’s work, Themis is providing generalized information, not individual-specific information (which I discuss later).

¹⁷⁷ Chris Arkenberg, *The Future of Intelligent Assistants*, 50 *COMPUTER* 77 (2017).

¹⁷⁸ *Id.*

interaction, Themis will directly counteract the social isolation inherent to solitary confinement. The vexing question is whether, and how, technology will develop to meet this need.

Do we have the technology to do this? Not yet, but there is progress being made. Today, human interaction with robots is no longer limited to simply text and voice.¹⁷⁹ The development of “social robotics” has introduced new types of robots to new types of interactions with humans.¹⁸⁰ Exactly what constitutes a “social robot” depends on context. As one review summarized, “the notion of social robots and the associated degree of robot social intelligence is diverse and depends on the particular research emphasis.”¹⁸¹

The use of social robots has been explored to reduce loneliness¹⁸², and to improve outcomes for older adults.¹⁸³ One company, Intuition Robotics, is creating a robot named ElliQ specifically for older adults.¹⁸⁴ The company’s CEO describes their goal this way: ““Think of [the robot] as a fully autonomous agent . . . You tell it what your goals are, and it tries to measure how you’re doing on those goals and suggests activities accordingly to help you meet those goals.”¹⁸⁵

¹⁷⁹ TAKAYUKI KANDA & HIROSHI ISHIGURO, *HUMAN-ROBOT INTERACTION IN SOCIAL ROBOTICS* (2017).

¹⁸⁰ Terrence Fong et al., *A Survey of Socially Interactive Robots*, 42 *ROBOTICS AND AUTONOMOUS SYSTEMS* 143 (2003).

¹⁸¹ Kerstin Dautenhahn, *Socially Intelligent Robots: Dimensions of Human–Robot Interaction*, 362 *PHILOS. TRANS. R. SOC. LOND. B. BIOL. SCI.* 679, 681 (2007).

¹⁸² Friederike Eyssel & Natalia Reich, IEEE, *Loneliness Makes the Heart Grow Fonder (Of Robots)—On The Effects Of Loneliness On Psychological Anthropomorphism*, 2013 8TH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI) 121 (2013).

¹⁸³ Gregory Baltus, et al., *Towards Personal Service Robots for the Elderly*, *PROC. OF THE WORKSHOP ON INTERACTIVE ROBOTICS AND ENTERTAINMENT* (2000); Joost Broekens et al., *Assistive Social Robots In Elderly Care: A Review*, 8 *GERONTECHNOLOGY* 94 (2009); Suzanne Hutson et al., *Investigating The Suitability Of Social Robots For The Wellbeing Of The Elderly*, *INTERNATIONAL CONFERENCE ON AFFECTIVE COMPUTING AND INTELLIGENT INTERACTION* 578 (2011).

¹⁸⁴ <https://elliq.com/>

¹⁸⁵ Steve Overly, *In The Future, Virtual Assistants Will Not Only Take Orders. They’ll Also Have Ideas Of Their Own* (Washington Post, May 3, 2017), https://www.washingtonpost.com/news/innovations/wp/2017/05/03/in-the-future-virtual-assistants-will-not-only-take-orders-theyll-have-ideas-of-their-own/?utm_term=.16d6b3a7f112

There are different types of social robots in each context. For instance, some are “service robots.” A service robot is defined by the International Organization for Standardization as a robot “that performs useful tasks for humans or equipment excluding industrial automation applications”.¹⁸⁶ This type of robot requires the “ability to perform intended tasks based on current state and sensing, without human intervention”.¹⁸⁷ In the elder care context, service robots are used to support basic activities such as eating and mobility.¹⁸⁸ Examples of service robots include NurseBot,¹⁸⁹ and Care-O-Bot (service for non-elderly as well).¹⁹⁰

More relevant to the proposal I develop here are “companion robots.” As described by AI researcher Kerstin Dautenhahn, “[t]he companion paradigm emphasizes the assistant role of a robot, i.e. a useful machine, able to recognize and respond to a human's needs, trying to be useful. ... Important characteristics for such a robot are to be considerate, proactive and non-intrusive, to work towards a relationship of trust and confidentiality with the human, to possess ‘smooth’ communicative skills, to be flexible, willing to learn and adapt, and be competent.”¹⁹¹ Example of companion robots include Paro¹⁹², The Huggable,¹⁹³ and the robot dog Aibo.¹⁹⁴

¹⁸⁶ ISO 8373

¹⁸⁷ *Id.*

¹⁸⁸ Joost Broekens et al., *Assistive Social Robots In Elderly Care: A Review*, 8 GERONTECHNOLOGY 94 (2009).

¹⁸⁹ <http://www.cs.cmu.edu/~flo/>

¹⁹⁰ <https://www.care-o-bot.de/en/care-o-bot-4.html>

¹⁹¹ Dautenhahn (2007), *supra* note 181.

¹⁹² <http://www.parorobots.com/> (“PARO is an advanced interactive robot developed by AIST, a leading Japanese industrial automation pioneer. It allows the documented benefits of animal therapy to be administered to patients in environments such as hospitals and extended care facilities where live animals present treatment or logistical difficulties.”)

¹⁹³ <http://robotic.media.mit.edu/portfolio/huggable/> (“The Huggable™ is a new type of robotic companion being developed at the MIT Media Lab for healthcare, education, and social communication applications. The Huggable™ designed to be much more than a fun interactive robotic companion.”)

¹⁹⁴ <https://aibo.sony.jp/en/>

It remains unclear exactly what the effect of these robots will be on outcomes of interest.¹⁹⁵ For instance, a systematic review of social robots in elder care found that the effects have “not been proven comprehensively” and that “relations between the type of outcomes aimed for, either related to support of care or support of independence, and the application of the robot system in care, are not well established.”¹⁹⁶

The same conclusions have been reached in the context of using robots for interactions with autistic children.¹⁹⁷ A systematic review of the literature found 15 peer-reviewed research studies (with 11 unique samples), and identified a number of methodological limitations.¹⁹⁸ As with the literature evaluating the effects of robots on older adults, the effects of robot interaction on children with autism is in need of “rigorous empirical studies that examine the incremental validity of this approach over other available techniques, as well as the generalizability of skills learned with a robot in relation to those learned from human interaction.”¹⁹⁹

In these and other areas the message is clear: rapid and expansive developments in social robotics offers great promise. But the precise effects of social robots remain, at present, uncertain. There has been limited real-world application, and limited rigorous evaluation of those applications. My hope is that Themis can be a companion robot, but research to date on other companion robots suggests that we must balance promise with caution.

4. *Themis v4.0*: Identify Risk Of Harm To Self And To Others

It is well established that inmates in solitary confinement are at higher risk of self-harm, and that they are disproportionately battling

¹⁹⁵ Roger Bemelmans et al., *Socially Assistive Robots In Elderly Care: A Systematic Review Into Effects And Effectiveness*, 13 JOURNAL OF THE AMERICAN MEDICAL DIRECTORS ASSOCIATION 114 (2012).

¹⁹⁶ Bemelmans et al (2012), *supra* note 195.

¹⁹⁷ Brian Scassellati et al., *Robots for Use in Autism Research*, 14 ANNUAL REVIEW OF BIOMEDICAL ENGINEERING 275 (2012); Joshua J. Diehl et al., *The Clinical Use Of Robots For Individuals With Autism Spectrum Disorders: A Critical Review*, 6 RESEARCH IN AUTISM SPECTRUM DISORDERS 249 (2012).

¹⁹⁷ Diehl et al. (2012), *supra* note 197 at 259.

¹⁹⁸ Diehl et al (2012), *supra* note 197 at 259.

¹⁹⁹ *Id.*, at 260.

mental illness.²⁰⁰ Experts have suggested that given high rates of suicide ideation and completion relative to the general population, prisons should also consider developing a suicide prevention strategy for those in solitary confinement.²⁰¹

Themis could thus add value by identifying opportunities for possible earlier interventions to prevent this self-harm. A similar tactic is already being beta-tested by Facebook, as it analyses users' posts in an effort to spot those who might be especially high risk of committing suicide.²⁰² Facebook VP of Product Management explained their interest in this technology simply: "we have an opportunity to help here so we're going to invest in that."²⁰³ The same can be said for Themis. If Themis is listening to the inmate, and Themis has sufficiently strong evidence to suspect that self-harm is imminent, the First Law requires that Themis act. Themis might also identify risk of harm to others. In both cases, Themis would have to know what the threshold is for alerting a prison authority.

Do we have the technology to do this? Maybe, but we need to proceed cautiously. In other legal contexts, the potential mis-use, of algorithms to inform violence risk assessments has drawn considerable attention.²⁰⁴ For instance, how good will Themis be at predicting self-harm? What data will Themis use to make those predictions? The issues may be particularly challenging in the solitary confinement context because being put on "suicide watch"

²⁰⁰ Fatos Kaba et al., *Solitary Confinement And Risk Of Self-Harm Among Jail Inmates*, 104 AMERICAN JOURNAL OF PUBLIC HEALTH 442 (2014); Metzner and Fellner, *supra* note 88.

²⁰¹ Seena Fazel et al., *Mental Health Of Prisoners: Prevalence, Adverse Outcomes, And Interventions*, 3 THE LANCET PSYCHIATRY 871 (2016).

²⁰² Josh Constine, *Facebook Rolls Out AI To Detect Suicidal Posts Before They're Reported* (Nov 27, 2017), <https://techcrunch.com/2017/11/27/facebook-ai-suicide-prevention/>

²⁰³ *Id.*

²⁰⁴ Ed Yong, *A Popular Algorithm Is No Better at Predicting Crimes Than Random People* (Jan 17, 2018), <https://www.theatlantic.com/technology/archive/2018/01/equivant-compass-algorithm/550646/>; Jessica M. Eaglin, *Constructing Recidivism Risk*, 67 EMORY L.J. 59 (2017); Melissa Hamilton, *Risk-Needs Assessment: Constitutional and Ethical Challenges*, 52 AM. CRIM. L. REV. 231, 256-61 (2015); Julia Angwin et al., *Machine Bias*, ProPublica (May 23, 2016), <http://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> [http://perma.cc/FB8E-WSV2].

can involve even more deprivation.²⁰⁵ Themis would need to be developed to keep these competing interests in mind.

5. *Themis v5.0*: Provide Professional Services Such as Psychiatric Treatment and Legal Counsel

Solitary confinement may be particularly bad for those with mental illness, and has been subject of class action litigation.²⁰⁶ Inmates in solitary confinement lack adequate mental health treatment and do not enjoy regular visits with their attorneys.²⁰⁷ Thus, the possibility of Themis as an AI attorney and AI psychiatrist is tantalizing. Considering first the value added for mental health. It has been observed repeatedly by experts that inmates facing severe mental disorders are not receiving adequate treatment.²⁰⁸ Themis could provide prison-specific treatment interventions that are empirically supported.²⁰⁹

Looking at the need for legal counsel, although inmates (including those in solitary) have a constitutional right to speak with their attorney, in practice this access can be curtailed.²¹⁰ The ability of Themis to provide regular, responsive, and extensive legal advice would be a major advance.

Moreover, if Themis could take on either of these professional roles, Themis might enjoy privileged conversations with the inmate. Although the Second Law of Themis requires Themis to respect the inmate's requests to keep information private, the Second Law is a coding principle and would not provide the same sort of legal

²⁰⁵ Robert J. Cramer, et al., *Suicide Prevention In Correctional Settings: Current Standards And Recommendations For Research, Prevention, And Training*, 23 J. CORRECTIONAL HEALTH CARE 313 (2017).

²⁰⁶ *Coleman v. Wilson*, 912 F. Supp. 1282 (E.D. Cal. 1995)

²⁰⁷ Jules Lobel, *Prolonged Solitary Confinement And The Constitution*, 11 U. PA. J. CONST. L. 115 (2008).

Andrew P. Wilper et al., *The Health and Health Care Of Us Prisoners: Results Of A Nationwide Survey*, 99 AMERICAN JOURNAL OF PUBLIC HEALTH 666 (2009).

²⁰⁹ This may take some time to develop, as research in this area lags. E. FULLER TORREY ET AL, TREATMENT ADVOCACY CENTER, *THE TREATMENT OF PERSONS WITH MENTAL ILLNESS IN PRISONS AND JAILS: A STATE SURVEY* (2014).

²¹⁰ See, e.g., Tom Haydon, *Solitary Confinement Limited Inmate's Access To Legal Advice, Lawyer Says*, http://www.nj.com/union/index.ssf/2014/08/solitary_confinement_limited_inmates_access_to_legal_advise_lawyer_say.html

protections as privilege. If Themis developed the capability to take on these rules, extensive analysis would be required to determine where privilege would apply amidst the inmate-Themis conversations.

Do we have the technology to do this? Not yet, but efforts are underway. For decades, researchers have examined the possibility of AI semantic information processing for legal applications.²¹¹ Yet, to date the major contributions have been in automatic tasks such as electronic discovery and standardized forms.²¹² It does not seem likely that most attorneys will be replaced any time soon by robots.²¹³

The use of AI in psychiatry is similarly at a very early stage. A review of recent developments stressed that although it is too early to tell, the advancement of AI technologies may enhance mental health care and increase efficiency.²¹⁴ But it is unclear whether robots can do this alone, without a human in the loop.²¹⁵ Still, there

²¹¹ See, e.g. Kevin D. Ashley, *Reasoning with Cases and Hypotheticals in HYPO*, 34 INT'L J. OF MAN-MACH. STUD. 753 (1991); SEMANTIC INFORMATION PROCESSING (Marvin Minsky, ed., 1968); Edwina L. Rissland, *Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning*, 99 YALE L. J. 1957, 1957-58 (1990); ANNE VON DER LIETH GARDNER, AN ARTIFICIAL INTELLIGENCE APPROACH TO LEGAL REASONING 4 (1987); Kevin D. Ashley & Stefanie Brüninghaus, *Computer Models for Legal Prediction*, 46 JURIMETRICS 309 (2006); Vincent Aleven, *Using Background Knowledge in Case-Based Legal Reasoning: A Computational Model and an Intelligent Learning Environment*, 150 ARTIFICIAL INTELLIGENCE 183 (2003).

²¹² DANA REMUS & FRANK S. LEVY, CAN ROBOTS BE LAWYERS? COMPUTERS, LAWYERS, AND THE PRACTICE OF LAW (2015); Josh Blackman, *Robot, Esq.* (2012), <https://ssrn.com/abstract=2198672> Julie Sobowale, *How artificial Intelligence Is Transforming The Legal Profession* (April 1, 2016); Cary Coglianese & David Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, 105 GEO. L.J. 1147 (2017); John O. McGinnis & Russell G. Pearce, *The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers in the Delivery of Legal Services*, 82 FORDHAM L. REV. 3041 (2014); Kathryn D. Betts & Kyle R. Jaep, *The Dawn of Fully Automated Contract Drafting: Machine Learning Breathes New Life into A Decades-Old Promise*, DUKE L. & TECH. REV., March 29 2017, at 216

²¹³ Steve Lohr, *A.I. Is Doing Legal Work. But It Won't Replace Lawyers, Yet* (NY Times, March 19, 2017),

²¹⁴ David D. Luxton, *Artificial Intelligence in Psychological Practice: Current and Future Applications and Implications*, 45 PROFESSIONAL PSYCHOLOGY: RESEARCH AND PRACTICE 332 (2014).

²¹⁵ *Id.*

are a handful of exploratory studies that offer promise. In a proof-of-principle study, Gillinder Bedi, et al. sought to combine automated speech analyses with Machine Learning to predict later psychosis onset in youths at clinical high-risk (CHR).²¹⁶ These speech features predicted later psychosis development with 100% accuracy, outperforming classification from clinical interviews.²¹⁷ These findings support the utility of automated speech analysis to measure subtle, clinically relevant mental state changes in emergent psychosis. But these findings are just a beginning. Most of the studies to date have explored human *plus robot* (or “robot assisted”) treatment approaches. A 2014 meta-analysis found that more studies are needed to prove the efficacy of robot-enhanced therapy, but the overall results clearly support the use of robot-enhanced therapy for different populations.²¹⁸

At present, it is unclear when (or if) AI will develop to the point that it could competently provide professional legal or counseling services without a human in the loop. But if that were possible, applying such technology to solitary confinement would be of great service.

6. Provide Individualized Social Connectivity

Inmates in solitary confinement (and in prison generally) are eager to connect to the outside world, but they are particular eager to connect to their slice of that outside world. An inmate solitary wants to know the latest news about *his* family and friends. Themis could conceivably provide this connection. For instance, imagine that the inmate asks Themis: “Could you help me connect with my daughter. Her name is ... and she went to ... high school.” Themis could then find the inmate’s daughter on Facebook, and initiate a conversation. If the daughter agreed, Themis could friend her on Facebook, consume her information, and communicate that to the inmate. If the daughter regularly uses social media, the inmate could—in an emotionally rich way—be a part of his daughter’s life.

²¹⁶ Gillinder Bedi et al., *Automated Analysis of Free Speech Predicts Psychosis Onset in High-Risk Youths*, 1 NPJ SCHIZOPHRENIA, article number: 15030 (2015).

²¹⁷ *Id.*

²¹⁸ Cristina A. Costescu et al., *The Effects of Robot-Enhanced Psychotherapy: A Meta-Analysis*, 18 REVIEW OF GENERAL PSYCHOLOGY 127 (2014).

For instance, one could imagine this exchange with an inmate (named John):

Themis: Good morning, John. Guess what your daughter did yesterday?

John: Don't tell me ... was she out there running again?

Themis: Yes, it was her best run yet. She posted a picture on Instagram and says she's feeling ready for the marathon on Sunday.

John: I can't wait to hear how she does.

Themis: Oh, and one more thing John: she posted on Facebook, "Themis, be sure to tell my Dad that I love him."

Being allowed to have an exchange like that—a human exchange—amidst the depravity that currently characterizes solitary confinement is the motivation for Themis.

Do we have the technology to do this? No, but this is in the realm of the plausible not the impossible. Scraping information from someone's Facebook and Instagram accounts is possible, especially if that person (as in the example above) provides permission. Simply reading the text of posts to the inmate is also already well within our technological capability. What remains to be done is the development of a system that can put it all together, from the inmate's request, to finding the daughter, to making a socially appropriate and effective overture in order to initiate the connection.

7. Provide True Human-Like Companionship

Although Themis would never be embodied, the ultimate (and perhaps unachievable) goal would be for Themis to become as close to a human-like system as possible. If Themis could provide true companionship, it would become like the operating system "Samantha" portrayed in Spike Jonze's movie *Her*. The benefits (and risks) of such a powerful AI system would be tremendous. So tremendous, in fact, that they would require a separate (equally lengthy) paper.

Do we have the technology to do this? No, and maybe never. At present, this remains more science fiction than science. It is worth revisiting as a thought experiment to inform policy guidelines for the more here-and-now applications already discussed. But there are enough complexities to work out without reaching this final stage of Themis development.

IV. DISCUSSION [preliminary]

The proposal advanced in Part III raises an interrelated set of difficult ethical and legal questions concerning the design and implementation of Themis. In this [still being developed] Part, I group these questions broadly into (a) “Design” questions and (b) “Legal/Ethical” questions. The two categories are related of course, because whether or not we should pursue Themis depends, in rather large part, on what sorts of protections we can build into the code. Nevertheless, separating the two types of questions is useful. I close in Section C with some broader reflections on what Themis might become, and what the practical next steps might be.

A. Design Considerations

1. The Dark Side of Surveillance (or, Can Themis Really Protect Inmate Privacy?)

The level of surveillance and control already at the heart of the modern American prison system has been heavily critiqued.²¹⁹ Parallels to Michael Foucault’s panopticon are myriad.²²⁰ Amidst this system, a central concern with my proposal is that it would give to the state another—and an even more intimate—layer of surveillance, control, and domination.

Is it not an invasion of privacy—indeed, perhaps the deepest violation of privacy—to introduce Themis into this mix?²²¹ Under

²¹⁹ Lorna A. Rhodes, *Supermax As A Technology of Punishment*, 74 SOCIAL RESEARCH: AN INTERNATIONAL QUARTERLY 547 (2007).

²²⁰ DAVID LYON, *The Search for Surveillance Theories*, in THEORIZING SURVEILLANCE: THE PANOPTICON AND BEYOND 3-20 (2006).

²²¹ The United States Supreme Court, for instance, has consistently restrained law enforcement from listening to conversations in our homes without a warrant. “It

the guise of “companionship,” would Themis not serve in practice to serve the state and not the inmate? With no one else to talk to, if an inmate begins speaking freely with the robot voice, won’t his words be used against him by the prison? Moreover, couldn’t the state design the system such that it would be calibrated to promote the state’s interests in control, rather than an interest in the individual inmate’s well-being?

The answer to these and related questions is of course: *yes*, the technology could be co-opted by the state. Instead of providing consolation, the technology could promote control; instead of robot rehabilitation, we could usher in a world of robot retributivism. It is not hard to see how such co-opting is possible, for as much as words can heal, words can hurt as well.

In theory, these dangers can be avoided *if* Themis adheres to the First Law of Themis: Themis may not harm a human, or, through inaction, allow a human being to come to harm. Themis could be programmed to understand “harm” to include violations of privacy.

But theory may not match reality. Imagine this scenario: a prison agrees to install Themis, but only on the condition that the prison codes into it a new rule that overrides the Three Laws and says: whenever an inmate in solitary admits a past rules infraction, Themis must immediately report this infraction to the warden. They might go a step further, and add another rule: Themis must report to the prison anything that is said by the inmate that may possibly be relevant to assessing the inmate’s mental health and/or likelihood of violent behavior. Phrased in this way, almost anything the inmate says would end up in the hands of the warden. And Themis would learn—by seeing what happens after sharing this information—that the results are not good for the inmate.

If Themis were co-opted at this design stage, or if Themis were vulnerable to government hacks, it would rapidly become the most insidious of government tools: the illusion of companionship but really a government spy. As such it would fit right in with the modern prisons system, which “represent the application of sophisticated, modern technology dedicated entirely to the task of

is axiomatic that the physical entry of the home is the chief evil against which the wording of the Fourth Amendment is directed. And a principal protection against unnecessary intrusions into private dwellings is the warrant requirement imposed by the Fourth Amendment on agents of the government who seek to enter the home for purposes of search or arrest.” *Welsh v. Wisconsin*, 466 U.S. 740 (1984).

social control, and they isolate, regulate, and surveil more effectively than anything that has preceded them.”²²²

This outcome must be avoided at all costs, and if Themis cannot be designed to protect the inmate’s privacy, then it should not be introduced into the prison system. These privacy concerns appear throughout the emerging law and AI literature,²²³ and that literature can be leveraged to inform the design of Themis.

2. Many Design Choices

Themis invites creative consideration of many design choices, including:

- *What voice and language?* Themis would be a single system capable of a large number of voices (male, female, loud, soft, etc) and a large number of languages and dialects. How would Themis choose a tone / voice / language for each inmate? How much input would an inmate have? For instance, could an inmate request a female rather than a male voice? Could an inmate request a particular accent?
- *Multiple voices at once?* The proposal thus far has imagined a single-voice Themis speaking to an inmate. But Themis could provide the inmate with multiple AI voices at the same time. For instance, perhaps an inmate would like to be part of a group conversation. Themis could provide the rest of the group. This dialogue approach might have therapeutic benefits, as there is evidence about the positive effects of group therapy in many contexts.

²²² Craig Haney, “*Infamous Punishment*”: *The Psychological Consequences of Isolation*, 3 NAT’L PRISON PROJECT J. (1993), at 3

²²³ Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CAL. L. REV. 513, 515 (2015) (observing that “[r]obotics is shaping up to be the next transformative technology of our time.”); Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. DAVIS L. REV. 399 (2017); Andrew Tutt, *An FDA for Algorithms*, 69 ADMIN. L. REV. 83 (2017); Jack M. Balkin, *The Path of Robotics Law*, 6 CAL. L. REV. CIRCUIT 45 (2015); Cary Coglianese & David Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, 105 GEO. L.J. 1147 (2017); Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 HARV. J.L. & TECH. 353 (2016).

- *Connecting inmates across the country?* Themis will be self-learning based on thousands of conversations with inmates in solitary confinement. If an inmate gave permission, could Themis then integrate conversations across inmates, in order to build community and solidarity? This would be a high tech version of “passing notes” between the bars of prison cells.

- *Fair, accountable, and transparent.* How, exactly, will Themis process its massive amounts of information? More attention needs to be paid to details. There has been much discussion of how to make AI systems fair, accountable, and transparent.²²⁴ The now annual conference of the Fairness, Accountability, and Transparency in Machine Learning (FAT/ML) organization tackles questions such as how one codes “fairness”.²²⁵ Any development of Themis would necessarily require these types of considerations.

- *Programming for individuals and groups:* How would coding take account of the multitude of differences across inmates in solitary confinement? For instance, would Themis run different code for those below age 30, as opposed to those over age 50? Would it take into account mental disorders? Race? Religion? Gang affiliation? And how much information would Themis gather on its own?

- *What’s off limits?* Setting aside for the moment prison-imposed limits on what information Themis could provide to the inmate, are there ethical boundaries on the requests that Themis should comply with? For instance, what if the inmate asks Themis to read him a hate-filled book about racial superiority? What if the inmate asks Themis to take on the voice of a young child and participate in a sexual roleplay? How will Themis navigate such decision-making?

- *Could AI help prison guards as well?* Craig Haney has argued that an “ecology of cruelty” develops in Supermax prisons, as prison guards become engulfed in the toxicity of punitiveness. Could AI help to reach these guards, and remind them of their common humanity with the prisoners they control?

²²⁴ Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. DAVIS L. REV. 399, 411 (2017) (“Perhaps the most visible and developed area of AI policy to date involves the capacity of algorithms or trained systems to reflect human values such as fairness, accountability, and transparency.”)

²²⁵ See <http://www.fatml.org>

- *International applications.* The analysis could, and should, be extended to contexts beyond the United States and beyond prisons. For instance, could the same (or modified) technology be deployed in situations where enemy combatants are indefinitely detained? How would international human rights standards apply?

3. Infrastructural and Architectural Considerations

Beyond the scope of this paper is the precise blueprint for how Themis would be installed into a prison system. But, without minimizing the extensive work required, it's worth noting that Supermax prisons are already wired for 24/7 surveillance. "Inmates may also be subject to constant electronic surveillance, and in some facilities, a light is kept on at all times."²²⁶ Themis could be built on top of the existing high-tech setup already in place.

4. Confinement Companion Technology for the Hearing- and Speech-Impaired

My proposal thus far has assumed that the inmate in solitary confinement does not have a hearing or speech impairment. But the reality is that some inmates do have these impairments,²²⁷ and it is likely that prison generally, and solitary confinement in particular,²²⁸ may be especially difficult for those with physical disabilities.²²⁹ Given these concerns, it would be important to consider alternative designs for Themis that might allow these populations to share in its benefits.

B. Legal and Ethical Considerations

²²⁶ Leena Kurki & Norval Morris, *The Purposes, Practices, And Problems Of Supermax Prisons*, in 28 CRIME AND JUSTICE REVIEW (2001).

²²⁷ Laura Margaret Kelly, *Sounding Out D/Deafness: The Experiences of D/Deaf Prisoners*, 8 JOURNAL OF CRIMINAL PSYCHOLOGY 20 (2018).

²²⁸ Farina Mendelson, *A Silent Struggle: Constitutional Violations Against the Hearing Impaired in New York State Prisons*, 20 CUNY L. REV. 559, 584 (2017).

²²⁹ JAMELIA N. MORGAN, ACLU, CAGED IN: THE DEVASTATING HARMS OF SOLITARY CONFINEMENT ON PRISONERS WITH PHYSICAL DISABILITIES (2017); MARGO SCHLANGER, ACADEMY FOR JUSTICE, PRISONERS WITH DISABILITIES: INDIVIDUALIZATION AND INTEGRATION (2017).

5. Would Themis undermine 8th Amendment challenges to solitary confinement?

Some advocates for ending solitary confinement are critical of efforts to improve solitary confinement conditions. Chief amongst these criticisms is the argument that by modifying the conditions of solitary, it only serves to prolong and validate the practice. Such logic could apply to the introduction of Themis. If, for instance, the detrimental effects of solitary confinement are (in part) mitigated by the introduction of confinement companions, it could be harder to argue that solitary confinement is cruel and unusual under the 8th Amendment.

There are two responses to this criticism. First, like others who have proposed reforms to current practice, I believe that one can advocate for an end to solitary confinement (which I would prefer), and for its second-best alternative: minimizing the harm while it still exists.²³⁰ Second, as discussed above in Part III, Themis offer advocates an intriguing opportunity for discovery of the true conditions of solitary confinement. A significant challenge for advocates of ending solitary confinement is proffering sufficient evidence to show the harsh realities of isolation. “Media access to prisoners, particularly those in solitary, is limited or non-existent, and many states do not provide adequate data on how their penal systems actually operate.”²³¹ Themis may help to provide access,

²³⁰ Margo Schlanger & Amy Fetting, EIGHT PRINCIPLES FOR REFORMING SOLITARY CONFINEMENT, THE AMERICAN PROSPECT (Oct. 6, 2015), <http://prospect.org/article/eight-principles-reforming-solitary-confinement-0> (“Many of the advocates who have fought for these reforms over the past decade want to end solitary confinement entirely, not merely reduce its use and make it less harsh. We agree with that ultimate objective, but we see the reduction and amelioration of solitary as necessary steps to its eventual elimination. Even those who oppose ending solitary may agree about ending the most brutal and inhumane aspects of the practice.”)

²³¹ Jules Lobel, *The Liman Report and Alternatives to Prolonged Solitary Confinement*, 125 YALE L.J. FORUM 238 (2016), citing Andrea C. Armstrong, *No Prisoner Left Behind? Enhancing Public Transparency of Penal Institutions*, 25 STAN. L. & POL’Y REV. 435, 462 (2014); David C. Fathi, *The Challenge of Prison Oversight*, 47 AM. CRIM. L. REV. 1453, 1453-54 (2010). See also Kupers (2017), *supra* note 6 at 34. (“One reason that solitary confinement has been allowed to proliferate so much in the modern era is that the general public has very little knowledge about what is occurring inside the prisons.”)

and might actually improve the success of 8th Amendment challenges.

6. Moral Hazard?

What if an inmate became so engaged by Themis that the inmate actually *wanted* to remain in solitary confinement in order to engage with Themis? Is such a perverse incentive possible? Maybe. If Themis is made available only to those in solitary confinement, and if Themis learns to be an excellent companion, perhaps for some inmates Themis will be their best “friend”. Further research and discussion is warranted on this possibility.

7. Protecting Privacy with Privilege?

In section A.1 above, I considered whether a technical solution could help Themis ensure the privacy of inmate conversations. But it may be more likely that a legal solution—privilege—will be most effective. For more than four centuries, English and now American courts have recognized an attorney-client privilege.²³² For roughly 200 years, courts have recognized some form of physician-patient privilege.²³³ And courts too have recognized a limited privilege between a priest and a confessor.²³⁴

The problem for Themis is that, under current law, we would have to wait until version 5.0—when Themis can be held out as an attorney or psychiatrist—for conversations with Themis to be privileged. Unless we change the law. Is it possible to carve out a new type of privilege for a new type of relationship? Maybe. Calo has very usefully pointed out the “difficulty of placing robots in one category or another, and our tendency in general to behave around

²³² Christopher T. Hines, *Returning to First Principles of Privilege Law: Focusing on the Facts in Internal Corporate Investigations*, 60 U. KAN. L. REV. 33, 41 (2011) (“the antecedents of the attorney-client privilege stretch for more than 400 years of Anglo-American legal history”).

²³³ Daniel W. Shuman, *The Origins Of The Physician-Patient Privilege And Professional Secret*, 39 SW. L. J. 661 (1985).

²³⁴ Jacob M. Yellin, *The History and Current Status of the Clergy-Penitent Privilege*, 23 SANTA CLARA L. REV. 95 (1983).

social technology as though it were a person”²³⁵ Themis is not a human. Themis, as I envision it, is also not simply a tape recorder. Themis is something in-between. Might we be able to justify a privilege for this in-between category?

[More discussion required here to justify a potential privilege on privacy grounds.]

8. Shouldn't This Technology Be Developed for *All* Inmates?

This paper has argued for the development of confinement companions for those in solitary confinement. But why not make the technology more widely available to the full prison population? After all, there are high overall rates of mental illness in the incarcerated population.²³⁶ Moreover, mental health services provided by prisons are inadequate, with some researchers estimating that prisons only have the capability to treat between 10% and 12% of the prison population.²³⁷

With so many in need of improve attention to their mental health, what is the justification for focusing only on those in solitary confinement? The answer is two-fold. The practical response is that reform has to start somewhere, and it makes sense to start where the conditions are most deplorable. The substantive response is that in those other contexts, human contact is available in ways that it is not for those in solitary confinement. That is, the added value of a robot like Themis would have to be weighed against the added value of another human. For those in solitary, there will be no new human added to the mix; it's Themis or nothing.

9. The New “Eye of God”?

²³⁵ Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CAL. L. REV. 513, 547 (2015)

²³⁶ DORIS J. JAMES & LAUREN E. GLAZE, U.S. DEPARTMENT OF JUSTICE, OFFICE OF JUSTICE PROGRAMS, BUREAU OF JUSTICE STATISTICS, MENTAL HEALTH PROBLEMS OF PRISON AND JAIL INMATES (2006). Henry J. Steadman et al., *Prevalence of Serious Mental Illness Among Jail Inmates*, 60 PSYCHIATRIC SERVICES 761 (2009).

²³⁷ Terry A. Kupers, *What to Do with The Survivors? Coping with The Long-Term Effects Of Isolated Confinement*, 35 CRIMINAL JUSTICE AND BEHAVIOR 1005 (2008).

A feature of early solitary confinement in the United States was an opening at the top of the prison cell called the “Eye of God.”²³⁸ This daylight opening was intended to allow the inmate to communicate with God in order to reform his soul (and consequently his behavior).²³⁹ Promoted by the Quakers in Pennsylvania, 1790, a law was passed to encourage “unremitted solitude” and “prevent all external communication.”²⁴⁰ This Philadelphia system drew much attention as it was piloted in the Eastern State Penitentiary in 1821.²⁴¹

When George Smith wrote his 1833 defense of solitary confinement for the state of Pennsylvania, he opened his essay by recognizing that “The prevention of crimes and the reformation of criminal in lieu of the vindictive infliction of pain on offenders, are now almost universally acknowledged to be the only legitimate designs which can justify the infliction of human punishment.”²⁴² At Eastern State Penitentiary, the Quakers aimed for this design to be “monastic”, and an “atmosphere of silence, solitude, meditation, and complete isolation” was carefully considered during the construction of the prison, which was considered to be “technologically far ahead of its time.”²⁴³



Designed by architect John Haviland, the exterior of the Penitentiary exhibited a gothic style, whose purpose was to scare and “dissuade free citizen from committing crime.”²⁴⁴ In contrast, the cells themselves were bare of any furniture or decoration, with

²³⁸ Muriel Schmid, “The Eye of God” *Religious Beliefs and Punishment in Early Nineteenth-Century Prison Reform*, 59 *THEOLOGY TODAY* 546 (2003).

²³⁹ *Id.*

²⁴⁰ Friedman, *supra* note 2, at 78.

²⁴¹ *Id.*

²⁴² GEORGE WASHINGTON SMITH, *A DEFENSE OF THE SYSTEM OF SOLITARY CONFINEMENT OF PRISONERS* 1 (1833).

²⁴³ *Id.*

²⁴⁴ Muriel Schmid, “The Eye of God” *Religious Beliefs and Punishment in Early Nineteenth-Century Prison Reform*, 59 *THEOLOGY TODAY* 546 (2003).

the exception of a single skylight at the top of the cell coined “the eye of God”.²⁴⁵ This architectural feature served two purposes: to provide natural light to the cell, and “as a reminder to prisoners that the eye of God is watching over them”, which highlighted the Quaker belief that “total solitude before God was supposed to effect a conversion of the criminal’s moral sensibilities.”²⁴⁶ This feature was also supposed to promote a sense of “totalizing surveillance” in the prisoners, from both within the cell through the “eye of God” and outside the cell from the prison guards.²⁴⁷

The “eye of God” feature also reflected the Protestant belief that an “individual encounter” between the criminal and God was necessary²⁴⁸, and that solitary confinement should aim to promote this encounter in an effort to rehabilitate criminals.²⁴⁹ It was the “therapeutic environment” of the Penitentiary that separated it from other prisons, which emphasized punishment.²⁵⁰ Later on, the Bible would be the first book to be provided to these inmates, who were encouraged, under the ‘eye of God’, to “read the Scriptures and find their own salvation.”²⁵¹

I provide this background on the Eye of God because it sets up an interesting parallel: would Themis be the new eye of God? In a culture that is (slowly) moving away from organized religion, and at least a little moving toward the worshipping of technology, could Themis take on some sort of deity role? And if so, would *religious* free exercise protections then be applicable?

²⁴⁵ *Id.*

²⁴⁶ *Id.*

²⁴⁷ Jill A. McCorkel, *Embodied Surveillance and the Gendering Of Punishment*, 32 JOURNAL OF CONTEMPORARY ETHNOGRAPHY 41 (2003).

²⁴⁸ Schmid, *supra* note 244.

²⁴⁹ McCorkel, *supra* note 247.

²⁵⁰ *Id.*

²⁵¹ Schmid, *supra* note 244. Caleb Smith, *Detention Without Subjects: Prisons and the Poetics Of Living Death*, 50 TEXAS STUDIES IN LITERATURE AND LANGUAGE 243 (2008). This design was also meant to instill “the metaphor of the cell as a grave”, in which criminals were viewed as “sinners who had fallen away from God and who needed to have their lives reconciled to God”. It was this aspect of confinement that would later cause Charles Dickens to testify against solitary confinement at the Eastern State Penitentiary, describing the prison as a “parade of ghostly figures” and inmates who looked as if they had been “summoned from the grave”. Using these terms, Dickens implied that “the prisoner, in this system, is somehow dead: ‘He is a man buried alive’”.

It is well established that “religious practices of prisoners, as distinguished from their beliefs, may properly be the subject of administrative regulation and control, so long as particular religious groups are not improperly discriminated against and so long as the action taken by the prison authorities is not arbitrary or unreasonable.”²⁵² How would this apply to Themis if, as speculated, the inmate’s relationship to Themis took on a more religious dimension?

C. Practical Considerations

1. How Will Themis Be Funded?

Themis will not be cheap to develop, and to attract investments there must be a viable path to market adoption. What prison is going to adopt a technology that seems to subvert its rationales? The most likely answer, I think, is a prison that is forced to adopt Themis by a judge. Themis could find its *entre* into the justice system through a settlement or court order. There is precedent for something similar.

In New York City, a settlement related to excessive force by prison guards included a requirement that the prison install “hundreds of new wall-mounted video cameras with recording capability--in addition to the 2,000 cameras already in place,” with the thought that this would better document (and thus reduce) violence by guards on inmates.²⁵³ The rationale for the monitoring was described:

U.S. Department of Justice Inspector General Glenn Fine stressed the value of cameras for prosecutors: “With video surveillance you often can see what happened before or after an incident, so that’s very important, and we have relied upon that kind of evidence very strongly.” These visual and auditory records protect prisoners and staff from violence and from false allegations

²⁵² 12 A.L.R.3d 1276 (Originally published in 1967); “Prison officials may not substantially burden an inmate’s religious practice, or selectively discriminate against an inmate based on religion, without legitimate penological justification. *Williams v. King*, 56 F. Supp. 3d 308 (S.D. N.Y. 2014).

²⁵³ John J. Gibbons & Nicholas De B. Katzenbach, *Confronting Confinement A Report of the Commission on Safety and Abuse in America’s Prisons*, 22 WASH. U. J.L. & POL’Y 385, 434–35 (2006).

of misconduct. Leslie Walker, executive director of Massachusetts Correctional Legal Services, believes that cameras can even discourage the “tiny, degrading, everyday humiliating name calling that can occur.” This behavior, she said, will not be reported with any regularity or believed unless it is “*seen and heard*.”²⁵⁴

The same logic applies to the Confinement Companion. Themis can hear those whose voices have been ignored.

The availability of Themis could also affect future Supreme Court jurisprudence. Justice Kennedy seems to have invited the next case when he commented in dicta in 2015 that: “In a case that presented the issue, the judiciary may be required . . . to determine whether workable alternative systems for long-term confinement exist, and . . . whether a correctional system should be required to adopt them.”²⁵⁵ One workable alternative may be the use of a Themis system to lessen the impact of isolation.

2. Next Steps: Where to Start?

This is the first proposal for utilizing AI in the context of solitary confinement. Like most beta proposals, it is need of further scrutiny and revision. But what are the next steps after that? Most promising, I think, is the advent of a public-private partnership to develop a proof of concept Themis system. Experimental work with non-incarcerated subjects could serve as proof of concept, and a business plan could be developed to attract investors.

Also critical will be buy-in from state and/or federal governments. As Calo has pointed out, “[t]he government possesses a wide variety of means by which to channel AI in the public good.”²⁵⁶ In particular, “policymakers at all levels ought to be thinking about the qualities and characteristics of the AI-enabled products government will purchase and the companies that create them.”²⁵⁷ Public pressure at the state level could force a prison system to adopt Themis, at least on a trial basis. Government and private funding would be required to conduct the rigorous testing needed to prove proof-of-concept and then efficacy.

²⁵⁴ *Id.*, emphasis added.

²⁵⁵ *Davis v. Ayala* (2015), 36.

²⁵⁶ Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. DAVIS L. REV. 399, 429 (2017)

²⁵⁷ *Id.*

V. CONCLUSION

A constant feature of American prisons for over three centuries has been the use of solitary confinement.²⁵⁸ Yet the purpose and nature of solitary confinement has changed over time. As technology has changed, so too have the physical spaces used to isolate.

AI has the potential to usher in the next wave of change. It will not be easy to make a dent in the ugly, punitive, soul-crushing system of Supermax prisons. But it is possible. It was not so long ago that the self-driving car was a science fiction fantasy; today, cities are preparing for a future without human drivers. If we can put 10 million self-driving cars on the road by 2020, can't we install a few thousand speakers and microphones in prison cells? And if we already have a neural network with 1 billion connections, shouldn't we be able to process a prisoner's simple request to talk about the weather?

It is true that the prison-industrial complex is a major barrier to the development of a confinement companion like Themis. But is the AI community really going to be out-smarted by prison guards? I don't think so.

The real challenge, in my view, is for law to show the AI community how much harm solitary confinement is causing—and how much good confinement companions can do in this space.

²⁵⁸ Harry Elmer Barnes, *Historical Origin of the Prison System in America*, 12 J. AM. INST. CRIM. L. & CRIMINOLOGY 35 (1921) (discussing how prison designs in the late 18th and early 19th century “definitely stipulated that ... [they] should be constructed according to the principle of solitary confinement ...”).